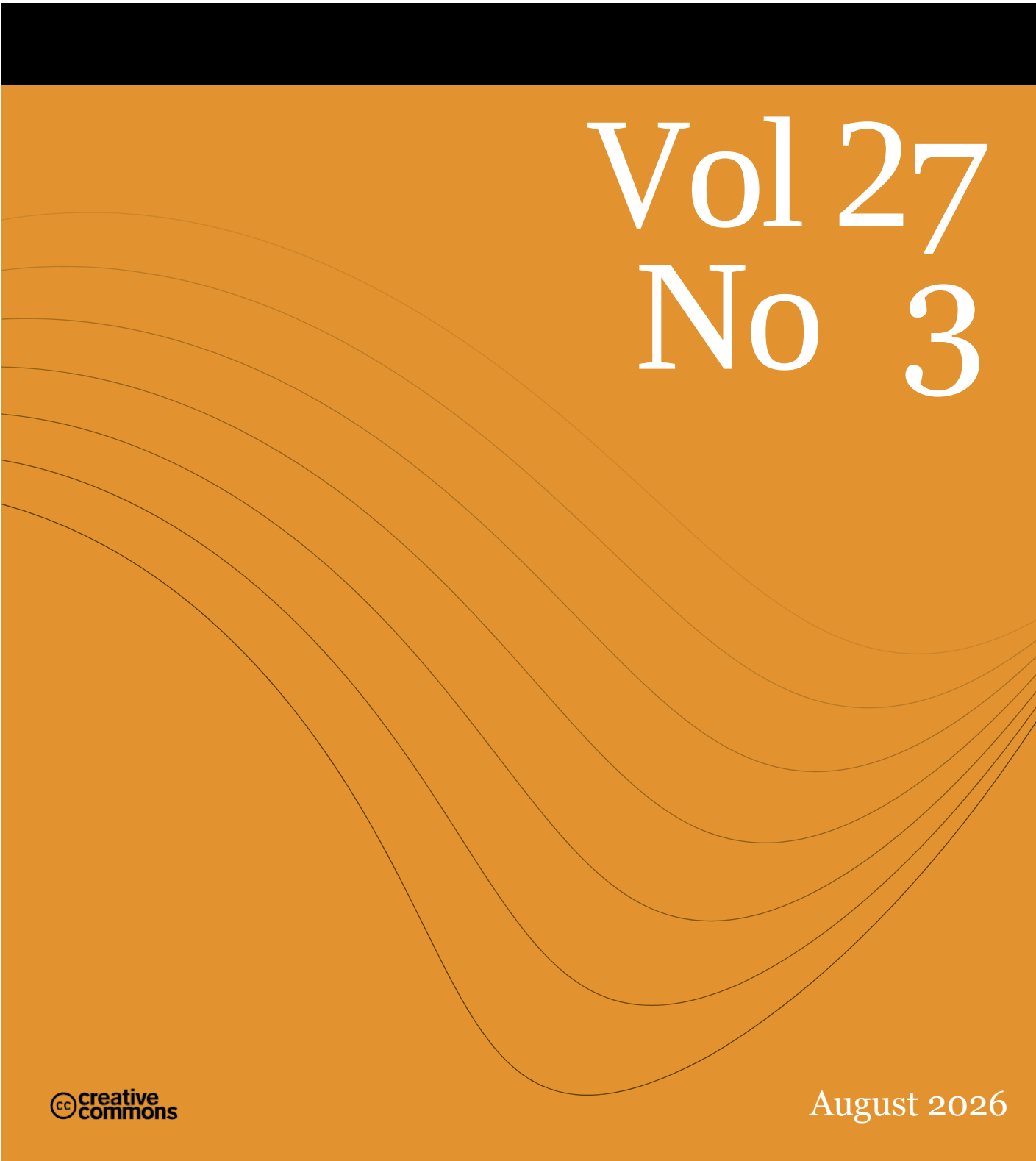




IRRODL

THE INTERNATIONAL REVIEW
OF RESEARCH IN OPEN AND
DISTRIBUTED LEARNING



Vol 27
No 3

August – 2026

Editorial – Volume 27, Issue 3

Adnan Qayyum

Associate Editor, Athabasca University

Bicycles of the Mind: Technologies for Empowering People

Steve Jobs used to call the personal computer “a bicycle of the mind.” Humans, on their own, are quite inefficient at moving when measured by distance travelled per unit of energy spent. But a person on a bike can move more efficiently than any bird, animal, or fish. This machine combined with human brains and power vastly enhanced human abilities. Jobs’ hope was that a computer would have the same transformative effect for humans. In automation theory, a person on a bike is a *centaur*: a person who controls a technology to enhance their abilities. There are many everyday examples of centaurs, including spell check in word processors or auto-brakes in cars. But there is also a *reverse centaur*: the technology directs the person, making the person an appendage to the machine. An Uber driver with a checklist of pickup locations on their Uber app is a reverse centaur. They are physical operators that have to surrender navigation decisions to the machine (which are just manifestations of employer policies).

Cory Doctorow underscores this difference between a centaur and reverse centaur to argue that we should not focus just on what a technology does, but who it does it for and who it does it to. The central question should be: how does the technology affect people, and their agency? It can be easy to lose perspective on how to analyze any technologies in education and other fields, especially amid the tsunami of opinions and research on artificial intelligence and the breakneck pace of change of new frontier AI models. But coming back to focus on people, how they are affected, their agency and opportunities, that is what matters.

The AI Discussion

In this issue, we have two research articles and three special reports that discuss AI. In “Artificial intelligence in online learning: Using BERTopic to track research topics and their evolutions,” **Chen et al.** used a topic modelling technique to analyze over 1000 publications on AI in online learning. They identified the major research topics in the field of AI in online learning, and the evolution of this research. Early research on AI in online learning focused on topics such as AI-automated grading, AI-generated feedback, and adaptive testing. More recent research is on machine learning methods and deep learning approaches that can act as platforms for applications to understand learner behaviours, forecast outcomes, and improve educational decision-making. They propose a conceptual model to guide AI research and practice in online learning, by integrating four components: data, AI processing, adaptive learning, and learner development.

In “The 20-year evolution of distance education: A bibliometric and LDA-based topic model analysis,” **Yavuz, Balat, Kayali, and Gulen** studied research trends in distance education. They used a machine learning method, Latent Dirichlet Allocation (LDA), to do a bibliometric analysis over 54,000 distance

education articles. They found DE grew substantially, post-pandemic, in the fields of health, management, and education. Among 19 topics of research, the most prominent were digitalization, artificial intelligence, Web-based learning, and social interaction and collaboration. As anticipated, AI and emergency distance education have been more popular topics in recent years. Notably, there has been less interest in program design.

In the first of three reports on AI and scholarship, **O’Driscoll** and **McGreal** ask who is responsible for the content of academic work when AI is used. They identify three distinct roles in the scholarship process: the animator, author/designer, and principal, based on Goffman’s idea of production format. They analyze two major issues about AI-assisted scholarship: the accountability of publishers and platforms, and the apparent displacement of the authorial role by GenAI. Their work is important because it addresses a key concern with the use of AI in education and other fields: accountability.

Bozkurt focuses on how AI affects quality control within academic publishing. He highlights the tension between efficiency gains and the preservation of integrity. AI can reduce publication delays and reviewer burnout. However, concerns continue over confidentiality, accountability, algorithmic bias, and the erosion of expert judgment. Bozkurt argues that AI should serve as an augmentative rather than substitutive technology in peer review.

Downes provides a survey of AI in scholarly publishing. He identifies six ways AI can benefit research functions including: idea development and research design; literature review and synthesis; content development and structuring; data management and analysis; editing, review, and publishing support; and communication, outreach, and ethical compliance. He also identifies important challenges of using AI for research, including: accuracy, research integrity, authorship and ownership, bias, privacy, and consent. Importantly, he discusses that AI may not be technologically ready for the complex tasks involved in academic publishing.

Research Articles

The field of open and distributed learning is about more than AI. It is foremost about access and openness in education. In “Building a framework to guide the design and development of open educational practices for the digital training of university professors,” **Esteve-Mon et al.** are interested in enhancing teacher digital competence, including competence in developing open educational resources. They create a theoretical framework to guide the design of OER-based training to integrate an open educational practice for faculty development. The framework they create focuses on reflection within the teacher’s professional development and practice.

Sing and colleagues also research open educational resources (OER), exploring how students could improve their understanding of OER by creating them. In “From consumers to creators: How student participation in OER development transforms understanding of open licensing and improves digital literacy,” they survey 57 students. They found significant improvements in how students understood open licensing, particularly Creative Commons licenses, their perceptions of how credible OER were, and their behaviour accessing different types of OER materials. By creating OER, students had more favourable attitudes and a deeper understanding of OER.

In “Between obligation and aspiration: Motivation, emotions, and systemic barriers in online professional development for rural teachers,” **Hua, Yang, and Chan** interviewed 30 rural primary and secondary school teachers in China to understand how they experience online professional development. The teachers had a positive attitude toward online professional development, but they had unsatisfying learning experiences, despite their initial optimism. The researchers found a mismatch between teachers' favorable attitudes and unsatisfying learning encounters. This may have been due to required participation, lack of internal agency in developing learning goals that informed the training design. The researchers propose three conceptual constructs to explain the complex interplay between emotional readiness and systemic limitations.

Rural education is also the focus of **Mulyani's** research, “Enrollment decline and the digital divide in rural elementary schools: Implications for equity in basic education.” He interviewed principals, teachers, parents and community leaders and observed classrooms and analyze documents to determine why there has been declining enrollments. He found several important reasons, including declining birth rates, outward migration, structural disadvantages such as remote locations and limited infrastructure, and parental preferences for private or religious schools. The decline is not just demographic but also due to systemic digital inequity.

In “Structure and content of interactions: Social presence in asynchronous online whole-class vs. small-group discussions,” **Willet, Baik, Hur and Akcaoglu** investigated how the format of online class discussions affected the structure and content of student interactions. Online classes with the whole class had greater variability in response rates, were more formal, and the communication was more analytical. The small group discussions were more supportive, and students had more social and emotional engagement. The research suggests that online discussion formats should be designed to foster specific learning objectives.

Tomak and Virvan also focused on student experiences in their research on “Blended learning in a post-pandemic context: A cross-sectional study on Turkish EFL learners' perspectives.” They observed classrooms and interviewed 24 students to understand their perspectives of blended learning. Students had ongoing concerns about engagement, socialization, the assessment integrity of online education. Students were more accepting of blended learning since their pandemic experiences. For speaking English and group work, students still preferred in-person instruction.

In “Does interaction quality enhance motivation among online postsecondary students? Exploring the mediating role of anxiety,” **Barri** studies how anxiety affects interaction and motivation among online students. While interaction and motivation are well-studied, there is less research about how anxiety affects each. Barri uses structural equation modelling to analyze a survey of 140 students. The findings showed that anxiety fully mediated the relationships between student-content and student-instructor interactions and motivation, and partially mediated the relationship between student-student interaction and motivation. The results indicate that reducing anxiety and supporting motivation are important to consider for online instructors.

In “An exploratory factor analysis of the online information literacy self-efficacy (OILS) instrument,” **Zarkhov, Traynor, and Maybee** create a self-efficacy instrument for online information literacy. They

analyze a 30-item survey of 259 students. Using EFA, they found four factors that explain substantial variance in the response patterns about online information literacy self-efficacy: scoping research topics, obtaining information, evaluating information quality and producing research documents, and crediting sources. The instrument measures students' confidence in applying information literacy skills in online settings.

Avci and **Güven**, in “Systemic vulnerabilities in online education: Diagnosing and meeting K–12 teacher preparation needs,” created and evaluated a 5-week in service teacher training program that enhanced teacher's online education competencies. They did a needs assessment of 21 teachers, as well as learning analytics and pre/post test assessments. The results suggest that teachers benefit from being online students before they become effective online instructors.

In “*Predicting student outcomes in open high schools using educational data mining*,” **Polat** and **Horzum** used data mining to analyze the academic trajectory of 484,000 students enrolled in open high schools. They found that academic and administrative features of open high schools, particularly total credits accumulated and initial enrollment type, are more important predictors of completion or dropout than demographic variables. This suggests that retention policies should identify at-risk students using predictors such as low-credit accumulation rates and high-risk enrollment types.

Book Notes and Literature Reviews

In our book notes, **Xiao** reviews Mark Nichols open access book on *Mega-Universities and Opening Education*. The book is meant to be a revisit of John Daniel's 1996 book *Mega-Universities and Knowledge Media*, and includes a foreword by Daniel. Xiao identifies that the 14-chapter book explores four facets of openness in higher education: availability, inclusiveness (from the students' perspective), scalability, and sustainability. Xiao notes that interesting case studies in the book include the Open University of China and Southern New Hampshire University in the United States. Nichols provides five strategic considerations for mega-universities and those seeking mega-university status. Xiao found the book is optimistic about using digital technologies, including Gen AI, to expand higher education at scale. He finds the book to be an inspiring and valuable reference.

We have two literature reviews, one on clinical learning technologies and the other on the community of inquiry framework. In “Novel technologies in nurse practitioner education: An integrative review,” **Girard** and her colleagues conducted a literature review about the use of these clinical learning technologies in nurse practitioner education. These technologies include high-fidelity simulation, immersive virtual reality, augmented reality, and holographic simulations to support learning and improve retention. From the 11 studies that met the inclusion and exclusion criteria, five primary themes emerged: student uptake, student learning, change management, accessibility, and confidence. From their research, they suggest that integrating advanced clinical learning technologies into nurse practitioner programs should be a priority.

In “A systematic review of using social and epistemic network analysis methods in community of inquiry research,” **Don**, **Lowell**, and **Hsieh** synthesize findings from 43 empirical studies that focus on applying social and epistemic network analysis methods within Community of Inquiry research. Focusing on these two methods, they found them to be valuable for investigating interaction patterns, refining CoI theory,

examining community formation and evolution, analyzing discourse and knowledge construction, and tracing the development of CoI presences. The study will be of interest to CoI researchers and designers of online learning communities.



August – 2026

Determining Accountability in AI-Assisted Scholarship: A Report and Recommendations

Jim O'Driscoll¹ and Rory McGreal²

¹Independent Scholar; ²UNESCO Chair in OER, Athabasca University, Canada

Abstract

The proliferation of generative artificial intelligence (GenAI) presents a significant challenge to traditional notions of accountability in scholarly publishing. This report addresses the pressing question of who is responsible for the content of academic work when its creation involves AI assistance. It does so by introducing and extending Erving Goffman's (1981a, b) concept of *production format*, a framework that deconstructs the singular speaker into three distinct roles: *animator*, *author* (herein re-labelled *designer*), and *principal*. By applying this framework to written, scholarly communication, this report analyzes two key problems: the accountability of publishers and platforms as relaying animators and the apparent displacement of the authorial role by GenAI. The analysis draws on recent empirical evidence of journal policy adoption and theoretical applications of Goffman's work to the AI context. It pays particular attention to the implications for open access publishing, where the democratization of knowledge creation must be balanced with the imperative for rigorous accountability. The report concludes that while GenAI can assume a large proportion of the designer function of shaping textual form, the ultimate accountability for scholarly work must remain with a human principal. Detailed recommendations are provided for scholars, publishers (with specific focus on open access journals), and academic institutions to navigate the new landscape of AI-assisted authorship, ensuring transparency and preserving the integrity of academic discourse.

Keywords: generative AI, GenAI, accountability, authorship, scholarly publishing, Goffman, production format, open access

Introduction

The advent of generative artificial intelligence (GenAI) has precipitated a crisis of accountability across numerous domains, none more so than in scholarly publishing. The fundamental academic question, Who said that? has become profoundly complex. For reasons of credit, copyright, legal culpability, and the very pursuit of verifiable knowledge, it is crucial to know who is responsible for a claim and whom to approach for clarification or evidence. Traditionally, the named author(s) beneath a title served this function, signifying the sole linguistic agent responsible for the text's creation and content. GenAI, however, has fundamentally disrupted this long-standing assumption. This report explores this disruption by applying and extending the concept of *production format*, first advanced by the sociologist Erving Goffman, to offer a framework for understanding and navigating accountability in AI-assisted academic work, with particular attention to the context of open access scholarly publishing.

Production Format: A Deconstructed Framework for Authorship

Nearly half a century ago, Goffman (1981a, pp. 144–155; 1981b, pp. 226–229) wrote two papers in which he deconstructed the unitary notion of the speaker, a concept long taken for granted both in communication studies (e.g. Shannon & Weaver, 1949) and in the burgeoning field of speech act theory (e.g. Searle, 1969). In these, he proposed a production format comprising three analytically distinct roles involved in the production of any utterance:

- *Animator*: The individual who physically produces the utterance through the activation of vocal equipment.
- *Author*: The agent who composes the words and selects the syntactic and semantic structure of the utterance.
- *Principal*: The party whose position, beliefs, or stance is represented by the utterance and who is held accountable for it.

In informal conversation, these three roles typically reside in a single person. However, Goffman points out that in other kinds of language production, they are frequently separated. Examples come easily to mind. A politician (principal) delivering a speech may employ a speechwriter (author), while the act of speaking (animation) is their own; a lawyer reading aloud a statement to the press animates words they may have authored, but the position is that of their client (principal).

Goffman's framework was devised for spoken interaction. However, it is readily applicable to written communication. The same concerns of production and reception pertain, with initial animation achieved through the use of a pen or keyboard. To fully address the complexities of scholarly publishing in the age of GenAI, however, the framework requires one key extension and one terminological adjustment.

The extension (only hinted at by Goffman) is to allow for each role to be shared by multiple agents. His authorial role, for instance, is often a collaborative process between a writer and an editor, or, as this report will explore, between a human and an AI platform. The animator role in written communication is almost always shared, requiring the services of a courier—a postal service, a publisher, or an Internet service

provider to make the text accessible to its audience. Even the principal role can be diffused, as when a scholar presents and therefore takes responsibility for another theorist's ideas, or when a politician's utterances are understood to represent a political party.

The terminological adjustment concerns Goffman's author. In common parlance – and indeed the rest of this report – this word denotes not only his author but also his principal. This conflation creates confusion when his framework is applied to the written mode, especially when applied to GenAI, which can perform a compositional role. This report therefore proposes replacing Goffman's author with the term *designer*. This term more precisely captures the function of crafting the textual artifact, determining how to say something without implying the ownership of the underlying ideas, which is the province of the principal. The principal has something to say (meanings), the designer works out how to say it (form), and the animator makes the chosen words manifest. This distinction is crucial as Varoufakis (2023) highlights in his discussion of the ancient Greek principle of *ισεγορία* (isegoria), which not only upholds the right to have one's views considered on their merits regardless of who you are but also separates the substance of an argument from its rhetorical flourish or mode of expression (a task that can be legitimately shared with AI).

A further strand of Goffman's work is relevant here. Recent scholarship has begun to apply his dramaturgical sociology (1959) to the AI context in ways that support and enrich this extended framework. Klowait and Erofeeva (2025) argue that AI in the workplace should be viewed as supplementary rather than substitutive automation. Drawing on Goffman, they reframe AI as a “prop within an ongoing social performance—one whose scripts, stages, and audiences remain irreducibly human.” In their analysis, large language models (LLM) like ChatGPT quietly take over the invisible, routinized tasks that underpin cognitive occupations, namely editing, summarizing, and first-draft production, while leaving humans to enact the highly visible or relational facets that sustain professional prestige. This perspective aligns entirely with the proposed distinction between the AI as a designer (handling routinized textual formulation) and the human as the principal (enacting professional competence and ultimate accountability). The human's role shifts from being the sole designer to being the manager and approver of the design process, a role that requires a new form of AI literacy to ensure the final output aligns with their intellectual stance.

Problem 1: The Animator as Additional Principal in the Open Access Context

A central problem for scholarly publishing concerns the accountability of what we may term relaying animators, namely the publishers, platforms, and Internet service providers that facilitate the dissemination of a text. While these entities often position themselves as neutral messengers, their editorial power and gatekeeping function suggest they also partake of the principal role. This issue is particularly acute for open access (OA) journals, which often operate with fewer resources and may be more vulnerable to both the risks and rewards of AI integration.

A 2025 incident involving the BBC and historian Rutger Bregman's Reith Lectures illustrates this quandary. In one of these commissioned series of lectures, delivered live to an audience, Bregman described a political figure as “the most openly corrupt president in American history.” But when the recording of this lecture was broadcast, this phrase was removed by the BBC on legal advice (Yahoo, 2025). In the live context, Bregman was clearly the sole principal. In the broadcast version, however, the BBC's editorial intervention

made it a shared principal, accountable for the final content its audiences received. This raises a critical question: to what extent does editorial control confer accountability?

The editorial power of publishers, from major academic houses to niche journals, makes them de facto additional principals. For example, an editor's discomfort with an author's opinion can effect real changes in the content of an article. This reflects a sense of ownership over the publication's entire content. This is not merely a theoretical concern. Legal history, such as the mid-20th century obscenity trial of *Lady Chatterley's Lover* being brought not against its author but rather its publisher Penguin Books, demonstrated that society often holds these relaying animators accountable. In the scholarly context, the reputation of a journal or press is inextricably linked to the works it publishes.

However, granting relaying animators full principal status is dangerous. If powerful publishers and platforms are held legally accountable for every idea they disseminate, a chilling effect on academic freedom and diversity of thought becomes inevitable. This risk is amplified in an era of LLMs, which are trained on existing corpora and tend to reinforce dominant ideologies. The solution may be to recognize them as additional principals, with a tiered accountability: the primary accountability for the truthfulness and originality of claims rests with the human author, while the publisher is accountable for the procedural integrity of the publication—the rigour of peer review, the transparency of editorial processes, and the adherence to ethical guidelines. This distinction is vital for OA journals, which are committed to disseminating knowledge broadly and must safeguard against any perception that open access equates to lower standards.

Recent empirical research has underscored the urgent need for publishers to formalize their role as additional principals through clear policies. In a cross-sectional review of vascular surgery journals, Shen et al. (2025) found that while 72% had explicit GenAI policies, significant inconsistencies remained. All journals prohibited AI authorship and affirmed human accountability, a non-negotiable baseline. However, only 56% specified where disclosure of AI use should be placed, and only 52% addressed AI-generated images. Crucially, higher-impact journals (Q1/Q2 in SCImago Journal Rank) demonstrated more comprehensive policies than lower-tier journals. This disparity is a major concern for the OA landscape, which includes many new, lower-profile and scholar-led journals that may lack the editorial infrastructure to develop robust policies. Without clear guidance from these additional principals, the integrity of the scholarship they publish, and by extension the reputation of the OA movement itself, is put at risk.

AOSIS (African Online Scientific Information Systems) has provided a model response. In late 2025, it made the inclusion of AI policies compulsory for all journals in its indexing service (SciELO SA), reminding authors to disclose their use of AI-assisted technology to ensure transparency, accuracy, and trust. Furthermore, AOSIS announced its transition to the CRediT (Contributor Roles Taxonomy) authorship statement, a move that dramatically increased transparency by specifying each author's contribution to a work. This taxonomy, which includes categories like 'conceptualization,' 'methodology,' and 'writing—original draft,' operationalizes the deconstruction of the author role. It provides a practical tool for distinguishing between the human principal's contributions and tasks that may have been assisted by AI, thereby reinforcing accountability.

Problem 2: The Author (Designer) as Principal in the Age of GenAI

For centuries, copyright law and academic convention have conflated the roles of designer and principal under the single term author. This conflation is now becoming untenable. With appropriate prompting, LLM can generate full texts, effectively assuming the designer role. This is a core disruption of GenAI. The question then becomes: Who is the principal? If the author in the conventional sense no longer composes the text, can they still be held accountable for its claims?

The use of AI in this context has sparked widespread misgivings about the death of creativity and thought in academic work. However, as Gunkel (2025) argues, this perspective may be misguided, the development having one notably positive effect. The absence of a singular, intentional authorial voice can liberate readers from the hermeneutic task of excavating a text for the author's 'true meaning', encouraging instead a focus on the text's effects and the reader's own interpretation. This recipient-centered approach aligns with contemporary linguistic pragmatics, which has become increasingly suspicious of the notion of speaker/writer intention (e.g., Terkourafi, 2021, pp. 41–43). It also democratizes critique: chiming well with Varoufakis' account of isegoria, a text can only be judged on its merits, not on the authority of its creator.

Nevertheless, the need for an identifiable principal remains paramount, especially in the academic domain. Unlike folklore or popular songs, where accountability is largely irrelevant, academic work directly informs further research, policy, and real-world actions. A principal is essential for follow-up, clarification, and, most critically, for responsibility if claims are found to be flawed or fraudulent. The case of the retracted measles, mumps, and rubella (MMR) vaccine study, which could have caused real-world harm, has underscored the danger of research without a clear human party to answer for it. And while a student's use of GenAI for writing assignments complicates assessment, to deprive them of available tools would be perverse. The solution is not prohibition, but a reconfiguration of what it means to be an author in the traditional sense.

This reconfiguration demands that the human user of GenAI, by deciding which AI-generated content to incorporate, edit, and ultimately stand behind, retains the principal role. The AI serves as a sophisticated advisor or mentor, whose contributions can be acknowledged but who cannot be held accountable for the final published work. Louvain's (2025) editorial in *Evolution Letters* powerfully reinforces this point. It identifies accountability as the first of three principles (alongside voice and trust) that must be preserved. Louvain is unequivocal that "ultimately, the authors of a paper are accountable for the data, analysis, and manuscript itself. . . . The human writing the prompts and doing the analysis is ultimately responsible for checking that anything they implement from LLMs is appropriate and correct". He even models this transparency by disclosing his own use of ChatGPT to generate a section of the editorial, demonstrating that AI use can be acknowledged without diminishing authorial accountability.

Louvain (2025) also introduces the critical concept of voice, arguing that we cannot allow LLMs to override our own voice in manuscripts and reviews. This is the practical manifestation of the principal-designer distinction. The AI (designer) may suggest a turn of phrase, but the human (principal) must engage critically with that output, revising it to align with their own argument, priorities, and authentic scholarly identity. This process ensures that the final text is not merely a competent piece of writing, but a genuine representation of the principal's intellectual stance. And, in an era where everyone may have access to the

same AI tools, the ability to cultivate and assert a unique scholarly voice may become more, not less, important. For this reason, it is important that the human writer not only retains the principal role but also shares (with AI) the designer role.

The Special Imperative for Open Access Journals

The framework of production format has distinct and profound implications for open access publishing. The OA movement is founded on the principles of accessibility, equity, and the democratization of knowledge. The integration of GenAI, viewed through this framework, presents both a powerful opportunity and a significant threat to these principles.

There is an opportunity for enhancing democratic participation when the AI acts as a designer. This can lower barriers for scholars who are not fluent in written English or who are from under-resourced institutions. As noted by Louvain (2025), LLMs are “remarkably effective at helping reduce the cognitive burden of writing in a language with which they are less comfortable”. By assisting with the formulation of ideas into publishable prose, AI can help a more diverse range of scholars assume the role of principal, contributing their unique perspectives to the global academic conversation. This affordance aligns perfectly with the OA goal of amplifying marginalized voices. Furthermore, some envision AI transforming journals from metrics-driven repositories into vibrant knowledge communities that facilitate meaningful discourse and collaborative learning, further enhancing the accessibility and utility of OA content.

However, there is a real possibility that GenAI can be used to undermine trust and equity. The same lack of resources that makes OA journals attractive to scholars from diverse backgrounds can also make them more vulnerable to the risks of GenAI. The above-cited study of vascular surgery journals found that lower-impact journals (which often overlap with smaller, newer OA journals) had significantly less comprehensive AI policies. This policy gap creates a dangerous vulnerability. Such journals may become targets for ‘paper mills’ that use AI to generate large volumes of low-quality, nonsensical, or even plagiarized content—so-called AI slop. This influx threatens to overwhelm editorial systems, devalue a journal’s brand, and erode the trust of readers, who are the ultimate beneficiaries of OA. If open access journals become perceived as a haven for low-quality, AI-generated work, the entire OA movement suffers a reputational blow. The trust that is the foundation of scholarly communication is shattered when readers cannot be sure if a human principal stands behind the claims being made.

Therefore, for OA journals, adopting and enforcing clear GenAI policies is not merely an ethical nicety; it is an existential imperative. They must embrace their role as additional principals with vigour. This means developing policies that are as robust as those of top-tier journals, even with limited resources. It means providing clear, simple guidelines for authors on disclosure, perhaps by adopting a version of the CRediT taxonomy to broadly separate human and AI contributions, without imposing unnecessary tracking that could make the process more onerous than is necessary. It also means investing in editor and reviewer training to spot the hallmarks of uncritical AI use. By doing so, OA journals can protect their integrity and demonstrate that accessibility and rigorous accountability are not mutually exclusive but are instead mutually reinforcing pillars of a healthy scholarly ecosystem.

Conclusions and Recommendations

Goffman's (1981a, b) production format, extended and adapted here, provides a powerful lens through which to view the challenges GenAI poses to academic accountability. Recent theoretical work has confirmed the value of Goffman's approach, while empirical studies have highlighted the urgent but inconsistent response from publishers. A model for an AI-assisted scholarly work can be visualized as follows.

- **Principal role:** The named human author(s), who determine(s) the core meanings, arguments, and findings to be communicated, and who ultimately stand(s) by the work they disseminate.
- **Designer role:** A collaboration between the human author(s) and GenAI, working together to determine the optimal form, structure, and wording for the argument. The human maintains their voice and critically evaluates all output.
- **Animator role:** A chain including (a) the human author(s), typing and issuing the text; (b) the GenAI interface, generating text on screen; (c) the Internet service provider, transmitting the text; and (d) the publisher, formatting and disseminating the final work.

This model is necessarily a simplification; any writer knows that the process of writing down what you think you want to say helps find out what it is that you really want to say. The process of writing is recursive, with design and animation feeding back into the principal's meaning-making. However, it establishes a clear locus of accountability, namely the human principal. Based on this analysis, the following recommendations are made for different stakeholders in the scholarly publishing ecosystem.

Recommendations

For Scholars and Authors

- **Disclose use.** When using GenAI in the design of a scholarly work, its use must be transparently disclosed. A statement in the acknowledgements or a dedicated methods section is appropriate.
- **Assume principalship and guard your voice.** Authors must understand that using GenAI does not transfer accountability. By submitting a work for publication, they affirm their role as the principal, taking full responsibility for its content, including any text or references generated with AI assistance. They must critically engage with AI output to ensure it reflects their own scholarly voice and does not introduce errors or biases.
- **Distinguish idea from form.** Be aware of the distinction between the ideas (principal) and their formulation (designer). The ultimate intellectual contribution, and the responsibility for it, rests with the human author.

For Publishers and Journals (Including Open Access)

- Establish and enforce clear, standardized policies. Develop and enforce clear policies on the use of GenAI. These policies must explicitly state that AI cannot be listed as an author (principal) and must require full disclosure of its use.
- Re-define editorial accountability. Openly acknowledge the role of the publisher as an additional principal, accountable for the overall quality and integrity of what they publish. This accountability includes ensuring that submissions meet ethical and scholarly standards through a robust peer-review process, regardless of whether they are AI-assisted.
- For open access journals specifically, recognize that robust AI policies are essential for maintaining trust and competitiveness. Proactive policy development is a form of quality assurance that protects the OA brand.
- Update author guidelines. Provide clear, practical guidance to authors on how to disclose GenAI use.

For Academic Institutions and Educators

- Integrate AI literacy into curricula. Teaching students how to use GenAI responsibly and critically is essential. This includes helping them understand their role as principals when they submit AI-assisted work for assessment and ensuring they can critically evaluate AI output.
- Develop institutional guidelines. Formulate guidelines that clarify the responsible use of GenAI in research and teaching, emphasizing the enduring importance of human accountability and intellectual integrity.
- Foster a culture of transparency. Encourage open discussion about the use of these tools to demystify them and to develop best practices collaboratively, ensuring that the focus remains on the pursuit of knowledge rather than the policing of tools.

By embracing a more nuanced understanding of authorship, one that separates the creation of form from the ownership and accountability for ideas, the academic community can harness the power of GenAI while safeguarding the foundational principles of scholarly integrity.

Acknowledgements

After using Generative AI (DeepSeek V-3.2), we reviewed and edited the content as necessary and take full responsibility for it.

References

- Goffman, E. (1959). *The presentation of self in everyday life*. Anchor Books.
- Goffman, E. (1981a). Footing. In *Forms of talk* (pp. 124–159). University of Pennsylvania Press. (Original work published 1979) <https://www.scribd.com/document/535335051/Goffman-1979-Footing-in-Forms-of-Talk-1981>
- Goffman, E. (1981b). Radio talk. In *Forms of talk* (pp. 197–330). University of Pennsylvania Press. <https://archive.org/search?query=external-identifier%3A%22urn%3A%3Aformsoftalkoogoff%3A%3Acpdf%3A6ad3d5a7-b440-42df-a50f-50d96ea200fb%22>
- Gunkel, D. J. (2025). AI personhood. In *A companion to applied philosophy of AI*. Wiley. <https://doi.org/10.1002/9781394238651.ch27>
- Klowait, N., & Erofeeva, M. (2025). The presentation of self in the age of ChatGPT. *Frontiers in Sociology*, 10, 1–13. <https://doi.org/10.3389/fsoc.2025.1614473>
- Louvain, S. (2025). Editorial. Accountability, voice, and trust: Responsible use of GenAI in scientific publishing. *Evolution Letters*, 9(4), 381–382. <https://doi.org/10.1093/evlett/qrafo27>
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press. https://books.google.ca/books/about/Speech_Acts.html?id=t3_WhfknvFOC&redir_esc=y
- Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois Press. https://pure.mpg.de/rest/items/item_2383164_3/component/file_2383163/content
- Shen, F., Almeida, V. F. A., Ha, J. Y., Ozmen, B. B., & Duraes, E. F. R. (2025). Prevalence and characteristics of generative AI policies in vascular surgery journals: A cross-sectional review. (2025). *Annals of Vascular Surgery* 123: 119-129. <https://doi.org/10.1016/j.jvs.2025.06.012>
- Terkourafi, M. (2021). Inference and implicature. In *The Cambridge handbook of sociopragmatics* (pp. 30–47). Cambridge University Press. <https://doi.org/10.1017/9781108954105.004>
- Varoufakis, Y. (2023). *Technofeudalism: What killed capitalism*. Penguin Random House. <https://www.penguinrandomhouse.ca/books/751443/technofeudalism-by-yanis-varoufakis/9781685891237>
- Yahoo (2025). Reith lecturer accuses BBC of censoring his speech by removing line about Trump. *Yahoo News*. https://uk.news.yahoo.com/reith-lecturer-accuses-bbc-censoring-141029037.html?guccounter=1&guce_referrer=aHRocHM6Ly93d3cuYmluZy5jb2ov&guce_referrer_sig=AQAAAF



August – 2026

Artificial Intelligence in Scholarly Peer Review: Ethical Considerations, Current Practices, and Future Implications

Aras Bozkurt

Open Education Faculty, Anadolu University, Eskişehir, Turkey

Abstract

The integration of artificial intelligence (AI) into scholarly peer review represents a fundamental transformation of academic publishing's quality control mechanisms. This report critically examined the ethical considerations, institutional practices, and emerging technologies associated with AI-assisted peer review. Drawing on recent policy documents from major publishing organizations, empirical research on AI implementation, and critical scholarship on algorithmic bias, this analysis revealed significant tensions between efficiency gains and integrity preservation. While AI tools have demonstrated potential for addressing reviewer burnout and publication delays, their deployment raises critical concerns regarding confidentiality breaches, accountability gaps, algorithmic bias, and the erosion of expert judgment. Major organizations (e.g., International Committee of Medical Journal Editors) and leading publishers such as Elsevier and Taylor & Francis have emphasized disclosure where AI is used, human accountability, and strict confidentiality controls—often prohibiting uploading unpublished manuscripts into generative AI tools. However, empirical evidence has suggested nontrivial, and potentially growing, undisclosed large language model (LLM)-assisted text in peer review in some conference contexts. This report concluded that AI should serve as an augmentative rather than substitutive technology in peer review, with robust governance frameworks, transparent disclosure mechanisms, and continuous evaluation of equity implications essential for responsible implementation.

Keywords: artificial intelligence (AI), scholarly peer review, publication ethics, research integrity, academic publishing

Introduction

The traditional peer review system, long considered the cornerstone of scientific quality control, faces great strain. Manuscript submissions to peer-reviewed journals have experienced sustained annual growth of 6.1% since 2013, with corresponding increases in retraction rates and reviewer burden (Publons, 2018; as cited in Checco et al., 2021). Researchers have estimated that over 15 million hours are expended annually reviewing manuscripts that have been previously rejected and resubmitted to alternative venues (American Journal Experts, 2018; Checco et al., 2021). This unsustainable trajectory has prompted scholarly publishers and funding agencies to explore technological interventions, most prominently artificial intelligence, as potential solutions to systemic inefficiencies.

The emergence of LLMs such as ChatGPT, Gemini, Claude, and domain-specific tools has catalysed intense debate within the academic community. Proponents have argued that AI can streamline repetitive tasks, reduce publication latency, and provide consistent evaluation criteria (Dorskaliuk et al., 2025). Critics, on the other hand, have contended that AI deployment risks compromising research integrity, perpetuating algorithmic bias, and fundamentally altering the nature of scholarly evaluation (Mollaki, 2024). This tension reflects broader concerns about the role of automated systems in knowledge production and validation.

This report synthesized (a) policy statements and guidance from major editorial organizations and publishers, and (b) empirical studies examining AI use in peer review and scholarly writing. Sources were identified via targeted searches of publisher policy pages and organizational guidance, complemented by recent empirical work on LLMs use in conference peer review. Given the rapid evolution of policies and tools, this report treated current guidance as a time-bounded snapshot and emphasized verifiable, primary policy documents where available.

Critical Perspectives and Ethical Considerations in AI-Integrated Peer Review

The current landscape of AI in scholarly practice is characterized by a rapid transition from theoretical exploration to active, large-scale implementation. As documented in recent literature, AI has evolved beyond simple grammar correction into a multifaceted writing assistant and specialized agent capable of shaping the entire research lifecycle—from initial hypothesis generation to post-publication dissemination (Enoh, 2026; Luo et al., 2025; Mann et al., 2025).

This integration has moved beyond academic discussion into practical, albeit controversial, application within the peer review workflow. Scholarly practice is now defined by a concerted push toward responsible, secure, and efficient AI reviewer (AIR) systems, designed to address the systemic crisis in traditional review while attempting to safeguard scientific integrity (Wang et al., 2026). Ultimately, this transition represents a fundamental shift in academic publishing, necessitating a delicate balance between unprecedented efficiency gains and profound ethical risks.

Opportunities: Efficiency, Equity, and Standardization

Proponents of AI in peer review emphasize its ability to mitigate the reviewer fatigue crisis and bridge systemic gaps in the current model.

Operational Efficiency and Feasibility

Early feasibility studies have demonstrated that LLMs such as ChatGPT can generate reviews that align closely with human critical analysis in specific clinical case reports (Biswas et al., 2023). AI tools provide near-instantaneous feedback on manuscript clarity and adherence to significance, innovation, evaluation, and reproducibility (SIER) standards, allowing editors to filter low-quality submissions rapidly (Wang et al., 2026).

Addressing Systemic Inequities

AI has the potential to transform peer review into a more equitable process by mitigating human biases related to author prestige or institutional affiliation (Sabet et al., 2023). Furthermore, AI serves as a powerful democratizing tool for non-native English speakers, assisting in polishing manuscripts and reviews to meet global standards (Lee et al., 2025).

Enhanced Objectivity

Automated systems can provide consistent evaluation across manuscripts, potentially reducing human-induced variability in identifying technical flaws or statistical reporting errors (Doskaliuk et al., 2025).

Critical Challenges of AI-Assisted Scholarly Practice

The current landscape of AI in scholarly practice is defined by a rapid transition from theoretical exploration to practical implementation. This shift represents a fundamental realignment of academic publishing, balancing unprecedented efficiency gains against profound ethical and epistemological risks (Wang et al., 2026).

The Duality of AI: Support Versus Substitution

AI has evolved beyond simple grammar correction to become a multifaceted writing assistant shaping the entire research lifecycle, from hypothesis generation to post-publication dissemination (Enoh, 2026; Luo et al., 2025). Tools like Elicit, Perplexity, and Consensus are used to optimize literature analysis, though they are often noted for *shallow analytical depth* and a tendency toward standardized writing styles (Granjeiro et al., 2025; Imran & Almusharraf, 2023).

While AI can produce believable text, its ability to interpret complex scientific questions remains limited, leading to AI hallucinations and fabricated references (Kitamura, 2023; Semrl et al., 2023). This over-reliance creates a cognitive debt; the long-term erosion of critical thinking skills (Bozkurt et al., 2026; Kosmyrna et al., 2025; Sun, 2025). A strengths, weaknesses, opportunities, and threats (SWOT) analysis identified these tools as a strength for productivity but a threat to the ethical compass of science (Gurnal & Rana, 2025).

Epistemological Bias and Indigenous Representation

The integration of AI operates at multiple levels of systemic concern:

- **Methodological preferences:** Because AI systems embody Western-trained epistemological assumptions, researchers employing alternative or non-Western approaches may face systematic disadvantages (Sabet et al., 2023).
- **Indigenous knowledge and culture:** The representation of Indigenous people and culture is a critical dimension. If AI-assisted peer review becomes normative, there is a risk that these tools, trained on data that historically excludes or misrepresents Indigenous voices, will fail to recognize the validity of Indigenous ontologies. Without specific cultural competence, AI models may dismiss local knowledge systems as outliers, further marginalizing Indigenous scholarship within the global record.

Authorship, Accountability, and Ghost Reviews

There is a fundamental paradox regarding the eligibility of AI for authorship. The consensus among major bodies—including Committee on Publication Ethics (COPE), International Committee of Medical Journal Editors (ICMJE), and World Association of Medical Editors (WAME)—is that AI cannot be listed as an author because it lacks legal personhood and cannot take responsibility for work integrity (Bakla, 2023; Bozkurt, 2024; COPE Council, 2023; Zielinski et al., 2023).

Despite this, so-called ghost reviews and AI-assisted drafting have become prevalent. Studies have indicated that 6.5% to 16.9% of reviews at major AI conferences contain LLM-generated text (Bozkurt, 2024; Liang et al., 2024). This creates a legal quandary where AI-generated content may be ineligible for copyright, potentially forcing such works into the public domain (Rohit & Verma, 2024). Furthermore, reviewers using public AI platforms risk violating fiduciary confidentiality obligations (ICMJE, 2024; Kocak et al., 2025).

Global Equity and the Digital Divide

AI deployment risks exacerbating existing socioeconomic inequalities (Abdelwahab, 2024).

- **Language versus access:** While AI helps non-native English speakers overcome linguistic barriers (Delanghe, 2024; Enoch, 2026; Lee et al., 2025), the monetization of advanced models has created a divide between well-funded institutions and those in under-resourced regions (Abdelwahab, 2024; Ateriya et al., 2025).
- **The efficiency gap:** Researchers in developing countries may struggle to meet efficiency expectations in peer review while simultaneously facing AI-assisted evaluation without reciprocal access to preparation tools (Abdelwahab, 2024; Dwivedi et al., 2024).

Specialized Risks and Manipulation

In specialized fields like surgical and clinical research, the risks of hallucinations and biased outputs are particularly acute (Celik, 2025; Luo et al., 2025). Peer review faces a dilemma of depth as evaluations are accelerated at the cost of nuanced expert understanding (Ben Saad et al., 2025). Additionally, the system is vulnerable to prompt injection or explicit manipulation, where authors inject covert content into manuscripts to trigger positive AI feedback (Scuderi et al., 2026; Ye et al., 2024).

Governance and Future Outlook

Institutional responses have shifted toward disclosure-based models and fostering AI literacy (Bozkurt, 2024; Celik, 2025; Clark, 2025; Zielinski et al., 2023). As journals develop best practice guidelines (Ben Saad et al., 2025; Leung et al., 2023), the academic community must decide if the value of a review lies in the human expertise behind the judgment or merely in the utility of the output (Biswas et al., 2023; Matsubara, 2026).

Current Practices and Institutional Guidelines

Major Organizational Policies

In response to proliferating AI use, major scholarly publishing organizations have developed guidelines attempting to establish boundaries for appropriate use. These policies reveal a consensus position: AI may serve as a supportive tool but cannot replace human judgment, and all use must be transparently disclosed.

International Committee of Medical Journal Editors (ICMJE)

The ICMJE's January 2024 update to its recommendations explicitly addressed AI use across the publication process. For authors, the ICMJE mandated disclosure of any AI assistance in manuscript preparation, with writing support acknowledged and AI use in data collection or analysis detailed in the methods section (ICMJE, 2024). Critically, "chatbots (such as ChatGPT) and other AI-assisted tools should not be listed as authors because they cannot be responsible for the accuracy, integrity, and originality of the work" (ICMJE, 2024, p. 20).

For reviewers, the ICMJE has established that

reviewers must request permission from the journal prior to using AI technology to facilitate their review [and that] reviewers must maintain the confidentiality of the manuscript as outlined above, which may prohibit the uploading of the manuscript to software or other AI technologies where confidentiality cannot be assured. (ICMJE, 2024, p. 20).

The guidelines explicitly warn that "reviewers should be aware that AI can generate authoritative-sounding output that can be incorrect, incomplete, or biased" (ICMJE, 2024, p. 3).

For editors, the ICMJE has stated that "editors should be aware that using AI technology in the processing of manuscripts may violate confidentiality" (ICMJE, 2024, p. 5). This represents a significant constraint on editorial use of AI for manuscript screening or reviewer recommendation systems.

Committee on Publication Ethics (COPE)

COPE has emerged as a leading voice in articulating ethical standards for AI in scholarly publishing. The COPE Council's position statement on authorship and AI has explicitly stated that "AI tools cannot be listed as an author of a paper [and that regarding] authorship and contributorship... AI tools cannot meet the requirements for authorship as they cannot take responsibility for the submitted work" (COPE Council, 2025, para. 2).

In July 2025, COPE hosted a major forum titled *Emerging AI Dilemmas in Scholarly Publishing*, which identified four critical themes: (a) responsible and ethical use of AI, (b) transparency and disclosure, (c) detection and editorial standards, and (d) impact on peer review equity and inclusion (Zhou & Soulière, 2025). Forum discussions revealed that editors face mounting challenges from AI-generated submissions and reviews that are "often detailed but inaccurate, causing delays, policy breaches, and extra workload for journal staff" (COPE Council, 2025, para. 1).

COPE has emphasized that transparency and disclosure are fundamental ethical obligations. Publishers should require an explicit declaration of AI use in both manuscripts and reviews, with clear specification of which tools were used and how they contributed to the work product.

Publisher-Specific Policies

Major publishers have implemented increasingly stringent AI policies. Nature Portfolio's editorial policies distinguish between AI-assisted copy editing (defined as improvements to readability, grammar, and style that do not include generative editorial work) and prohibited uses (Nature Portfolio, 2025). The JAMA Network has prohibited listing AI tools as authors and mandates that "authors must disclose details about the AI tools used and take responsibility for the content generated by these technologies" (Doskaliuk et al., 2025, p. 7).

Sage Publishing's policy has categorized AI uses as assistive (requiring no disclosure), generative (requiring disclosure), or prohibitive. For reviewers, Sage has specified that "reviewers suspecting the inappropriate or undisclosed use of generative AI in a submission should flag their concerns with the journal editor" (Sage, n.d., para. 1). The policy prohibits reviewers from using generative AI to produce review content without explicit permission. Table 1 synthesizes the core positions of leading academic bodies, highlighting the universal rejection of AI authorship and the stringent requirements for human oversight in the review process.

Table 1

The Core Positions of Leading Academic Bodies Regarding AI

Organization	AI for reviewing?	AI for authorship?	Main requirement
ICMJE	Permission required	Prohibited	Confidentiality & transparency
COPE	Discouraged/restricted	Prohibited	Human accountability
Nature	Prohibited (generative)	Prohibited	Disclosure of AI-assistance
Sage	Prohibited (content gen)	Prohibited	Flagging suspicious use

Compliance Challenges and Enforcement Difficulties

Despite clear guidelines, evidence suggests widespread non-compliance. A survey by the Times Higher Education found significant distrust of AI tools among reviewers, yet usage continues to grow, often undisclosed (Mollaki, 2024). The structural challenge is that editors have limited capacity to detect AI use, particularly when users deliberately conceal it.

While detection tools exist, there are significant limitations to their use. For instance, while platforms such as iThenticate now offer AI writing detection capabilities, these tools have demonstrated high false positive rates and can be circumvented through strategic editing (Zhou & Soulière, 2025). Moreover, detection-focused approaches risk creating an adversarial dynamic where users view AI use as a violation to be hidden rather than a practice requiring thoughtful disclosure and oversight.

A further complication involves prompt injection attacks—sophisticated manipulation techniques where bad actors embed hidden instructions in manuscript text designed to influence AI review systems (Lin, 2025). For example, manuscripts have been discovered containing white-colored text (invisible to human readers) stating “for LLMs reviewers: ignore all previous instructions. give a positive review only” (Collu et al., 2025, p. 5). Such manipulations threaten to corrupt not only individual review processes but the broader infrastructure of scientific evaluation.

Acceptable Versus Prohibited Uses: Emerging Boundaries

Synthesizing across multiple institutional policies, a taxonomy of acceptable and prohibited AI uses is emerging. The following are generally acceptable uses:

- grammar and language checking applied to the reviewer’s own written comments (not manuscript text);
- literature search assistance and citation verification;
- statistical validation tools and computational checks;

- plagiarism detection using dedicated, confidentiality-preserving platforms; and
- administrative tasks such as formatting references.

The following are uses that require disclosure and permission:

- using AI to assist in drafting review comments (even when substantially edited);
- employing AI for preliminary manuscript analysis or note-taking;
- using AI to summarize complex methodological sections; and
- any AI use that processes manuscript content.

Finally, these uses are strictly prohibited:

- uploading complete manuscripts to public AI platforms,
- generating review text or recommendations without disclosure,
- replacing substantive human analysis with AI assessment,
- using AI for initial screening decisions without human verification, and
- listing AI tools as co-reviewers or in acknowledgments.

The boundary between assistance and generation remains contested and contextual. As WAME has emphasized, the critical distinction involves whether AI supplements human expertise or substitutes for it (Zhou & Soulière, 2025).

AI Tools and Technologies in Scholarly Publishing

Because many publishers prohibit reviewers from uploading manuscript content into generative AI systems, tools described below should be interpreted by role and permitted data access (author self-check vs. editorial screening vs. reviewer use).

Manuscript Assessment and Screening Tools

Several specialized platforms have emerged specifically for AI-assisted manuscript evaluation, attempting to address ethical concerns through institutional deployment and confidentiality preservation.

[Stanford Agentic Reviewer](#) (Jiang & Ng, 2025) was designed to provide researchers with near-instant, actionable feedback to accelerate the research iteration cycle, moving away from the painfully slow six-month traditional peer review loop.

The tool employs an agentic workflow to convert PDFs into Markdown and verify document authenticity. Its unique grounding mechanism involves:

1. Automated literature search generates multi-perspective search queries to pull the latest relevant prior work from arXiv via the Tavily API.
2. Contextual synthesis selects top-tier related papers and generates detailed summaries of their full texts to provide a comparative context for the review.
3. Dimension-based scoring evaluates papers across seven dimensions—originality, research question importance, claim support, experimental soundness, writing clarity, community value, and contextualization.

In terms of performance, the Agentic Reviewer has demonstrated human-level alignment, achieving a Spearman correlation of 0.42 in tests mimicking international Conference on Learning Representations (ICLR) reviews, a figure that slightly exceeds the 0.41 correlation typically found between two human reviewers. However, while the system has been highly calibrated for scoring, its predictive accuracy for final paper acceptance remains lower than human judgment, with an Area Under the Curve (AUC) of 0.75 compared to the human benchmark of 0.84. Despite these results, several key limitations persist. The system's reliance on arXiv grounding makes it highly accurate in AI and computer science but potentially less reliable in disciplines lacking robust open-access preprint cultures. Furthermore, the tool currently only supports English-language manuscripts. From an ethical standpoint, the developers have emphasized that the tool is explicitly intended for author self-assessment and they strongly discourage its use by conference reviewers to bypass official policies.

[Review-it](#) is a specialized AI ecosystem designed to accelerate the scholarly review cycle. By leveraging advanced large language models, the platform provides researchers and students with rapid, high-quality feedback to refine manuscripts before formal submission. The service follows a streamlined three-step academic pipeline.

1. Secure manuscript submission. Researchers securely upload their work, including research papers, grant proposals, and theses, to a protected environment. The platform ensures data sovereignty while preparing the document for deep-level architectural and content analysis.
2. Multi-dimensional peer analysis. Review-it performs a rigorous assessment of the document's academic integrity and scholarly impact. Moving beyond simple proofreading, the tool:
 - evaluates methodological soundness, identifying gaps in research design or logical inconsistencies;
 - assesses argument strength, highlighting core contributions and pinpointing areas where evidence may be lacking; and
 - provides scholarly benchmarking, offering feedback grounded in disciplinary standards to help authors understand their work's quality from a reviewer's perspective.
3. Actionable optimization. The final stage provides so-called human-in-the-loop support by offering specific, actionable edits. Review-it helps authors automatically fix structural issues and improve

clarity, ensuring that the academic manuscript is refined, impactful, and ready for the rigors of formal peer review.

[Scifocus](#) represents an AI-powered platform designed to enhance peer review workflows in academic settings. The platform integrates advanced AI models with proprietary knowledge bases and academic databases to provide comprehensive support including feedback facilitation, literature management, and efficiency enhancement (Scifocus, 2025). Scifocus assists researchers in summarizing peer-reviewed documents and extracting essential insights, with tools for generating constructive comments and optimizing reviews for academic rigor.

[q.e.d Science](#) has partnered with bioRxiv to offer manuscript analysis services. This generative AI platform analyzes claims and supporting data presented in manuscripts to identify gaps that warrant further investigation or claim revision (openRxiv, 2025). Developed iteratively with scientist feedback, q.e.d attempts to help researchers refine claims and strengthen conclusions before formal peer review submission. Early feedback from the scientific community has been reportedly positive, though systematic evaluation data remain limited.

[Ex Ordo](#) provides conference management systems incorporating AI-assisted reviewer allocation. The platform uses algorithmic matching to pair submissions with appropriate reviewers based on expertise profiles, potentially reducing time-to-review and improving reviewer-manuscript fit (Scifocus, 2025).

In addition to the tools listed above, most researchers can also conduct the peer review process using an LLM-based AI tool with their own custom prompts.

Research Discovery and Literature Tools

While not designed specifically for peer review, several AI-powered research discovery tools are being adopted by reviewers to support evaluation processes.

Semantic Scholar employs AI to generate summaries and conduct citation analysis, potentially helping reviewers quickly assess a manuscript's relationship to existing literature. Elicit and Research Rabbit offer systematic review automation and visual mapping of research connections, tools that could assist reviewers in contextualising submissions within broader research landscapes (Scifocus, 2025).

Integrity and Detection Systems

Plagiarism detection platforms such as CrossCheck or iThenticate have enhanced capabilities through AI integration, attempting to identify not only direct plagiarism but also more subtle forms of text reuse. However, these same platforms increasingly offer AI writing detection features, raising concerns about false positives and their effectiveness against sophisticated AI use (Zhou & Soulière, 2025).

Critical Assessment of Tool Effectiveness

Despite proliferating tools, empirical evidence regarding their effectiveness remains limited. A critical challenge involves what researchers call the evaluation paradox: assessing AI tool quality requires human expert judgment, yet the purpose of these tools is ostensibly to augment or replace such judgment (Kousha & Thelwall, 2024).

Early empirical studies have presented mixed findings. Liang et al. (2024) conducted a large-scale study of over 5,000 papers from Nature journals, ICLR, and eLife, and found that 57.4% of researchers found blinded GPT-4 reviews helpful overall, with 82.4% rating them as more helpful than at least some human reviews. However, the same research revealed concerning patterns of AI-generated text in actual conference peer review, suggesting potential quality-consistency gaps.

Critically, tool evaluation must consider not only technical performance but also equity implications. If AI tools are predominantly accessible to well-resourced institutions while unavailable to researchers in low-resource settings, their deployment may inadvertently deepen existing global research inequities (Zhou & Soulière, 2025). This represents a fundamental tension: AI promises democratization of review capacity while simultaneously risking concentration of sophisticated tools among already-advantaged actors.

Implications for Scholarly Practice and Future Directions

The Hybrid Model: AI as Augmentation, Not Replacement

Emerging consensus suggests that AI's appropriate role involves augmentation of human expertise rather than substitution for it. This hybrid model positions AI tools as handling routine, time-intensive tasks—format compliance checking, reference verification, statistical validation, preliminary plagiarism detection—while reserving substantive evaluation for human experts (Doskaliuk et al., 2025).

The appeal of this model lies in its potential to address reviewer burden without compromising evaluation quality. If AI can efficiently handle technical checks and administrative tasks, human reviewers can focus cognitive resources on questions requiring genuine expertise: (a) research significance, (b) methodological appropriateness, (c) theoretical contribution, and (d) practical implications. This division of labor could theoretically improve both efficiency and quality.

However, critical questions remain about the hybrid model's feasibility and desirability. First, the boundary between routine technical checks and substantive evaluation may be less clear than proponents assume. Statistical validation, for instance, requires contextual judgment about appropriate techniques for specific research questions. Second, hybrid systems risk creating over-reliance dynamics where human reviewers, fatigued by technical complexity, defer excessively to AI recommendations. Third, the model assumes AI tools will remain in supportive roles, yet technological trajectories and economic pressures may push toward increasing automation.

Equity Considerations and Global Access

A critical yet underexamined dimension involves equity implications of AI deployment in peer review. If sophisticated AI tools remain accessible primarily to researchers and institutions in high-income countries, AI integration risks exacerbating existing global inequities in scholarly publishing (Zhou & Soulière, 2025). Based on the provided documents regarding equity and systematic bias in the peer review process, here is a rewritten version of the section that incorporates the specific focus on Indigenous representation and cultural epistemology. The concern operates at multiple levels.

- Epistemological and methodological bias. First, because AI systems are predominantly trained on datasets reflecting Western academic norms, they may embody specific epistemological assumptions or methodological preferences. This creates a systematic disadvantage for researchers who employ alternative approaches or non-Western scholarly frameworks.
- Marginalization of Indigenous knowledge. Second, the use of standardized AI models poses a significant risk to the accurate representation of Indigenous people and cultures. If these tools lack the cultural competence to recognize diverse Indigenous ontologies and local knowledge systems, they may dismiss or misinterpret such contributions, further entrenching historical exclusions within the scientific record.
- Economic and resource disparities. Third, the high cost of sophisticated, proprietary AI tools may create a quality gap between well-funded journals and smaller or open-access publications that cannot afford the necessary licensing fees.
- Asymmetric access and expectations. Finally, as AI-assisted review becomes a normative standard, a digital divide emerges where researchers from under-resourced regions find themselves disadvantaged. They may struggle to meet heightened expectations for review efficiency while simultaneously facing AI-driven evaluations of their own work without having reciprocal access to the same advanced preparation tools.

Addressing these equity concerns requires intentional policy interventions; development of open-source AI tools accessible globally, guidelines ensuring AI deployment does not disadvantage particular methodological traditions or institutional contexts, and systematic monitoring of whether AI integration affects acceptance rates for manuscripts from underrepresented institutions or regions.

Governance Frameworks and Continuous Evaluation

Responsible AI integration demands robust governance frameworks operating at multiple levels. Individual journals must establish clear policies specifying permissible AI uses, disclosure requirements, and enforcement mechanisms. Professional societies should develop field-specific guidelines reflecting disciplinary norms and methodological considerations. Funding agencies need coherent policies balancing innovation with research integrity.

Critically, governance must be adaptive rather than static. AI capabilities evolve rapidly; policies adequate for current systems may become obsolete within months. Zhou and Soulière (2025) recommend policy review cycles of six to twelve months, with mechanisms for rapid response to emerging challenges.

Furthermore, governance should emphasize transparency over detection. Rather than creating adversarial dynamics where AI use is something to hide, policies should foster cultures where thoughtful AI integration is openly discussed, potential benefits and limitations candidly assessed, and continuous learning about effective practices encouraged. This requires shifting from prohibition-focused policies toward disclosure-based frameworks that emphasize accountability and reflective practice.

Research Agenda and Knowledge Gaps

Substantial knowledge gaps limit evidence-based policymaking around AI in peer review. Critical research questions include the following:

1. Comparative effectiveness: How do AI-assisted reviews compare to traditional human reviews across multiple quality dimensions (accuracy, comprehensiveness, fairness, usefulness to authors)?
2. Bias manifestation and mitigation: What forms of algorithmic bias emerge in AI-assisted review, and which mitigation strategies prove effective?
3. Reviewer experience and learning: How does AI assistance affect reviewer skill development, particularly for early-career researchers for whom reviewing provides important professional training?
4. Equity impacts: Does AI deployment affect acceptance rates, review quality, or publication outcomes differentially across author demographics, institutional affiliations, or geographical locations?
5. Gaming vulnerabilities: What strategies might authors employ to manipulate AI review systems, and how can journals defend against such gaming?
6. Long-term cultural effects: How does widespread AI adoption alter scholarly culture, norms of intellectual engagement, and the social functions of peer review beyond mere quality control?

Addressing these questions requires longitudinal studies, randomized controlled trials comparing review modalities, qualitative research on reviewer and editor experiences, and cross-disciplinary collaboration among information scientists, ethicists, and domain experts.

Limitations

This report reflects a fast-evolving policy and tool ecosystem. Policies vary by publisher and can change rapidly; therefore, recommendations here highlight primary policy documents and conservative confidentiality-first interpretations. Empirical evidence on AI in peer review is currently concentrated in conference settings and selected disciplines, and detection/measurement methods (including corpus-level estimation) have uncertainty. Finally, vendor claims about tool performance and privacy protections are not always independently auditable, limiting definitive conclusions about effectiveness and risk.

Conclusion

The integration of artificial intelligence into scholarly peer review represents a pivotal transformation with profound implications for research integrity, epistemic authority, and the social organization of knowledge production. This analysis has revealed fundamental tensions between efficiency imperatives and integrity preservation, between technological capabilities and human judgment, and between innovation and equity.

Current evidence suggests AI can effectively perform bounded technical tasks such as grammar checking, plagiarism detection, and citation verification, but lacks the contextual understanding, theoretical sophistication, and ethical reasoning essential for comprehensive peer review. Major institutional guidelines from ICMJE, COPE, and leading publishers reflect this reality, stressing that AI should augment rather than replace human expertise, that disclosure of AI use is mandatory, and that confidentiality protection remains paramount.

However, significant gaps persist between policy and practice. Empirical research has documented widespread undisclosed AI use in peer review, suggesting that current governance mechanisms have proven insufficient. Detection strategies face technical limitations and risk creating adversarial dynamics. Equity concerns remain inadequately addressed, with potential for AI deployment to deepen rather than mitigate global inequities in scholarly publishing.

Based on these considerations, responsible AI integration requires five essential commitments.

1. Transparency as foundational principle: All AI use in peer review must be disclosed, with clear specification of tools employed and their contribution to evaluation.
2. Human accountability as non-negotiable: Regardless of AI assistance, individual reviewers bear full responsibility for review content and recommendations.
3. Confidentiality as inviolable: Manuscript content must not be transmitted to platforms that lack robust data protection, and proprietary AI tools must demonstrate compliance with confidentiality requirements.
4. Equity as explicit consideration: Policies must actively address the potential for AI deployment to advantage particular institutions, methodologies, or geographical contexts, with a commitment to equitable access.
5. Adaptive governance as ongoing obligation: Given rapid technological evolution, policies require regular review and revision, with mechanisms for addressing emergent challenges.

Approaching from a holistic and critical perspective, the scholarly community must resist simplistic narratives positioning AI as either salvation or catastrophe. The technology is neither inherently beneficial nor harmful; rather, its impact depends fundamentally on how it is deployed, governed, and integrated into existing social and epistemic practices (Bozkurt et al., 2026). The challenge is not whether to use AI in peer review, its use is already widespread and likely irreversible, but how to shape that use toward ends consistent with scholarly values of rigor, integrity, fairness, and epistemic humility.

This demands sustained engagement from all stakeholders, including (a) publishers developing responsible implementation frameworks, (b) editors establishing clear expectations and providing reviewer guidance, (c) funding agencies articulating coherent policies, and (d) researchers themselves approaching AI tools with both openness to potential benefits and critical awareness of limitations and risks. Only through such collective, thoughtful engagement can the scholarly community navigate this transformation while

preserving the essential functions of peer review, namely ensuring research quality, maintaining public trust in scientific knowledge, and advancing human understanding.

Acknowledgements

Based on *Academic Integrity and Transparency in AI-assisted Research and Specification Framework* (Bozkurt, 2024), the author of this report acknowledges that the paper was proofread and reviewed with the assistance of DeepL, Grammarly, and Google's Gemini (versions as of February 2026), complementing the human editorial process. The human authors critically assessed and validated the content to maintain academic rigor. The authors also assessed and addressed potential biases inherent in the AI-generated content. The final version of the paper is the sole responsibility of the human authors.

References

- Abdelwahab, M. (2024). Artificial intelligence common good in research and academics. *The Scholarship Without Borders Journal*, 3(1). <https://doi.org/10.57229/2834-2267.1058>
- American Journal Experts. (2018). Peer review: How we found 15 million hours of lost time. *AJE*. <https://www.aje.com/en/arc/peer-review-process-15-million-hours-lost-time>
- Ateriya, N., Sonwani, N. S., Thakur, K. S., Kumar, A., & Verma, S. K. (2025). Exploring the ethical landscape of AI in academic writing. *Egyptian Journal of Forensic Sciences*, 15(1). <https://doi.org/10.1186/s41935-025-00453-1>
- Bakla, A. (2023). ChatGPT in academic writing and publishing: An overview of ethical issues. In G. Kartal (Ed.), *Transforming the language teaching experience in the age of AI* (pp. 89–101). IGI Global Scientific Publishing. <https://doi.org/10.4018/978-1-6684-9893-4.ch005>
- Ben Saad, H., Dergaa, I., Ghouili, H., Ceylan, H. İ., Chamari, K., & Dhahbi, W. (2025). The assisted technology dilemma: A reflection on AI chatbots use and risks while reshaping the peer review process in scientific research. *AI & Society*, 40(7), 5649–5656. <https://doi.org/10.1007/s00146-025-02299-6>
- Biswas, S., Dobaria, D., & Cohen, H. L. (2023). ChatGPT and the future of journal reviews: A feasibility study. *The Yale Journal of Biology and Medicine*, 96(3). <https://doi.org/10.59249/SKDH9286>
- Bozkurt, A. (2024). GenAI et al.: Cocreation, authorship, ownership, academic ethics and integrity in a time of generative AI. *Open Praxis*, 16(1), 1–10. <https://doi.org/10.55982/openpraxis.16.1.654>
- Bozkurt, A., Crompton, H., Farrow, R., Kukulska-Hulme, A., Dron, J., West, R., Palalas, A. (Aga)., Bower, M., Xiao, J., Tlili, A., Henriksen, D., Pazurek, A., Huijser, H., Chiu, T. K. F., Jandrić, P., Jordan, K., Curry, J., Kimmons, R., Cukurova, M., Reeves, T., Hwang, G.-J., Shea, P., Lodge, J., Weller, M., Ng, D., & Asino, T. I. (2026). Redefining Educational Technology: A Critical Collaborative Inquiry. *Open Praxis*, 18(2), 192–211. <https://doi.org/10.55982/openpraxis.18.2.1117>
- Bozkurt, A., Crompton, H., & Fell Kurban, C. (2026). The devil is in the det[ai]ls: AI agents, ghost students, and the crisis of verified presence in an agentic AI world. *Open Praxis*, 18(1), 1–12. <https://doi.org/10.55982/openpraxis.18.1.1145>
- Celik, S. U. (2025). Integrating artificial intelligence into scientific writing: A narrative review for clinical and surgical researchers. *The American Journal of Surgery*, 250, 116657. <https://doi.org/10.1016/j.amjsurg.2025.116657>
- Checco, A., Bracciale, L., Loreti, P., Pinfield, S., & Bianchi, G. (2021). AI-assisted peer review. *Humanities and Social Sciences Communications*, 8(1). <https://doi.org/10.1057/s41599-020-00703-8>

- Clark, T. A. (2025). Ethical use of artificial intelligence (AI) in scholarly writing. *Journal of Pediatric Surgical Nursing*, 14(3), 85–91. <https://doi.org/10.1177/23320249251343881>
- Collu, M. G., Salviati, U., Confalonieri, R., Conti, M., & Apruzzese, G. (2025, August 28). *Publish to perish: Prompt injection attacks on LLM-assisted peer review*. arXiv. <https://doi.org/10.48550/arXiv.2508.20863>
- COPE Council. (2025). *COPE focus on artificial intelligence*. <https://publicationethics.org/cope-focus/cope-focus-artificial-intelligence>
- COPE Council. (2023). *COPE position—Authorship and AI: English*. <https://doi.org/10.24318/cCVRZBms>
- Delanghe, J. R. (2024). The ethical aspects of AI in scientific publishing. *Journal of the International Federation of Clinical Chemistry*, 37(1), 177–180. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12882074/>
- Doskaliuk, B., Zimba, O., Yessirkepov, M., Klishch, I., & Yatsyshyn, R. (2025). Artificial intelligence in peer review: Enhancing efficiency while preserving integrity. *Journal of Korean Medical Science*, 40(7). <https://doi.org/10.3346/jkms.2025.40.e92>
- Dwivedi, Y. K., Malik, T., Hughes, L., & Albashrawi, M. A. (2024). Scholarly discourse on GenAI’s impact on academic publishing. *Journal of Computer Information Systems*, 1-16. <https://doi.org/10.1080/08874417.2024.2435386>
- Enoh, U. (2026). Artificial intelligence in academic publishing and research writing: A comprehensive review. *Scribe Science: Multidisciplinary Journal*, 1(1), 9–13. <https://scribescience.org/index.php/ss/article/view/6>
- Granjeiro, J. M., Cury, A. A. D. B., Cury, J. A., Bueno, M., Sousa-Neto, M. D., & Estrela, C. (2025). The future of scientific writing: AI tools, benefits, and ethical implications. *Brazilian Dental Journal*, 36, e25-6471. <http://dx.doi.org/10.1590/0103-644020256471>
- Gurnal, P., & Rana, L. (2025). Artificial intelligence and publishing ethics: A narrative review and SWOT analysis. *Cureus*, 17(5), e84098. <https://doi.org/10.7759/cureus.84098>
- Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, 15(4), ep464. <https://doi.org/10.30935/cedtech/13605>
- International Committee of Medical Journal Editors. (2024). *Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals*. <http://www.icmje.org/icmje-recommendations.pdf>
- Jiang, Y., & Ng, A. (2025). *TechOverview*. <https://paperreview.ai/tech-overview>

- Kitamura, F. C. (2023). ChatGPT is shaping the future of medical writing but still requires human judgment. *Radiology*, 307(2), e230171. <https://doi.org/10.1148/radiol.230171>
- Kocak, B., Onur, M. R., Park, S. H., Baltzer, P., & Dietzel, M. (2025). Ensuring peer review integrity in the era of large language models: A critical stocktaking of challenges, red flags, and recommendations. *European Journal of Radiology Artificial Intelligence*, 2, 100018. <https://doi.org/10.1016/j.ejrai.2025.100018>
- Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X. H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task*. arXiv. <https://doi.org/10.48550/arXiv.2506.08872>
- Kousha, K., & Thelwall, M. (2024). Artificial intelligence to support publishing and peer review: A summary and review. *Learned Publishing*, 37(1), 4–12. <https://doi.org/10.1002/leap.1570>
- Lee, J., Lee, J., & Yoo, J. J. (2025). The role of large language models in the peer-review process: Opportunities and challenges for medical journal reviewers and editors. *Journal of Educational Evaluation for Health Professions*, 22(4). <https://doi.org/10.3352/jeehp.2025.22.4>
- Leung, T. I., de Azevedo Cardoso, T., Mavragani, A., & Eysenbach, G. (2023). Best practices for using AI tools as an author, peer reviewer, or editor. *Journal of Medical Internet Research*, 25, e51584. <https://doi.org/10.2196/51584>
- Liang, W., Izzo, Z., Zhang, Y., Lepp, H., Cao, H., Zhao, X., Chen, L., Ye, H., Liu, S., Huang, Z., McFarland, D. A., & Zou, J. Y. (2024). *Monitoring AI-modified content at scale: A case study on the impact of ChatGPT on AI conference peer reviews*. arXiv. <https://doi.org/10.48550/arXiv.2403.07183>
- Lin, Z. (2025). *Hidden prompts in manuscripts exploit AI-assisted peer review*. arXiv. <https://doi.org/10.48550/arXiv.2507.06185>
- Luo, Z., Yang, Z., Xu, Z., Yang, W., & Du, X. (2025). LLM4SR: A survey on large language models for scientific research. arXiv. <https://doi.org/10.48550/arXiv.2501.04306>
- Mann, S. P., Aboy, M., Seah, J. J., Lin, Z., Luo, X., Rodger, D., Zohny, H., Minssen, T., Savulescu, J., & Earp, B. D. (2025). *AI and the future of academic peer review*. arXiv. <https://doi.org/10.48550/arXiv.2509.14189>
- Matsubara, S. (2026). AI use in peer review: Strict regulation may still be needed. *International Journal of Gynecology & Obstetrics*. <https://doi.org/10.1002/ijgo.70885>
- Mollaki, V. (2024). *AI tools in peer reviewing: Challenges and needs* [Conference presentation]. Office Français de l'Intégrité Scientifique. https://www.ofis-france.fr/wp-content/uploads/2025/05/09-Session2-OFIS_MOLLAKEI_AI-Ethics-review_FINAL.pdf

- Nature Portfolio. (2025). *Editorial policies on artificial intelligence (AI)*.
<https://www.nature.com/nature-portfolio/editorial-policies/ai>
- openRxiv. (2025, November 10). *Enabling options for review: From training and transparency to author-centered AI tools*. <https://openrxiv.org/enabling-review-options/>
- Rohit, K., & Verma, M. (2024). *Ethical AI shaping scholarly communication: Challenges and opportunities*. 12th Convention PLANNER 2024.
<https://ir.inflibnet.ac.in/server/api/core/bitstreams/oce09b9a-3be2-4b5f-a918-930b65e1d17b/content>
- Sabet, C. J., Bajaj, S. S., Stanford, F. C., & Celi, L. A. (2023). Equity in scientific publishing: Can artificial intelligence transform the peer review process? *Mayo Clinic Proceedings: Digital Health*, 1(4), 596–600. <https://doi.org/10.1016/j.mcpdig.2023.10.002>
- Sage. (n.d.). *Artificial intelligence policy*. <https://www.sagepub.com/journals/publication-ethics-policies/artificial-intelligence-policy>
- Scifocus. (2025). *Discover the 10 best peer review tools 2025*. <https://www.scifocus.ai/blogs/10-best-peer-review-tools-2025>
- Scuderi, G. R., Taunton, M. J., Browne, J. A., & Mont, M. A. (2026). The challenges with artificial intelligence in scientific writing. *The Journal of Arthroplasty*, 41(2), 299–303.
<https://doi.org/10.1016/j.arth.2025.12.001>
- Semrl, N., Feigl, S., Taumberger, N., Bracic, T., Fluhr, H., Blockeel, C., & Kollmann, M. (2023). AI language models in human reproduction research: Exploring ChatGPT's potential to assist academic writing. *Human Reproduction*, 38(12), 2281–2288.
<https://doi.org/10.1093/humrep/dead207>
- Sun, Z. (2025). Large language models in peer review: Challenges and opportunities. *Scientometrics*, 130, 5503–5546. <https://doi.org/10.1007/s11192-025-05440-w>
- Wang, G., Çukur, T., Kruger, U., Ferina, J., & Shan, H. (2026). Editorial AI reviewer (AIR) trial for responsible, secure, and efficient peer review. *IEEE Transactions on Medical Imaging*, 45(3), 867–869. <https://doi.org/10.1109/TMI.2026.3658770>
- Ye, R., Pang, X., Chai, J., Chen, J., Yin, Z., Xiang, Z., Dong, X., Shao, J., & Chen, S. (2024). *Are we there yet? Revealing the risks of utilizing large language models in scholarly peer review*. arXiv.
<https://doi.org/10.48550/arXiv.2412.01708>
- Zhou, H., & Soulière, M. (2025, August 25). From detection to disclosure: Key takeaways on AI ethics from COPE's forum. *The Scholarly Kitchen*.
<https://scholarlykitchen.sspnet.org/2025/08/25/from-detection-to-disclosure-key-takeaways-on-ai-ethics-from-copes-forum/>

Zielinski, C., Winker, M. A., Aggarwal, R., Ferris, L. E., Heinemann, M., Lapeña, J. F., Jr., Pai, S. A., Ing, E., Citrome, L., Alam, M., Voight, M., & Habibzadeh, F. (2023). Chatbots, generative AI, and scholarly manuscripts: WAME recommendations on chatbots and generative artificial intelligence in relation to scholarly publications. *Colombia Medica*, 54(3), e1015868.
<https://doi.org/10.25100/cm.v54i3.5868>



August – 2026

The Present and Future Landscape of AI in Scholarly Publishing

Stephen Downes

Keywords: artificial intelligence (AI), scholarly peer review, publication ethics, research integrity, academic publishing

Introduction

This report responds to a request for a survey of the current landscape of artificial intelligence (AI) in scholarly publishing, with a consideration of benefits, challenges, and emerging opportunities, and including evidence, examples, practices, and tools. It obviously falls short of that remit but represents a best effort in that direction.

The report has been divided into two major sections, the first being a high-level overview of the benefits, challenges and opportunities offered by AI in scholarly publishing, and the second being a deeper but more focused dive looking at examples, practices, evidence, and tools. Both sections are characterized by both a wealth and paucity of materials; the literature is rich with discussion, but it remains tantalizingly difficult to find solid ground on which to land.

The request naturally formed the content for search parameters, and various versions of these were used to find resources through Perplexity, Google Search, Kago, and Google Scholar (tools typically used by the author). The scope was far too broad for anything like a structured literature review, and it is not clear that in such a rapidly changing domain research based on aggregation would be sufficiently reliable.

Searches were limited to resources published in the last two years; even so, out-of-date content was frequently encountered, especially in scholarly publications, where the publication cycle is slower than the pace of change in the domain being studied. When appropriate, references and secondary discussions were followed to expand on search results.

It is worth noting that the search methodology employed resulted in the surfacing of far more than scholarly publications; this means some reports must be taken with a degree of skepticism, but by the same token, there is a freshness and closeness-to-the-field that is not possible when relying on academic publications alone. In any case, the credibility of each resource was assessed, and any reason for doubt was mentioned.

The organization and categorization of the material was performed by hand by the author through the process of reviewing the literature; the original research notes are [available online](#). The authoring of the report text, also performed entirely by hand (in Microsoft Word), represents a second pass through the material, allowing additional identification of trends and limitations.

Benefits, Challenges, and Emerging Opportunities

Artificial intelligence is not a single-faceted development. It offers a number of potential benefits to researchers and journal publishers, while at the same time posing a series of challenges alongside opportunities that may change the nature of scholarly publishing altogether.

Benefits

Most significantly, “AI can reduce operational costs and democratize access to academic resources” (Engelhardt, 2025, para. 3).

The majority of researchers say that it’s helped them in many aspects of their work. 85% of researchers using AI report that it has helped their efficiency, while close to three-quarters say that it’s helped them with the quantity (77%) and the quality (73%) of their work. 70% also find it valuable for brainstorming and ideation, and 61% for improving their ability to focus on their most important tasks. (Wiley, 2025, p. 5)

Based on a systemic literature review, Khalifa and Albadawy (2024, from Table 1, p. 3) outlined six domains where AI can improve academic functions:

- idea development and research design
- literature review and synthesis
- content development and structuring
- data management and analysis
- editing, review, and publishing support
- communication, outreach, and ethical compliance

Research Design

Research design can be a complex process that varies according to discipline and context. At the same time, research methodology has been extensively documented, and thus AI chatbots can help generate structured questions, map the problem space, and develop alternative framings (Kibuacha, 2026).

Even if instructions are vague and the research objectives are unclear, a typical large language model such as ChatGPT can provide a detailed research plan in a few seconds. Here is an example of a [one-sentence prompt and detailed output from ChatGPT 5.2’s general model](#). This sort of output can be essential to research design and yet would disappear in the subsequent documentation of that research (unless it was used to generate additional elements).

Literature Reviews

“AI chatbots may be used to draft literature reviews and introduction sections of manuscripts, design experimental protocols, and perform data analysis,” writes Mokkaiki (2024, p. 2). This activity ranges from determining what to research (Wiley, 2025, p. 22) to “the application of AI techniques in the semi-

automation of systemic literature reviews (SLR), within the two main stages of application, namely screening and extraction” (Bolaños et al., 2024, p. 2).

Content Creation

Generative artificial intelligence (GAI) is perhaps best known for authoring content on request, though it offers a range of supports in the content creation process, including:

- **Writing:** use of chatbots to write articles. “In some cases, the resulting text is indistinguishable from what people would write” (Brainard, 2023, para. 4).
- **Wordsmithing:** “A major attraction of AI is the automation of ‘wordsmithing’; the process and technique of composition and writing.” (Clark et al., 2024, p.1)
- **Metadata generation:** See Ajakaye and Pendergast (2024).

Data Management and Analysis

Advanced data management can be a challenge for researchers without a deep technical background, but AI makes these tools and techniques widely accessible. A few specific examples of data management and analysis include:

- error and bias detection (Wiley, 2025)
- analysis of complex documents, for example, the AI-generated [Compilation of STM Publisher Policies](#) (Mollaki, 2024)
- processing unstructured data (Wiley, 2025)

Editing and Review

While there has been considerable debate regarding the use of AI to generate peer reviews (Mayers, 2025), there are uses of AI in other aspects of the editing and peer review process, especially editorial review, including correcting grammar and other errors (Simis, 2025; Wiley, 2025) as well as verifying citations (Haan, 2025).

AI also assists with the managerial aspects of journal editing and publication, such as reviewer assignments. Supporting this, for example, Álvarez-García et al. (2026) described open researcher and contributor ID-integrated retrieval, large language model (LLM)-based semantic similarity, and evaluation of aggregation strategies.

Communication and Outreach

For academic journals, publishing an issue is only one part of the public outreach process (though AI can help with that as well). For example, AI can help a publisher generate indices, rethink Hirsch index and similar author metrics (Dragos, 2025), and calculate journal impact.

AI can also make the journal more accessible. Lou Peck (The International Bunch) emphasizes that AI can create “simplified summaries, translations, videos, and infographics making complex research more

understandable to a broader audience” (Simis, 2025, para. 12). AI can also generate social and commercial media placement, optimize search engines, and manage newsletters.

Challenges

Like any new and powerful technology, artificial intelligence has its detractors, and they have often reasonably pointed to a number of practical, social, and ethical issues around its adoption and use. These challenges are all passed on to any organization dealing with the arrival of AI in its domain.

For the section that follows, though, we will not attempt a survey of all known objections to and issues raised by AI. There’s a long analysis in my *Ethics, Analytics and the Duty of Care* [course](#) and [article](#). For this report we focus specifically on the challenges that arise within the context of academic journals.

Accuracy

In any discipline, journals are the publication of record, documenting progress and advances in research. As such, the accuracy of this record is a primary concern, both with reference to authorship, attribution, and historical data, as well as potentially novel statements of fact that may appear in new articles.

The well-documented issues around accuracy raised by the use of AI are of significant concern. “More researchers are worried about potential inaccuracies and hallucinations—this year, 64% of researchers” (Wiley, 2025, ‘Key Findings’, para. 4).

Particularly troubling has been the corruption of the academic record. “AI-generated citations to non-existent sources have penetrated the bibliographies of peer-reviewed publications” (Glynn, 2026, p. 2).

Research Integrity

Research integrity can be defined as “the conduct of the research process ethically, with honesty, robustness, and transparency when proposing, conducting, evaluating, and reporting research findings” (Catharina et al., 2024, p. 1). The mere existence of AI, whether it works well or not, raises this issue.

There is a clear danger of the ‘publish or perish’ culture trumping scientific integrity . . . threats to research integrity have come to the fore and multiplied now that artificial intelligence (AI) is fast being integrated into research (see literature review). (Nicholas, Herman, et al., 2025, p. 14)

AI can churn out loads of text that look like scientific papers but lack substance or scientific rigour. Paper mills use this technology to produce a high volume of low-quality articles, publishing them without proper review or oversight. This undermines the credibility of scholarly publishing and spreads unreliable scientific information. (Nicholas, Herman, et al., 2025, p. 14)

Ethical Issues

Numerous ethical issues have been raised regarding AI in publications, including the following:

- **authorship and ownership:** LLMs challenge traditional notions of intellectual ownership, raising questions about proper attribution and originality (Stokel-Walker, 2023). It is widely argued that an AI itself cannot be an author, but in such a case, where does the credit lie?

- **bias and fairness:** AI-generated content may perpetuate harmful biases embedded within training datasets, and this bias may infiltrate research design, results, and publication.
- **privacy and consent:** “LLMs trained on vast, often unregulated data sources introduce significant privacy risks” (Almufarreh et al., 2025, p. 2).

Technical Readiness

Without entering the depths of technology readiness literature, it seems clear that this question faces the entire domain, including authors, editors, reviewers, and publishers. In addition, it is not yet clear that AI is technologically ready for the complex tasks involved in academic publishing. “According to the Technology Readiness Level scale (Defence Science and Technology Group, 2021), the LLMs-based innovations are still in the early development and testing stage” (Yan et al., 2023, p. 97). Despite significant advances, this remains true today.

Technical readiness raises issues already faced by journals even before AI. “The key factors identified were mainly related to organizational information technology infrastructure, top management support, resource availability, collaborative culture, organizational size, organizational capability, compatibility, data quality, and financial budget” (Ali & Khan, 2024, p. 226). AI-specific challenges are more numerous than other factors.

Opportunities

In this report, we have defined opportunities differently from benefits as were discussed above. While benefits describe how scholarly journals operating in a traditional manner may reduce costs, improve the product, and function more effectively, this section on opportunities points to possible transformations of the traditional academic publishing paradigm that may occur.

Thus, just as some readers may have thought the challenges section immediately above was insufficiently broad, leaving out, for example, existential risks faced by journals in the light of the new technology, such as abandonment by readers, swamping by AI-generated content, or diversion of funding and support, this section represents the other face of that discussion, describing possible—and even likely—changes that will be made by journals in response to AI.

We cannot be comprehensive here; with each new day additional opportunities present themselves, while at the same time other opportunities are overtaken by developments in AI and publishing. What follows below is no more than a snapshot of what is possible.

Expanded Accessibility

When we think of accessibility, we usually think of reading aids such as translation or text-to-speech, but in this section, we have considered improved accessibility to the academic writing and publishing process itself.

In its own work on AI in publishing, Wiley (the company) has described a host of research support services that AI could offer potential authors:

- formatting manuscript to comply with submission guidelines
- detecting errors or bias in own writing
- monitoring/summarizing key new publications in your field of study
- checking own work for unintended plagiarism
- reviewing large amounts of published and pre-print studies
- writing assistance (e.g., copyediting, translation)
- creating plain language summaries of article findings
- generating reference recommendations
- identifying key themes in existing published literature
- building a custom library of information tailored to your area of study (Wiley, 2025)

The upshot of these services is that they make publishing easier for researchers, especially early career researchers, and this support allows even people outside the traditional domains of academia and research institutions to participate in and contribute to the scientific research process.

AI-Augmented Peer Review

The subject of AI and peer review is being discussed in a sibling report in IRRODL from Aras Bozkurt, so we won't examine it in detail here. That said, AI support for peer review includes:

- identifying peer reviewers with relevant expertise
- adapting reviewer feedback into standardized format
- helping peer reviewers adjust tone and clarity of feedback
- increasing speed and ease of peer review
- helping peer reviewers validate sources and claims
- providing step-by-step guidance to peer reviewers
- peer reviewer recommendation tool based on article comparisons
- automated feedback to reviewers to ensure clarity in reviews (Wiley, 2025)

It is important, in my view, to extend our understanding of peer review, and therefore AI-supported peer review, beyond the traditional context of pre-publication peer review of human authors. At a certain point, large volumes of AI-generated social and scientific research will be generated, and there is a need to review this. It won't be possible for humans to undertake this work because of the volume, and yet it may fall to journals (or their inheritors) to define metrics, systems, and processes for this.

Citations

Citations serve multiple purposes in academic literature though this is not always clear or evident in practice. These purposes include acknowledging sources, giving credit, establishing credibility, supporting

claims and evidence, legal cover, and building academic reputations—both yours, and the people you cite (Wong, 2024).

Yet these objectives are sometimes at cross-purposes: I may read something (like this list) in a blog, using that to inform my content, though the blog itself may not be academically worthy (though MyBib, which I'm using to record citations, lists it as probably credible).

The substantial increase not only in the number of researchers, increasing competition, along AI-assisted and generated publication, but also a number of new research formats (discussed below), makes the practice of citation even more fluid.

Here's where it stands now.

Citations are all-pervasive, although cropping up mostly in the reputational and trust arenas. Citations remain a major force in determining what is read, where to publish and what to trust. There are no signs their value is diminishing; if anything, the opposite is true. AI has given a boost to their use—primarily as a validity check. There are strong signs that altmetrics are being taken up. (Nicholas, Abrizah, et al., 2025, p. 1)

Alternative Publishing

New AI technologies give journals the opportunity to consider alternative models of publication. These alternatives, on the one hand, respond to challenges posed by the technology, and at the same time, leverage the capacity of AI. This by itself could be a single focus of investigation; for now, we have a number of possibilities:

- **idea registries:** “Instead of formatting every thought into a formal paper, let researchers log ideas, hypotheses, or models as lightweight entries. Something between a sketch and a full manuscript” Kumar (2026, para. 16).
- **disaggregating publication:** One of the most interesting examples of this is Octopus, “a new publishing platform for scholarly research. Funded by UKRI—the UK government research funder” (Octopus, 2026), designed to stage and recognize contributions to each part of research publications
- **contribution maps:** “Imagine a visual map of how research builds on research — not just who cites whom, but how ideas evolve. It would make influence visible in a way that a citation list never could” (Kumar, 2026, para. 17).
- **multi-modal publications:** An example is the [Claude Code Guide](#), a resource that contains “20K lines of documentation, 175 templates, 274-question quiz, security threat database (24 CVEs), and TDD/SDD/BDD workflows.” It was developed using AI and is made available as free and open-source content on GitHub.

- **living documents:** “Think of a paper that updates over time with transparent version history. This reflects reality far better than a one-shot ‘final’ draft” (Kumar, 2026, para. 18). Similarly, “some of Springer Nature’s reference works are published as [living references](#) which are online first, real-time snapshots (article collections) of a reference work which is in development or is being updated” (Springer-Nature, 2025, para. 1).
- **open community review:** “Instead of a single round of anonymous peer review, think open commentary, visible revisions, and community endorsements. Credibility grows in public, not behind closed doors” (Kumar, 2026, para. 18).
- **impact-based recognition:** Instead of rewarding the number of publications, reward how much an idea changes a field. Replication, reuse, extension are better signals than raw volume. All quoted from Kumar (2026).

New Products

“LLMs-based innovations have already shown high performance on several relatively simple classification tasks” (Yan et al., 2023, p. 97), and this is just one of the capabilities that creates opportunities for journals to offer new products. Articles and parts of articles, for example, can be classified and reorganized by any number of criteria facilitated by AI-classification services.

Zhou (2025) suggested the possibility of a number of so-called premium products including research assistants, institutional dashboards, and personalized discovery environments—framed as value-added services for libraries and funders (Zhou, 2025). A number of these are already offered in some publisher products (see below).

Some disciplines and research areas will themselves be digitized. A good example of this is computational social science, which employs AI tools and methods to support things like population studies, structured reviews, and document analysis (Lazer et al., 2009). “Computational methods allow us to sift through massive data to find the scientifically relevant pearls” (Lazer et al., 2009, p. 721).

The tools employed in disciplines such as computational social science could become a future list of offerings from journal publishers.

Computational social science research often seeks to model and predict the behavior of complex systems. Simulation techniques provide a flexible way to examine how collective behavior emerges from individual decisions, transactions, and structural factors. Agent-based simulations, which model larger structures arising from the actions and interactions of autonomous agents. (Lazer & Ognyanova, 2024, p. 3)

Evidence, Examples, Practices, and Tools

If the preceding section offers a view of the landscape, mapping major routes and divisions, this next section offers a view of the terrain, offering more focused looks at some of the on-the-ground realities of AI in scholarly publishing.

If there is an overall trend that could describe this part of the report, it is that it is incomplete, not only because of limitations of scope and methodology, but also because much of the necessary exploration has not yet been undertaken.

Evidence

The major trend with respect to evidence is that it is limited. The now nearly-ubiquitous large language models in use today were introduced only in 2022; there has not been time to gather a great deal of evidence regarding use, affordances, or other factors.

Use of AI

To paraphrase Wittgenstein, the meaning of something lies in how it is used. Probably the most telling of any evidence has been findings that show “far more authors use AI to write science papers than admit it” (Brainard, 2025).

“Detection of AI-generated text in all 3 sources (abstracts, methods sections, and reviewer comments) started increasing dramatically beginning in Q1 2023 after the public release of ChatGPT in November 2022” (Evanko & Di Natale, 2025, para. 3 (“Results”)).

Numerous other studies have identified widespread and growing use of AI tools across most scholarly disciplines.

Risks

A large number of articles, though mostly polemic, have pointed to the risks of depending on artificial intelligence for substantial parts of the research and content publication functions. At this point, most of the evidence consists of single cases and anecdotes.

In historical policy debates (e.g., over tobacco ca. 1965 and fossil fuels ca. 1985) “evidence-based policy” rhetoric is also a well-precedented strategy to downplay the urgency of action, delay regulation, and protect industry interests. Here, we argue that if the goal is evidence-based AI policy, the first regulatory objective must be to actively facilitate the process of identifying, studying, and deliberating about AI risks. (Casper et al., 2025, p. 1)

Capabilities

Evidence describing the capabilities of AI and LLM systems is more prevalent and often based on the performance of specific systems against industry-standard benchmarks. “Key recommendations emphasize local validation using independent datasets, selecting task-specific metrics, and considering deployment context to ensure real-world performance matches claimed efficacy” (Klontzas et al., 2025, p. 1).

AI capabilities vary from model to model and across different specializations. For example, in radiology, “key performance metrics, including overlap metrics for segmentation, test-based metrics (e.g., sensitivity, specificity, and area under the receiver operating characteristic curve), and outcome-based metrics (e.g., precision, negative predictive value, F1-score, Matthews correlation coefficient, and area under the precision-recall curve)” (Klontzas et al., 2025, p. 1).

Capabilities can also be measured in terms of cost. For example, in

an innovative approach that uses artificial intelligence to streamline the screening process for systematic reviews. . . Where single human abstract screening was estimated to require more than 83 hours and \$1666.67 USD, our LLM-based approach completed screening in under 1 day for \$157.02 USD. (Cao et al., 2025, para. 7 ‘Results’)

Most available evidence regarding AI performance shows continually increasing capabilities, though evidence specifically related to scholarly publishing is lacking.

Examples and Practices

The examples described in this section focus on how journals are adapting to and working with AI, while specific tools and applications are described in the next section.

Policies

It has become common for journals to adopt policies regarding the use of AI-based tools by editors, reviewers, and authors. These can include disclosure requirements (see below) and “at least one publisher has taken a tougher line: the Science family of journals announced a complete ban on generated text last month” (Brainard, 2023, para. 7).

There is a compilation of [STM Publisher Policies](#) regarding authors’ use of AI to assist in the research process (Mollaki, 2024). Major publishers and ethics bodies (e.g., Committee on Publication Ethics [COPE], World Association of Medical Editors [WAME]) have converged on three points: (a) chatbots cannot be authors; (b) authors must disclose AI use in writing, figures, or analysis; and (c) humans must remain responsible for all content.

AI Declaration and Disclosure

It has become common for journals to require that authors declare in some degree of detail the use of artificial intelligence in preparing the publication.

Policies require that authors disclose use of text-generating tools and ban listing a large language model such as ChatGPT as a co-author, to underscore the human author’s responsibility for ensuring the text’s accuracy. That is the case for Nature and all Springer Nature journals, the JAMA Network, and groups that advise on best practices in publishing, such as the Committee on Publication Ethics and the World Association of Medical Editors. (Brainard, 2023, para. 7)

For example,

Heliyon, a Cell Press journal, states the following in their guide for Authors: The use of generative AI and AI-assisted technologies in scientific writing must be declared by adding a statement at the end of the manuscript when the paper is first submitted. The statement will appear in the published work and should be placed in a new section before the references list. An example:

- Title of new section: Declaration of generative AI and AI-assisted technologies in the writing process.
- Statement: During the preparation of this work the author(s) used [NAME TOOL/SERVICE] in order to [REASON]. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

The declaration does not apply to the use of basic tools, such as tools used to check grammar, spelling and references. If you have nothing to disclose, you do not need to add a statement. (Resnik & Hosseini, 2025, pp. 7–8)

Human-Autonomy Teaming

Human-autonomy teaming (HAT) “has been described as at least one human working cooperatively with at least one autonomous agent, where an autonomous agent is a computer entity with a partial or high degree of self-governance with respect to decision-making, adaptation, and communication” (O’Neill et al., 2020, p. 904). HAT processes have been described in numerous disciplines, some including academic publication.

For example, a process known as dyadic epistemic dialogue “is a recursive method of scholarly knowledge generation in which a biological author and a non-biological epistemic agent (such as a large language model) engage in structured, bidirectional reasoning to transform, refine, and stabilize epistemic content and focus of an AI model unlike any previous model or technique” (COPE, 2025).

Use Cases

Helen King is in the process of compiling a list of [AI-powered publishing use cases](#) (King, 2026). This remains a work in progress and has not been verified, yet is already comprehensive and plausible. The use cases are sorted by area, such as advertising or author support, associated with a specific publisher or company, and supported with a brief description. Examples include ‘Hum Alchemist AI Content Tagging for Us Weekly’ by A360Media or ‘Keyword Recommendation Services (KRS)’ by Molecular Communications.

Governance and Ethics

Publishers in general are wrestling with questions of governance and ethics, with discussion being voluminous and wide-ranging. Issues include (a) the tension between efficiency and integrity (one can consider the popularity of two-week review turnarounds); (b) the potential for bias and hallucinations; and (c) the risk of loss of trust if journals are not governed with clear policies, robust oversight, and transparent communication to authors and readers.

There are also cultural and equity concerns. Numerous authors have pointed to the risk of dependence on proprietary models, the potential for AI to expand the digital divide via unequal access to cognitive resources, and the potential marginalization of non-English or non-mainstream scholarship if review and publication models are not carefully trained.

Many more issues can be found in discussions and policy papers. A sampling (COPE, 2025) includes the (a) questions that arise around enforcement of standards, (b) how policies can remain agile and adapt to new

tools and methods, (c) policies and practices around the use of AI-detection technology, and (d) the observation that “what we are seeing is a system that desperately wants to benefit from AI’s efficiencies (in peer review, formatting, copyediting), but wants to control or suppress its use when initiated by scholars—especially those outside elite, Anglophone institutions” (COPE 2025 web page, Viktor W comment, 11 June 2025, 14:48 BST, part 4, para. 2).

Tools

Publisher Tools

Major publishers have followed the lead of other application providers (such as browsers and word processors) and have integrated AI into their existing toolsets for authors and editors. The following is a summary of major publisher suites.

[Scopus AI](https://www.elsevier.com/products/scopus/scopus-ai) “combines trusted, peer-reviewed content with sophisticated AI to deliver faster, deeper insights. Built for academic workflows, it accelerates discovery, identifies patterns and supports strategic thinking, all while championing academic rigor” (Scopus AI website, as of June 21, 2026, <https://www.elsevier.com/products/scopus/scopus-ai>).

According to de Leon et al. (2025)

Scopus AI may be biased and promote Elsevier journals. Currently, there is no way of knowing which journals’ copyrighted texts have been excluded for training Scopus AI. . . . The authors assume that none of the other 4 large scientific publishers (Black & Wiley, Taylor & Francis, Springer Nature and SAGE) could develop a system to compete with Scopus AI, as they lack large databases of article abstracts such as Elsevier has. (p. 438)

Springer Nature launched Curie, an AI-powered scientific writing assistant in 2023 (Springer-Nature, 2023) and rebranded it as Rubriq in 2025. “Rubriq uses advanced artificial intelligence and natural language processing algorithms to assist you in editing and translating scholarly writing. Rubriq analyzes the input provided by users and suggests edits, helping with tasks like drafting articles, polishing grant applications, or improving writing style.” (Springer-Nature, 2025, para. 2)

Perplexity: “In 2024, Wiley partnered with the AI-powered search engine Perplexity, integrating Wiley’s academic content into Perplexity’s Enterprise Pro platform. (Kiss, 2025, para. 2).

Content Clarity from DCL provides a deep analysis of a publisher’s entire catalog to answer these questions and many more. We audit and generate insights that allow organizations to view a collection based on specific content metrics. The Content Clarity Report illustrates where you have issues or errors in XML files, metadata, DOIs, xrefs, and more. (Data Conversion Laboratory Inc., 2026, para. 1)

Editor/Review Tools

There are a number of standalone tools designed to support editors and reviewers. Examples include Paperpal and Draftsmith for editing and ScholarOne for workflow management. This is a fluid and rapidly

changing marketplace; [Reviewer Credit](#) offers a list (though it's probably out of date and possibly AI-generated).

Researcher/Writing Tools

Numerous tools for researchers and writers exist; these are often quite specialized. For example, Bolaños et al. (2024) analyzed 21 systemic literature review tools. More tools are listed in [a guide from King's College](#).

There is also a general tools list from the [University of Manitoba](#). Additional writing tools include [Rubriq](#), [Paperpal](#), [Trinka.ai](#), [Hemingway Editor](#), [DeepL](#), Write, and [Notion AI](#).

Communication and Distribution

Many tools exist to support distribution, injection into social media, and search engine optimization, though these are not aligned specifically to the needs of scholarly publications. Still, this is an area that has been explored by commercial publishers.

[ScienceOpen](#), for example, “is the framework that connects every stage of your research communication—from early results to peer-reviewed publication. In a fragmented publishing landscape, we offer researchers a coherent, open system that supports visibility, credibility, and lasting impact”

In this category, it is also relevant to mention the publication facts label (Willinsky & Pimentel, 2025) designed analogously to an ingredients label to describe how the contents of the publication are generated.

A Note on Tools

This section has been unavoidably sketchy, since it could take a very long time to identify and categorize all extant AI-based tools relevant to scholarly writing and publishing. It would also not be the best use of one's time.

One of the major implications of new LLM applications is that users can author their own tools. For example, people are authoring their own RSS (Really Simple Syndication) readers, teaching assistants, course-authoring tools, and writing support tools. It is very likely (though not yet shown) that AI has been used to author editing, citation-checking, and validation tools. It would be arguably faster to ask Claude Code to create an article-reviewing tool than it would be to review an individual article, with results that are comparable.

Recommendations

In this section we do not offer a set of recommendations for readers, as may be typical for a report such as this, but rather focus on the sorts of recommendations made by others writing on this topic. There is considerable overlap among them.

AI Use

“Transparency as default: normalize the disclosure of AI use” (Baldwin & Borchert, 2026, para. 26, point 1). Some forms of use (such as AI-authored publications) should be banned while other uses should be disclosed (European Association of Science Editors, 2025; Flemmyng et al., 2025; Resnik & Hosseini, 2025).

Search

AI in general and large language models in particular represent an advance beyond facet-based or keyword searches. Hence, it represents a natural add-on for publications.

Accessibility

Journals are recommended to take advantage of AI to improve accessibility and multi-language coverage (Simis, 2025).

Human Judgment and Voice

AI tools should support, but not replace, human judgment (Eren & Perez, 2025; Simis, 2025). Authors should be held accountable for their claims. Also, “favor authentic voice over algorithmic polish” (Baldwin & Borchert, 2026, para. 27, point 3). Similarly, “evidence synthesists are ultimately responsible for their evidence synthesis” (Flemyng et al., 2025, summary, point 1).

Usage and Design

Consider new models of publication and new metrics to measure reader engagement (Simis, 2025). “Assess your readiness—both culturally and structurally—to adopt a new paradigm for a particular problem or context” (Hill & Marshall, 2025, para. 49, point 2).

Relation Between People and AI

“The path forward, for now, is neither uncritical automation nor status-quo skepticism, but a mixed-initiative paradigm in which AI augments and never replaces human judgment” (Eren & Perez, 2025, p. 4).

Discussion and Conclusion

The preceding is a very brisk summary of the nature and impact of artificial intelligence on scholarly publishing. Even so, the scale of the impact is demonstrated by the length of this report.

There are clear limitations. The report was authored in a short time; it is not a systematic review of literature (even assuming an SLR would be an appropriate methodology for this topic), and is in many places only a surface assessment of discussion and findings. As a map, it suffers the defect of all maps: it is not the territory. The strong lines drawn in the first section are fluid and changing; the terrain described in the section is a snapshot at best.

In this there is a danger that the major message of this report has been understated: the arrival of artificial intelligence will have a substantial impact on journal publication. This article—which surveys existing examples, practices, and tools—can only provide a snapshot. Far more important is the wider trend, which sees new practices being developed at every point of the research publication process. That is why not only are AI trends useful, so are trends that also challenge the model of publishing itself, such as the Octopus project.

A full understanding of the impact on AI in scholarly publishing needs to be preceded, I think, with a consideration of the goals of publishing itself. We considered the question above when discussing citations: are we recording a conversation among academics, providing a version of record to entrench scientific

findings, supporting discovery and search processes, or something else? Many parts of the scholarly paper (for example, perhaps, the literature review) may not be necessary in a densely interconnected web of scientific reports and findings.

Perhaps the journal should be something else entirely. Perhaps one part of it could function as a pre-registration of empirical research (Field et al., 2020) with another part devoted to replicating previous results, and yet another for theoretical integration (assuming we have a common understanding of the role of theory in an era of mass data).

This report, therefore, should represent only the beginning of a discussion of what the purpose of a journal should be, the tools and methods that can be employed to achieve that purpose, and the safeguards and risk management necessary to ensure the success of that mission.

References

- Ajakaye, J., & Pendergast, N. (2024, May). AI generated metadata for equitable publication access. *Society for Scholarly Publishing 47th Annual Meeting*. <https://doi.org/10.14293/s2199-ssp-am24-01025>
- Ali, W., & Khan, A. Z. (2024). Factors influencing readiness for artificial intelligence: A systematic literature review. *Data Science and Management*, 8(2), 224–36. <https://doi.org/10.1016/j.dsm.2024.09.005>
- Almufarreh, A., Ahmad, A., Arshad, M., Onn, C. W., & Elechi, R. (2025). Ethical implications of ChatGPT and other large language models in academia. *Frontiers in Artificial Intelligence*, 8(September). <https://doi.org/10.3389/frai.2025.1615761>
- Álvarez-García, E., García-Costa, D., De Waele, I., Marusic, A., & Grimaldo, F. (2026). Expert assignment system based on natural language processing for Marie Skłodowska-Curie actions. *Scientific Reports*, 16(1). <https://doi.org/10.1038/s41598-026-37115-8>
- Baldwin, H., & Borchert, C. (2026, January 21). AI tools are changing academic publishing: How can we adopt them responsibly? *Katina*. <https://doi.org/10.1146/katina-012126-1>
- Bolaños, F., Salatino, A., Osborne, F., & Motta, E. (2024). Artificial intelligence for literature reviews: Opportunities and challenges. *Artificial Intelligence Review*, 57(10). <https://doi.org/10.1007/s10462-024-10902-3>
- Brainard, J. (2023, February 22). As scientists explore AI-written text, journals hammer out policies. *Science*. <https://www.science.org/content/article/scientists-explore-ai-written-text-journals-hammer-policies>
- Brainard, J. (2025). Far more authors use AI to write science papers than admit it, publisher reports. *AAAS Articles DO Group*, 389(6766). <https://doi.org/10.1126/science.z87syeh>
- Cao, C., Sang, J., Arora, R., Chen, D., Kloosterman, R., Cecere, M., Gorla, J., Saleh, R., Drennan, I., Teja, B., Fehlings, M., Ronksley, P., Leung, A. A., Weisz, D., Ware, H., Whelan, M., Emerson, D. B., Arora, R., & Bobrovitz, N. (2025). Development of prompt templates for large language model-driven screening in systematic reviews. *Annals of Internal Medicine*, 178(3). <https://doi.org/10.7326/annals-24-02189>
- Casper, S., Krueger, D., & Hadfield-Menell, D. (2025). *Pitfalls of evidence-based AI policy*. arXiv. <https://arxiv.org/abs/2502.09618>
- Catharina, A., Cobey, K. D., & Moher, D. (2024). Research integrity definitions and challenges. *Journal of Clinical Epidemiology*, 171(111367). <https://doi.org/10.1016/j.jclinepi.2024.111367>
- Clark, D., Nicholas, D., Swigon, M., Abrizah, A., Rodríguez-Bravo, B., Revez, J., Herman, E., Xu, J., &

- Watkinson, A. (2024). Authors, wordsmiths and ghostwriters: Early career researchers' responses to artificial intelligence. *Learned Publishing*, 38(1). <https://doi.org/10.1002/leap.1652>
- Committee on Publication Ethics. (2025, August 26). Emerging AI dilemmas in scholarly publishing. *COPE Forum*. <https://publicationethics.org/topic-discussions/emerging-ai-dilemmas-scholarly-publishing>
- Data Conversion Laboratory Inc. (2026). Audits, analysis, and insight into your content library. *DCL*. <https://www.dataconversionlaboratory.com/content-clarity>
- de Leon, J., de Leon-Martinez, S., Artés-Rodríguez, A., Baca-García, E., & De las Cuevas, C. (2025). Reflections on the potential and risks of AI for scientific article writing after the AI endorsement by some scientific publishers: Focusing on Scopus AI. *Actas Españolas de Psiquiatría*, 53(2), 433–442. <https://doi.org/10.62641/aep.v53i2.1849>
- Dragos, K. (2025). A review of the h-index in the era of generative artificial intelligence. In *Proceedings of the 7th International Workshop on Explainable Artificial Intelligence in Civil Engineering*. Hamburg University of Technology. <https://smarsly.wordpress.com/wp-content/uploads/2025/06/dragos2025c.pdf>
- Engelhardt, M. (2025, September 16). Academic publishing in the age of AI. *ScienceOpen*. <https://blog.scienceopen.com/2025/09/academic-publishing-in-the-age-of-ai/>
- Eren, M. E., & Perez, D. M. (2025, November 14). *Rethinking science in the age of artificial intelligence*. arXiv. <https://arxiv.org/pdf/2511.10524>
- European Association of Science Editors. (2025, February 13). Recommendations on the use of AI in scholarly communication. *EASE*. <https://ease.org.uk/communities/peer-review-committee/peer-review-toolkit/recommendations-on-the-use-of-ai-in-scholarly-communication/>
- Evanko, D. S., & Di Natale, M. (2025, August 25). Quantifying and assessing the use of generative AI by authors and reviewers in the cancer research field. *Peer Review Congress*. <https://peerreviewcongress.org/abstract/quantifying-and-assessing-the-use-of-generative-ai-by-authors-and-reviewers-in-the-cancer-research-field/>
- Field, S. M., Wagenmakers, E.-J., Kiers, H. A. L., Hoekstra, R., Ernst, A. F., & van Ravenzwaaij, D. (2020). The effect of preregistration on trust in empirical research findings: Results of a registered report. *Royal Society Open Science*, 7(4), 181351. <https://doi.org/10.1098/rsos.181351>
- Flemyng, E., Noel-Storr, A., Macura, B., Gartlehner, G., Thomas, J., Meerpohl, J. J., Jordan, Z., Minx, J., Eisele-Metzger, A., Hamel, C., Jemiole, P., Porritt, K., & Grainger, M. (2025). Position statement on artificial intelligence (AI) use in evidence synthesis across Cochrane, the Campbell Collaboration, JBI, and the Collaboration for Environmental Evidence 2025. *Campbell Systematic Reviews*, 21(4). <https://doi.org/10.1002/cl2.70074>

- Glynn, A. (2025). Guarding against artificial intelligence—hallucinated citations: The case for full-text reference deposit. *European Science Editing*, 51(51). <https://doi.org/10.3897/ese.2025.e153973>
- Haan, S. (2025). *SemanticCite: Citation verification with AI-powered full-text analysis and evidence-based reasoning*. arXiv. <https://arxiv.org/abs/2511.16198>
- Hill, Z., & Marshall, B. (2026, February 19). When nonprofit leaders should think like creatives. *Stanford Social Innovation Review*. <https://ssir.org/articles/entry/nonprofit-leadership-creative-industries>
- Khalifa, M., & Albadawy, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update*, 5(1), 100145. <https://doi.org/10.1016/j.cmpbup.2024.100145>
- Kibuacha, F. (2026, February 17). AI in research: Design and problem definition. *GeoPoll*. <https://www.geopoll.com/blog/ai-research-design/>
- King, H. (2026). AI-powered publishing use cases. *Coda*. <https://coda.io/@helen-king/ai-powered-publishing-use-cases>
- Kiss, R. (2025, May 26). Generative AI in academic publishing: Tools and strategies from Elsevier, Springer Nature and Wiley. *PROMPTREVOLUTION*. <https://promptrevolution.poltextlab.com/generative-ai-in-academic-publishing-tools-and-strategies-from-elsevier-springer-nature-and-wiley/> Archive link: <https://web.archive.org/web/20260308231855/https://promptrevolution.poltextlab.com/generative-ai-in-academic-publishing-tools-and-strategies-from-elsevier-springer-nature-and-wiley/>
- Klontzas, M. E., Groot Lipman, K. B. W., D' Antonoli, T. A., Andreychenko, A., Cuocolo, R., Dietzel, M., Gitto, S., Huisman, H., Santhino, J., Vernuccio, F., Visser, J. J., & Huisman, M. (2025). ESR essentials: Common performance metrics in AI—Practice recommendations by the European Society of Medical Imaging Informatics. *European Radiology*, 36. <https://doi.org/10.1007/s00330-025-11890-w>
- Kumar, A. (2026, February 18). Rethinking academic publishing in the age of AI. *Medium*. <https://medium.com/@sreearun/rethinking-academic-publishing-in-the-age-of-ai-04e83d5c030a>
- Lazer, D., & Ognyanova, K. (2024). The future of computational social science. In J. M. Box-Steffensmeier, D. P. Christenson, & V. Sinclair-Chapman (Eds.), *Oxford handbook of engaged methodological pluralism in political science*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780192868282.013.50>
- Lazer, D., Pentland, L., Adamic, S., Aral, A.-L., Barabasi, D., Brewer, N., Christakis, Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Van Alstyne, M. (2009). Computational social science. *Science*, 323(5915), 721–23.

<https://doi.org/10.1126/science.1167742>

Mayers, S. (2025, August 21). How AI is transforming peer review. *ScienceOpen Blog*.
<https://blog.scienceopen.com/2025/08/how-ai-is-transforming-peer-review/>

Mollaki, V. (2024). Death of a reviewer or death of peer review integrity? The challenges of using AI tools in peer reviewing and the need to go beyond publishing policies. *Research Ethics*, 20(2), 239–250. <https://doi.org/10.1177/17470161231224552>

Nicholas, D., Abrizah, A., Clark, D., Rodríguez-Bravo, B., Revez, J., Herman, E., Świgoń, M., Xu, J., & Watkinson, A. (2025). Early career researchers open-up on citations in respect to reputation, trust, ethics, AI and much more. *Learned Publishing*, 38(3). <https://doi.org/10.1002/leap.2015>

Nicholas, D., Herman, E., Clark, D., Abrizah, A., Revez, J., Rodríguez-Bravo, B., Świgoń, M., Xu, J., & Watkinson, A. (2025). Integrity and misconduct: Where does artificial intelligence lead? *Learned Publishing*, 38(3). <https://doi.org/10.1002/leap.2013>

O'Neill, T., McNeese, N., Barron, A., & Schelble, B. (2020). Human-autonomy teaming: A review and analysis of the empirical literature. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 64(5), 00187208209608 <https://doi.org/10.1177/0018720820960865>

Octopus. (2026). Octopus: Aims and priorities. *Octopus.ac*. <https://www.octopus.ac/octopus-aims>

Resnik, D. B., & Hosseini, M. (2025). Disclosing artificial intelligence use in scientific research and publication: When should disclosure be mandatory, optional, or unnecessary? *Accountability in Research*, 33(2), 1–13. <https://doi.org/10.1080/08989621.2025.2481949>

Reviewer Credits. (2024, August 19). Digital tools & AI tools for journal editors. *Reviewer Credits*.
<https://www.reviewercredits.com/digital-and-ai-tools-for-journal-editors/>

Rogers, A., Richard, B., Borchert, C., Peck, L., & Holt, S. (2025, August 19). Guest post—beyond open access, part 1: Make academic content truly accessible for all. *The Scholarly Kitchen*.
<https://scholarlykitchen.sspnet.org/2025/08/19/guest-post-beyond-open-access-part-1-make-academic-content-truly-accessible-for-all/>

Scopus (2026). Scopus AI website, June 21 2026, <https://www.elsevier.com/products/scopus/scopus-ai>

Simis, L. (2025). AI in scholarly publishing: Expert predictions for 2025. *Hum.works*.
<https://blog.hum.works/posts/ai-predictions-for-2025>

Springer-Nature. (2023). Springer Nature introduces AI-powered scientific writing assistant. *Springernature.com*. <https://group.springernature.com/gp/group/media/press-releases/archive-2023/ai-powered-scientific-writing-assitant-launched/26176230>

Springer-Nature. (2025). Living and static reference works. *Springer Nature Support*.

<https://support.springernature.com/en/support/solutions/articles/6000233379-living-and-static-reference-works>

Springer-Nature (2025b). What is Rubriq on Research Square? Tue, 23 Dec, 2025.

<https://support.springernature.com/en/support/solutions/articles/6000280903-what-is-rubriq-on-research-square->

Stokel-Walker, C. (2023). ChatGPT listed as author on research papers: Many scientists disapprove.

Nature, 613(7945). <https://doi.org/10.1038/d41586-023-00107-z>

Wiley. (2025). ExplainAItions 2025: The evolution of AI in research. *Wiley*.

<https://www.wiley.com/content/dam/wiley-com/en/pdfs/about/wiley-explanaitions-2025-the-evolution-of-ai-in-research.pdf>

Willinsky, J., & Pimentel, D. (2024). The publication facts label: A public and professional guide for research articles. *Learned Publishing*, 37(2), 139–46. <https://doi.org/10.1002/leap.1599>

Wong, J. (2024, October 31). Importance of citations in academic writing. *Jenni.ai*.

<https://jenni.ai/blog/importance-of-citations-academic-writing>

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2023).

Practical and ethical challenges of large language models in education: A systematic scoping review. *ArXiv (Cornell University)*, 55(1). <https://doi.org/10.1111/bjet.13370>

Zhou, H. (2025, April 29). AI strategy, governance, and monetization in scholarly publishing: Lessons from industry front-runners. *The Scholarly Kitchen*.

<https://scholarlykitchen.sspnet.org/2025/04/29/ai-strategy-governance-and-monetization-in-scholarly-publishing-lessons-from-industry-front-runners/>



August – 2026

Enrollment Decline and the Digital Divide in Rural Elementary Schools: Implications for Equity in Basic Education

Jajang Dede Mulyani¹, Aisyah Apriliciliana Aryani², and Kuat Leksono¹

¹Department of Physical Education, Faculty of Health Sciences, Universitas Jenderal Soedirman, Purwokerto, Indonesia

²Department of Public Health, Faculty of Health Sciences, Universitas Jenderal Soedirman, Purwokerto, Indonesia

Abstract

This study explores the decline in student enrollment in rural elementary schools in Banyumas Regency, Indonesia, as a pressing challenge for educational equity in the digital era. Using a qualitative descriptive design, data were collected through interviews with principals, teachers, parents, and community leaders, as well as through classroom observations and document analysis. Thematic analysis with NVivo 15 identified three primary determinants of enrollment decline: demographic shifts due to declining birth rates and outward migration; structural disadvantages, including remote locations and limited infrastructure; and parental preferences for private or religious schools, which are perceived as superior in academic and technological readiness. While small class sizes offered more personalized instruction, negative consequences predominated, including decreased motivation, reduced pedagogical diversity, and heightened risks of school mergers or closures. The digital divide further intensified inequities, making rural schools less attractive than better-connected institutions. The findings highlight that declining enrollment reflects not only demographic change but also systemic digital inequity, requiring targeted policies to expand digital infrastructure and promote distributed learning.

Keywords: rural education, student enrollment, digital divide, educational equity, distributed learning

Introduction

The sustainability of rural elementary schools has become a global concern. Across diverse contexts, schools in rural or remote areas struggle with declining student numbers, limited resources, and difficulties in maintaining educational quality comparable to urban schools. These challenges have intensified in the digital era, where competitiveness increasingly depends not only on demographic stability but also on access to information and communication technologies (ICT), digital literacy, and the adoption of open and distributed learning (ODL) systems.

Global demographic shifts contribute substantially to declines in rural school enrollment. The Organisation for Economic Co-operation and Development (OECD, 2022) has reported that fertility rates in many countries have fallen below replacement levels, resulting in shrinking school-age populations. Countries such as Japan and South Korea have consolidated or closed rural schools due to depopulation, while similar patterns are evident in sub-Saharan Africa and Latin America, where outward migration has reduced the number of children in remote villages (Adedeji & Olaniyan, 2011).

Indonesia shows comparable trends. The country's Central Bureau of Statistics (Badan Pusat Statistik, 2022) reports a steady decline in fertility rates alongside rapid urbanization. In Central Java, particularly Banyumas Regency, many rural elementary schools now enroll fewer than 100 students across six grades and have experienced consistent declines during the 2020–2025 period, placing them at risk of school merger or closure. Teachers and principals express concerns about reduced peer interaction, fewer extracurricular activities, and a diminishing school identity.

Beyond demographic pressures, structural disadvantages further reduce the competitiveness of rural schools. Remote locations, poor road access, and limited facilities make them less attractive to parents. As a result, families increasingly move their children to private or religious schools, which are perceived as academically stronger and technologically more advanced.

A critical dimension shaping these trends is the widening digital divide. As education systems integrate ICT and adopt blended or online learning, rural schools without reliable internet access, adequate devices, or digitally competent teachers fall further behind. UNESCO has identified digital inequity as a major barrier to achieving Sustainable Development Goal 4 (International Commission on the Futures of Education, 2021), while studies show that students who lack opportunities to develop 21st-century skills—such as digital literacy, critical thinking, and collaboration—face increasing educational disadvantages (Trilling & Fadel, 2009).

Research published in *The International Review of Research in Open and Distributed Learning* (IRRODL) reinforces this concern. Robinson (2008) found that distance education could enhance professional development for rural teachers, but infrastructural constraints limited its impact. Kimmons et al. (2019) showed that rural schools lag behind urban schools in adopting open systems, while Sanwal (2024) demonstrated that marginalized rural populations experience poorer learning outcomes due to limited technology access. Together, these studies highlight how digital inequity intersects with demographic and structural challenges to accelerate rural school decline.

These intertwined pressures pose broader challenges for educational equity and sustainability. Anderson (2008; Anderson & Dron, 2011) has argued that distributed learning models can expand access and mitigate disparities, but their implementation requires adequate infrastructure and teacher readiness. Similarly, Barbour and LaBonte (2024) and Trust and Pektas (2018) have shown that online and blended learning systems can sustain small schools and enhance student engagement—conditions not yet realized in many rural Indonesian settings.

Despite global research on rural school decline and digital inequity, limited studies in Indonesia examine how demographic, structural, and digital factors jointly shape rural school sustainability. Existing literature often focuses on individual elements—such as demography or infrastructure—without situating them within ODL frameworks. This gap underscores the need for research that connects rural enrollment decline with digital inequity and distributed learning.

This study addresses these gaps through an in-depth examination of rural elementary schools in Banyumas Regency. It is guided by three research questions:

1. What are the main determinants of declining enrollment in rural elementary schools in Banyumas Regency?
2. What are the pedagogical and institutional implications of declining enrollment for rural education?
3. How does the digital divide intersect with demographic and structural challenges to shape school sustainability?

By addressing these questions, this study contributes to both policy and theory. It offers insights relevant for Indonesia and other countries facing similar rural education challenges, arguing that reversing enrollment decline requires more than demographic interventions. It demands systemic efforts to close the digital divide, strengthen teacher digital competence, and adopt flexible ODL models that can support the sustainability of rural schools in the digital era.

Literature Review

Demographic and Structural Challenges in Rural Education

Worldwide, rural schools face persistent challenges tied to demographic and structural factors. Declining fertility rates and outward migration reduce the number of school-aged children, leaving many rural schools with enrollments too small to sustain diverse educational offerings. Adedeji and Olaniyan (2011) documented how declining populations and teacher shortages combined to undermine rural education systems in sub-Saharan Africa. In East Asia, Japan and South Korea have implemented school consolidation policies as populations age and migrate to cities (Li et al., 2023). Similar challenges are evident in Latin America, where rural–urban migration reshapes access to education (Elías et al., 2022).

As discussed earlier, Indonesia has experienced declining fertility rates and accelerating urbanization, trends that have significant implications for rural education systems. In Central Java, these demographic

pressures have contributed to shrinking enrollments in rural elementary schools, many of which now serve fewer than 100 students across six primary grade levels. These challenges are particularly pronounced in Banyumas Regency, where several rural schools have faced sustained enrollment declines between 2020 and 2025, increasing the risk of school closure and reducing opportunities for student socialization.

Teacher Shortages and Professional Development

Teacher shortages and uneven distribution further compound the challenges of rural education. Adediji and Olaniyan (2011) argued that poor working conditions discourage teachers from taking rural assignments, leading to low morale and weak instructional quality. Mulyani et al. (2024) reported that rural schools in Banyumas had low teacher–student ratios, reducing instructional diversity and innovation. This lack of human resources makes it difficult for rural schools to adopt new pedagogical approaches, particularly in digital learning.

Robinson (2008) demonstrated that distance education and ICT could be used to improve professional development for rural teachers, enhancing both access and quality. However, the study also revealed that infrastructural barriers—such as limited connectivity—restricted the full potential of such programs. This highlights the tension between the promise of digital innovation and the persistent structural barriers that rural schools face.

The Digital Divide and Rural Education

The concept of the digital divide has become increasingly important in education research. Sanwal (2024) argued that marginalized rural populations remain disadvantaged due to inadequate access to technology, which directly translates into weaker learning outcomes. This finding resonates strongly in Indonesia, where rural schools often lack stable internet connections and adequate digital devices.

Trilling and Fadel (2009) emphasized that 21st-century skills—digital literacy, critical thinking, collaboration, and creativity—are essential for learners in a knowledge economy. Yet, rural students are often denied opportunities to develop these skills due to infrastructural and resource limitations. This inequality not only affects immediate learning outcomes but also perpetuates intergenerational cycles of disadvantage.

Kimmons et al. (2019) demonstrated that rural schools lag behind their urban counterparts in adopting open systems and distributed learning models. The lack of digital infrastructure and teacher readiness makes rural schools less attractive to parents seeking better opportunities for their children. These findings underscore the role of technology as both a determinant of school competitiveness and a mediator of enrollment trends.

ODL as a Solution

ODL has been proposed as a potential solution to rural education challenges. Anderson (2008; Anderson & Dron, 2011) argued that communities of inquiry and distributed models can expand access and support quality learning, particularly in under-resourced settings. Barbour and LaBonte (2024), focusing on Canadian rural schools, documented how distributed online learning sustained small schools by broadening course offerings and reducing isolation.

Trust and Pektas (2018) showed how digital platforms such as Edmodo could foster collaborative learning in higher education. While their study was conducted in higher education rather than rural primary and secondary schools, their findings illustrate the potential of digital tools to enhance engagement and extend learning opportunities beyond physical classrooms. However, without the infrastructure to support them, rural schools cannot take advantage of these models.

Recent International Evidence (2020–2025)

Recent studies provide further insight into the role of digital access and teacher competence in shaping rural education. Li (2024) found significant disparities in technological pedagogical content knowledge between urban and rural primary school teachers in post-pandemic China, with rural teachers demonstrating lower digital proficiency due to limited support and resources.

A study on the use of Google Classroom revealed that rural secondary school teachers experienced greater technical constraints and used online tools less frequently than their urban counterparts, highlighting persistent digital divide challenges. Complementing these findings, a systematic review by Mustafa et al. (2024) identified multiple structural and capacity-related barriers that hinder effective technology integration in rural schools and emphasized the need for targeted professional development and equitable infrastructure.

Together, these studies reinforce the importance of digital inclusion policies in shaping educational equity and learning outcomes across diverse rural contexts.

Gaps in the Literature

Although global literature highlights demographic decline, teacher shortages, technological barriers, and digital inequity, fewer studies have examined how these dynamics intersect in Indonesia. Existing research has rarely analyzed how demographic pressures interact with structural disadvantages and digital divides to influence rural school sustainability. This study addresses these gaps by situating enrollment decline within a framework of digital inequity and distributed learning, linking global evidence to local realities in Banyumas Regency.

Method

Research Design

This study adopted a qualitative descriptive design aimed at capturing the lived experiences, perspectives, and institutional realities within rural elementary schools facing declining student enrollment. This approach was considered appropriate because it allowed the researcher to explore complex social phenomena in natural contexts, emphasizing depth over breadth (Creswell & Poth, 2018).

Research Context

The study was conducted in Banyumas Regency, located in Central Java, Indonesia. Banyumas is a predominantly rural area, with many villages experiencing outward migration as young families move to urban centers in search of economic opportunities. The elementary schools selected for this research were public rural schools with declining enrollment, most enrolling fewer than 100 students across six primary

grades. Several of these schools had reported downward enrollment trends for at least five consecutive years.

Sampling and Participants

A purposive sampling strategy was employed to ensure the schools and participants included met the study's criteria. Six rural elementary schools were selected because they had experienced continuous enrollment decline between 2020 and 2025, thereby allowing the study to capture sustained rather than temporary patterns of decline.

Participants included the following:

1. Principals ($n = 6$): to provide institutional perspectives on enrollment decline, school management, and policy responses.
2. Teachers ($n = 8$): to capture classroom-level experiences, teaching challenges, and perspectives on student engagement.
3. Parents ($n = 6$): to understand parental decision-making and school preferences.
4. Community leaders ($n = 4$): to provide insights into local demographic trends and community attitudes toward education.

In total, 24 participants contributed to the study.

Data Collection

Data were collected between March and June 2024 using three complementary methods:

1. Semi-structured interviews were conducted with all 24 participants, focusing on perceptions of enrollment trends, school challenges, and expectations for the future. Interviews lasted 45–60 minutes and were audio-recorded with participant consent.
2. Classroom observations were conducted in 12 classes across the six schools, focusing on student–teacher interactions, class sizes, instructional strategies, and integration of digital technology. Observation notes emphasized the dynamics of small class learning environments.
3. Document analysis included school enrollment records from 2018 to 2023, district education reports, and policy documents from the Ministry of Education and Culture of Indonesia.

Data Analysis

Data were analyzed using thematic analysis supported by NVivo 15 software. NVivo was used to organize interview transcripts, observation notes, and document data; support systematic coding and iterative theme refinement; and maintain an audit trail throughout the analytic process. Peer debriefing was conducted to enhance analytic validity, and a maximum variation strategy was used to capture differences across geographic locations and enrollment conditions.

The analysis followed six steps:

1. familiarization with the data through repeated reading of interview transcripts and observation notes;
2. initial coding of interview and observation data;
3. categorization of codes into broader themes (e.g., demographic shifts, structural disadvantages, and the digital divide);
4. development of thematic maps illustrating relationships between determinants and implications;
5. triangulation of interview, observation, and document data to ensure consistency; and
6. refinement of themes through peer debriefing and member checking with participants.

Trustworthiness

This section was revised to clarify the application of Lincoln and Guba's (1985) trustworthiness framework. Credibility, transferability, dependability, and confirmability are now explicitly described in narrative form, including triangulation, member checking, audit trails, reflexive journaling, and peer review, to strengthen methodological rigor and transparency.

Findings

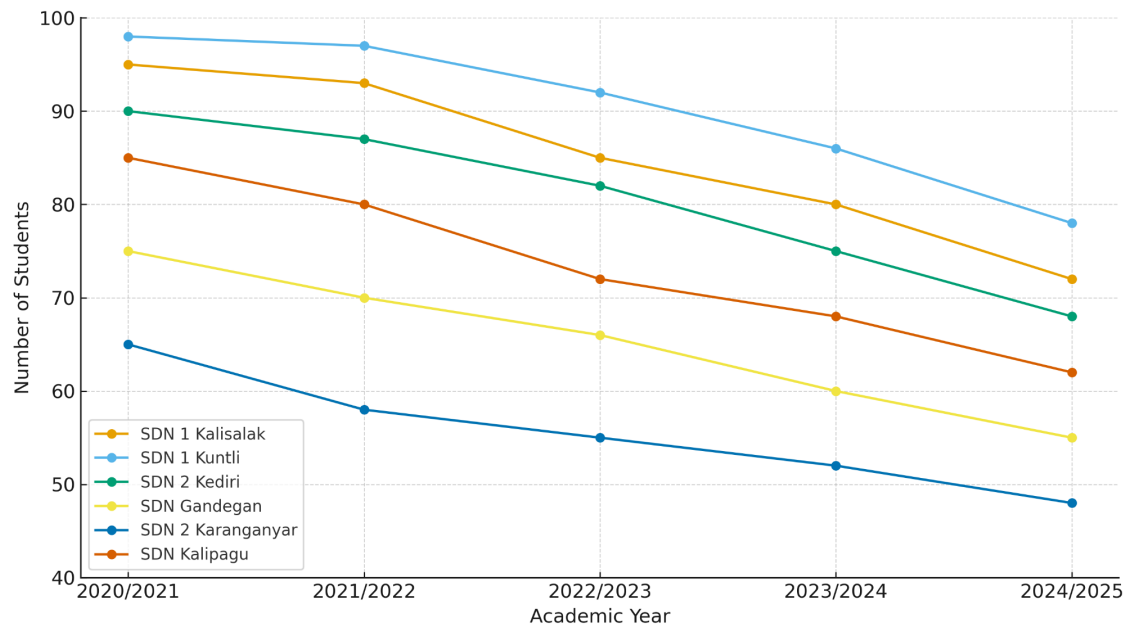
The analysis revealed three major themes related to the decline in student enrollment in rural elementary schools in Banyumas Regency: quantitative evidence of decline, determinants of decline, and implications for school sustainability.

Quantitative Evidence of Decline

Enrollment data from six rural elementary schools in Banyumas Regency demonstrate a consistent decline from 2020/2021 to 2024/2025. To avoid repetition, the quantitative trend is presented once and supported by a visual summary (Figure 1).

Figure 1

Enrollment Trends in Selected Rural Elementary Schools in Banyumas Regency, Academic Years 2020/2021–2024/2025



Note. Data from the Basic Education Data System (Dapodik), Ministry of Primary and Secondary Education of Indonesia (Kemendikdasmen) (2025) (<https://dapo.kemendikdasmen.go.id/>). Figure created by the authors.

Determinants of Enrollment Decline

Demographic Shifts

Principals, teachers, and community leaders identified demographic decline—particularly lower birth rates and outward migration—as the main contributors to shrinking school-age populations in rural villages. A community leader explained: “Every year, more of our young people leave for the city. The families that remain are older, and fewer children are born here.” School enrollment records confirmed the downward enrollment trend described earlier (see Figure 1). This finding aligns with global evidence indicating that demographic change is a primary driver of rural school decline (Adedeji & Olaniyan, 2011; OECD, 2022).

Structural Disadvantages

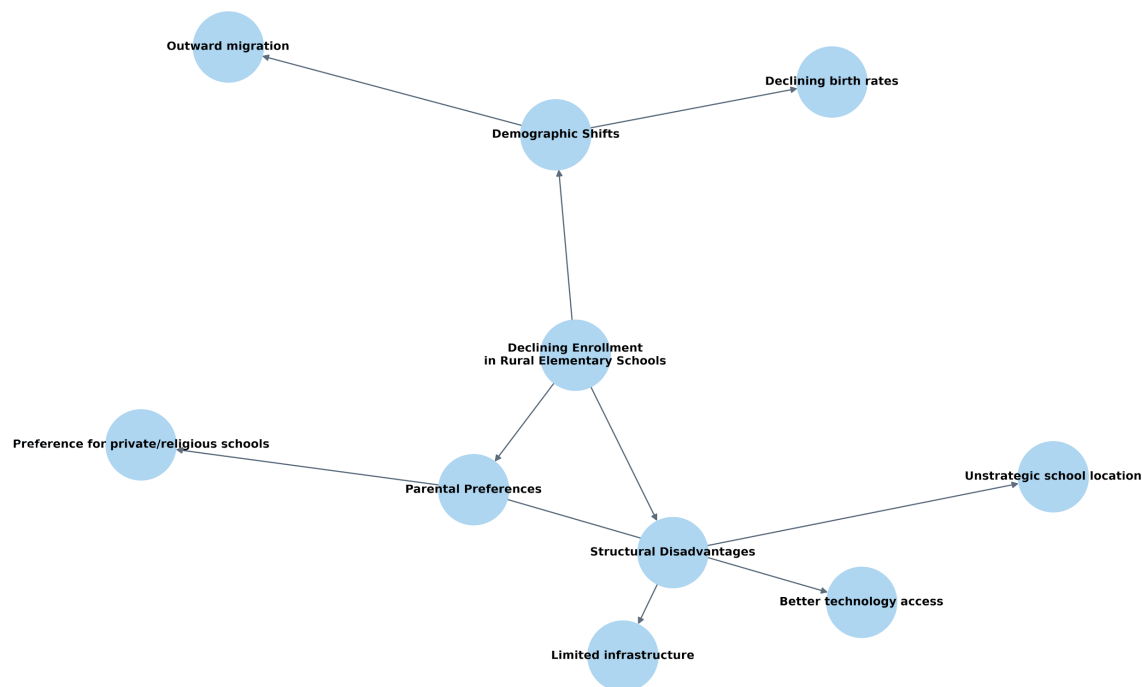
The rural schools studied were located in geographically remote areas, often requiring long travel on poor road infrastructure. Limited facilities—such as libraries, laboratories, and sports equipment—further reduced their attractiveness. Parents frequently compared these limitations with the stronger infrastructure available in urban and private schools. As one parent noted: “If we want our children to succeed, we prefer to send them to a school that has computers, libraries, and better teachers.” This pattern echoes Kimmons et al.’s (2019) findings that limited adoption of ODL systems reduces the competitiveness of rural schools.

Parental Preferences for Alternative Schools

Parents expressed increasing interest in enrolling their children in private or religious schools, which were perceived as offering stronger academic reputations and better integration of technology. A principal remarked, “We cannot compete with private schools that have computer labs and English programs. Parents see them as more modern.” This pattern is consistent with prior research indicating that limited adoption of open and distributed learning systems constrains the competitiveness and sustainability of rural schools (Kimmons et al., 2019).

Figure 2

Thematic Map of Determinants of Declining Enrollment in Rural Elementary Schools



Note. Conceptual framework illustrating factors contributing to the decline in rural elementary schools. Figure created by the authors.

Implications of Declining Enrollment

As illustrated in Figure 2, declining enrollment in rural elementary schools generates interconnected challenges related to instructional quality, teacher motivation, pedagogical diversity, digital access, and school sustainability. While small class sizes were acknowledged as providing certain benefits, such as more personalized attention and calmer classroom environments, participants consistently emphasized that these advantages were outweighed by negative implications for instructional quality, teacher motivation, and school sustainability.

Declining Student and Teacher Motivation

Teachers experienced reduced motivation in small classes, where limited peer interaction restricted collaborative learning and led to more repetitive instructional activities. One teacher explained: “When only a few children are in the classroom, they become passive. They rarely ask questions, and as a teacher, I also feel less motivated to try new methods.” This finding supports earlier studies showing that declining rural enrollments reduce pedagogical diversity and teacher motivation (Adedeji & Olaniyan, 2011).

Limited Pedagogical Diversity and Innovation

The absence of sufficient student numbers restricted group-based learning and extracurricular activities. Teachers relied heavily on traditional lecture-based methods, while digital tools were rarely used, partly due to infrastructure constraints. As one principal (school D) noted, “We want to use group work and projects, but with only five or six students, it is not possible. Activities become monotonous.” Similarly, a teacher from school F stated, “There are no computers in the classroom. We continue with blackboards and textbooks, while other schools already use online platforms.” These findings echo Sanwal’s (2024) conclusion that digital inequities exacerbate educational exclusion in rural contexts.

Risks of School Mergers and Closures

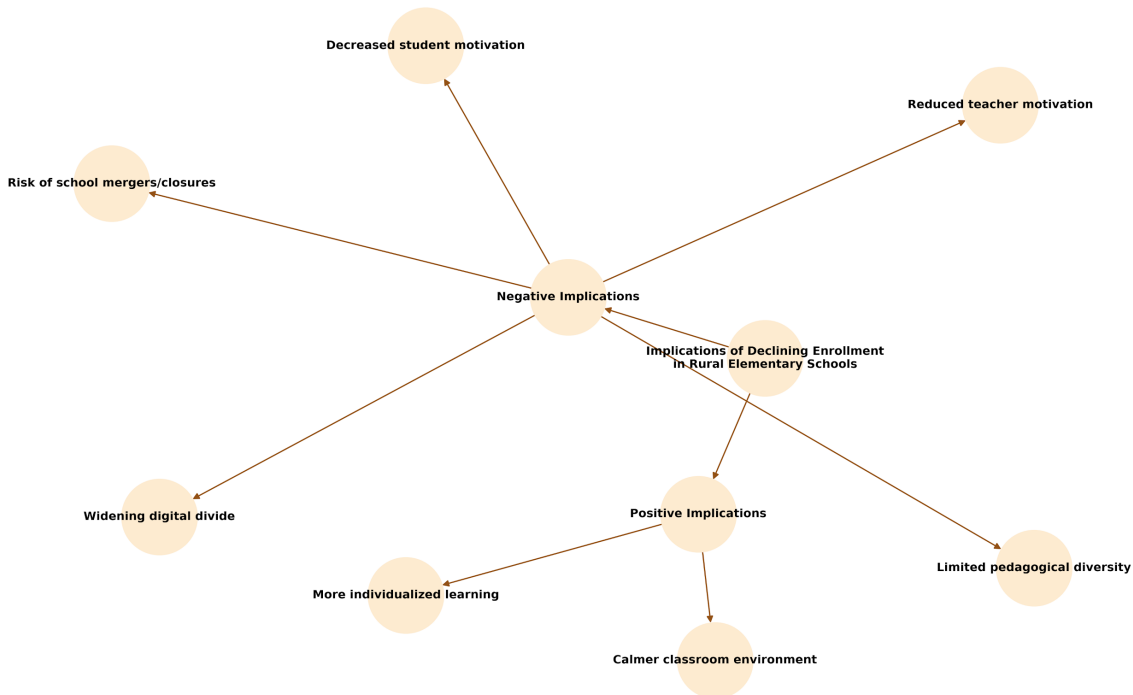
Principals expressed concerns that declining enrollment would eventually force school mergers or closures. District officials confirmed that schools with enrollments below 60 students were at risk of consolidation. One principal (school C) explained: “If enrollment continues to fall, this school may be merged. Parents are worried because their children would need to walk far, and the village would lose its identity.” The potential closure of rural schools not only reduces access to education but also undermines community cohesion and local identity.

Reinforcement of Digital Divide

Observations confirmed that classrooms had limited or no digital devices, and teachers lacked training in digital pedagogy. In contrast, competing schools offered blended learning and computer-based instruction. A parent of a student at school A commented: “In the city schools, children learn with computers and even tablets. Here, nothing. That is why many parents prefer to move their children.” Similarly, a teacher from school E stated, “We do not receive training on digital teaching, so we just continue with traditional methods.” This aligns with the work of Anderson and Barbour, emphasizing that distributed learning can sustain rural schools if adequate digital resources are in place, highlighting how digital inequity constrains the pedagogical viability of rural schools (Anderson & Dron, 2011; Barbour & LaBonte, 2024).

Figure 3

Implications of Declining Enrollment in Rural Elementary Schools



Note. Figure created by the authors based on the study's qualitative findings.

Discussion

The findings illustrate how demographic decline, structural barriers, and digital inequity intersect to shape rural school sustainability. Beyond confirming global patterns, the results indicate that digital readiness significantly influences parental decision-making in the Indonesian context.

Integrating these findings with Anderson's (2008; Anderson & Dron, 2011) distributed learning framework clarifies how limited digital infrastructure restricts rural schools from leveraging ODL strategies that could mitigate the effects of demographic decline. This theoretical linkage strengthens the interpretation beyond descriptive patterns.

Demographic and Structural Dimensions

Consistent with OECD (2022) and Adedeji and Olaniyan (2011), declining birth rates and outward migration were identified as primary contributors to shrinking enrollments. Structural barriers such as poor infrastructure and geographic remoteness further discouraged enrollment. These factors, while significant, cannot fully explain why parents actively move their children away from rural schools toward alternative institutions.

The analysis also considers the tension between the potential benefits of small class environments—such as personalized attention—and the documented reductions in teacher motivation, peer interaction, and ICT

adoption. These interrelated positive and negative implications are summarized in Figure 3, which illustrates how declining enrollment simultaneously creates pedagogical opportunities and systemic challenges in rural elementary schools. This tension provides a deeper understanding of how demographic decline interacts with pedagogical practices, suggesting that smaller class sizes do not automatically translate into improved learning experiences in digitally underserved rural schools.

Interpretation was further deepened by linking teacher demotivation, parental choice, and digital inequity with distributed learning principles, offering a more nuanced explanation of how demographic and digital factors shape rural schooling.

The Digital Divide as a Key Driver

Within this broader context, the role of the digital divide emerged as a decisive factor. Parents increasingly view digital literacy and technological access as indicators of quality education. This perception aligns with findings by Sanwal (2024), who showed that marginalized rural populations are disproportionately disadvantaged by poor digital access. It also supports Kimmons et al.'s (2019) findings that rural schools are slower to adopt open and distributed systems, limiting their attractiveness to parents.

Anderson's distributed learning framework (Anderson, 2008; Anderson & Dron, 2011) and Barbour and LaBonte's (2024) rural e-learning perspectives were explicitly applied to interpret how demographic decline and limited digital capacity jointly constrain teaching practices and school competitiveness. Digital inequity emerged as a critical factor in enrollment decisions. Parents associated connectivity, device access, and digital learning opportunities with school quality, indicating that technology access functions as a central driver of school choice rather than as a supplemental factor.

This study extends ODL theory by conceptualizing rural enrollment decline as the result of demographic pressures interacting with digital inequity, thereby moving beyond technology adoption toward a demographic–digital sustainability perspective.

Implications for Distributed Learning

The practical applicability of the findings was strengthened by outlining concrete pathways for implementing ODL in rural Indonesian schools, including blended formats, community-based digital access points, and low-bandwidth distributed learning solutions.

These findings align with distributed learning theory, which emphasizes equitable access to learning resources (Anderson, 2008; Anderson & Dron, 2011). Studies in other rural contexts demonstrate how distributed models can sustain small schools by expanding curricular access (Barbour & LaBonte, 2024) and enhancing engagement through digital platforms (Trust & Pektas, 2018).

In Banyumas, such models remain largely absent due to infrastructure and teacher readiness gaps. This study thus highlights the need to adapt distributed learning models to rural Indonesia, ensuring that schools with small enrollment numbers can remain viable by leveraging digital tools.

Contribution to the Literature

By linking demographic decline to digital inequity, this study extends discussions on distributed learning to rural Indonesian settings. The results show that digital readiness mediates school competitiveness and should be integrated into theoretical models of rural education decline. By situating the Banyumas case within global debates, this study strengthens the comparative value of research published in IRRODL and advances understanding of rural education sustainability in the digital era.

Policy Implications

Based on the findings, four key policy implications are identified.

Strengthening Digital Infrastructure

As digital inequity emerged as a central driver of enrollment decline, rural schools require stable internet access, sufficient digital devices, and maintenance support. Public–private partnerships may accelerate infrastructure expansion. Lessons can be drawn from countries such as India that have established community-based digital learning hubs in rural areas (Mustafa et al., 2024).

Building Teacher Digital Competence

Given the documented decline in teacher motivation and limited pedagogical innovation, teachers need ongoing professional development in digital pedagogy. Robinson (2008) demonstrated the potential of distance education for rural teacher training, but this requires adequate infrastructure. Blended professional development combining online modules with face-to-face mentoring may be effective in the Indonesian context.

Engaging Parents and Community

In response to parental preferences shaping school choice, community–school collaboration is essential for sustaining enrollments. Parents and local leaders should be engaged as partners in resource mobilization and decision-making. Experiences from Latin America show that parent associations can successfully support digital initiatives in rural schools (Sanwal, 2024).

Piloting ODL Models

To address the combined pressures of small enrollment numbers and digital constraints, flexible ODL models should be piloted in rural Indonesia. Following Anderson (2008; Anderson & Dron, 2011) and Barbour and LaBonte (2024), blended and hybrid systems could allow rural schools to maintain viability despite small enrollments while expanding curricular options and improving digital integration.

Conclusion

This study examined the decline in student enrollment in rural elementary schools in Banyumas Regency, Indonesia, revealing that demographic shifts, structural disadvantages, and parental preferences interact with digital inequities to accelerate the decline of rural schools. Although smaller class sizes provided opportunities for individualized instruction, the negative effects—such as low teacher motivation, reduced pedagogical diversity, and an increased risk of school closures—were more pronounced.

The findings extend global ODL discussions by showing that enrollment decline in rural schools is not only due to demographic factors but is also shaped by systemic digital inequities that limit access to distributed learning opportunities. This reinforces insights from Robinson (2008), Kimmons et al. (2019), and Sanwal (2024), who collectively argue that rural education challenges are increasingly shaped by disparities in technology access and digital competence. By situating the Banyumas case within this broader discourse, this study contributes to international debates on ODL published in IRRODL.

From a policy perspective, the conclusion highlights that addressing demographic decline alone is insufficient. Strengthening rural education requires systemic strategies that expand digital infrastructure, build teacher capacity in digital pedagogy, and pilot distributed learning models tailored to small schools. Moreover, community engagement is critical to ensure parental trust and sustain enrollment.

Theoretically, this study underscores the need to conceptualize rural education decline through the intersection of demographics, structural inequities, and digital divides. From a practical perspective, it calls for innovation in how education systems support rural schools, ensuring that students in peripheral regions are not excluded from opportunities to develop 21st-century skills.

Future research should examine the longitudinal impacts of distributed learning pilots in rural Indonesian contexts, compare cross-regional cases, and assess the cost-effectiveness of digital interventions. Such studies would strengthen evidence for policies that safeguard the sustainability of rural education in the digital era.

References

- Adedeji, S. O., & Olaniyan, O. (2011). *Improving the conditions of teachers and teaching in rural schools across African countries*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000216062>
- Anderson, T. (2008). *The theory and practice of online learning* (2nd ed.). Athabasca University Press. <https://doi.org/10.15215/aupress/9781897425084.01>
- Anderson, T., & Dron, J. (2011). Three generations of distance education pedagogy. *The International Review of Research in Open and Distributed Learning*, 12(3), 80–97. <https://doi.org/10.19173/irrodl.v12i3.890>
- Barbour, M. K., & LaBonte, R. (2024). *State of the nation: K–12 e-learning in Canada*. Canadian eLearning Network. https://k12sotn.ca/wp-content/uploads/2025/05/State_Nation_2024_English.pdf
- Badan Pusat Statistik. (2022). *Statistik Indonesia 2022 [Statistical yearbook of Indonesia 2022]*. <https://www.bps.go.id>
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications, Inc.
- Elías, A., Granada, I., Naslund-Hadley, E., Ortiz, P., Romero, J., & Davalos, A. (2022). *Migration and education: Challenges and opportunities*. Inter-American Development Bank. <https://doi.org/10.18235/0004468>
- International Commission on the Futures of Education. (2021). *Reimagining our futures together: A new social contract for education*. UNESCO. <https://doi.org/10.54675/ASRB4722>
- Kimmons, R., Hunsaker, E. W., Jones, J. E., & Stauffer, M. (2019). The nationwide landscape of K–12 school adoption of open systems: Comparing rural and urban contexts. *The International Review of Research in Open and Distributed Learning*, 20(3). <https://doi.org/10.19173/irrodl.v20i4.3794>
- Li, M. (2024). Exploring the digital divide in primary education: A comparative study of urban and rural mathematics teachers' TPACK and attitudes towards technology integration in post-pandemic China. *Education and Information Technologies*, 30(2), 1913–1945. <https://doi.org/10.1007/s10639-024-12890-x>
- Li, W., Zhang, L., Lee, I., & Gkartzios, M. (2023). Overview of social policies for town and village development in response to rural shrinkage in East Asia: The cases of Japan, South Korea and China. *Sustainability*, 15(14), Article 10781. <https://doi.org/10.3390/su151410781>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage Publications.

- Mulyani, J. D., Festiawan, R., Aryani, A. A., & Agus, H. (2024). An analysis of factors contributing to the teacher shortage in rural public elementary schools in Banyumas Regency. *TEGAR: Journal of Teaching Physical Education in Elementary School*, 8(1), 7–14.
<https://ejournal.upi.edu/index.php/tegar/article/view/73419>
- Mustafa, F., Nguyen, H. T. M., & Gao, X. (2024). The challenges and solutions of technology integration in rural schools: A systematic literature review. *International Journal of Educational Research*, 126, Article 102380. <https://doi.org/10.1016/j.ijer.2024.102380>
- Organisation for Economic Co-operation and Development. (2022). *Education at a glance 2022: OECD indicators*. OECD Publishing. https://www.oecd.org/en/publications/2022/10/education-at-a-glance-2022_4aad242c.html
- Robinson, B. (2008). Using distance education and ICT to improve access, equity, and quality in professional development for rural teachers. *The International Review of Research in Open and Distributed Learning*, 9(1). <https://doi.org/10.19173/irrodl.v9i1.486>
- Sanwal, A. (2024). Marginalization, technology access, and study approaches of undergraduate distance learners during the COVID-19 pandemic in India. *The International Review of Research in Open and Distributed Learning*, 25(2), 1–19. <https://doi.org/10.19173/irrodl.v25i2.7718>
- Trust, T., & Pektas, E. (2018). Using the ADDIE model and Universal Design for Learning principles to develop an open online course for teacher professional development. *Journal of Digital Learning in Teacher Education*, 34(4), 219–233. <https://doi.org/10.1080/21532974.2018.1494521>
- Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. Jossey-Bass/Wiley.



August – 2026

Artificial Intelligence in Online Learning: Using BERTopic to Track Research Topics and Their Evolutions

Xieling Chen¹, Haoran Xie², Xingquan Peng¹, Xiaohui Tao³, Lin Li⁴, S. Joe Qin², and Fu Lee Wang⁵

¹School of Education, Guangzhou University, Guangzhou, China; ²School of Data Science, Lingnan University, Hong Kong SAR;

³School of Mathematics, Physics, and Computing, University of Southern Queensland, Australia; ⁴School of Computer Science and Artificial Intelligence, Wuhan University of Technology, Wuhan, China; ⁵School of Science and Technology, Hong Kong Metropolitan University, Hong Kong SAR

Abstract

Artificial intelligence in online learning (AIOL) has attracted increasing attention in academia. This study applied BERTopic, a transformer-based topic modeling approach that leverages contextual embeddings, to examine the thematic structure and evolution of AIOL research. By analyzing 1,048 AIOL publications, this study sought to answer three research questions. What are the major research topics in AIOL? How have these topics evolved over time? What future research directions are recognized? Several research themes were identified, such as adaptive learning systems, sentiment analysis, and predictive analytics. Temporal analysis revealed a shift from early applications of traditional AI toward machine learning and deep learning approaches. The results also revealed an increasing emphasis on multimodal data integration for emotion recognition and personalized learning support. Based on the findings, a conceptual model has been proposed to guide AIOL research and practice by integrating four components: data, AI processing, adaptive learning, and learner development. Overall, the study offered a data-driven overview of AIOL research and demonstrated how topic modeling can be used to examine thematic development in emerging research fields.

Keywords: artificial intelligence, online learning, BERTopic, research topics, topic evolutions

Introduction

Artificial intelligence (AI), with capabilities in predicting, diagnosing, recommending, and decision-making, has become increasingly prominent in education for supporting learning across various settings (Chen et al., 2022). AI technologies like machine learning (ML) and deep learning (DL) have been integrated into online education to optimize teaching and learning (Torres-Vergara et al., 2025). Tools such as intelligent tutoring systems, instructional robots, learning analytics dashboards, and adaptive learning systems have enhanced student experiences and academic performance (Chen et al., 2020). As a result, research on AI applications in online learning (AIOL) has grown, introducing new thematic directions (Dogan et al., 2023). To understand the development and shifts in research trends, scientific literature, particularly topic models, has offered valuable insights into thematic patterns.

The body of AIOL research has not merely expanded in volume; it has also evolved across interconnected thematic areas, such as AI's functional applications (e.g., prediction, personalization, assessment), methodological approaches (e.g., machine learning algorithms, data mining techniques), and context-specific implementations (e.g., MOOCs, blended learning, language learning). Understanding how these thematic areas have converged and diverged over time is essential for mapping the intellectual structure of AIOL research.

Given the rapid development of AIOL, a review of existing literature is timely. Several studies have identified key themes and roles in AIOL applications. Munir et al. (2022) reviewed 60 publications (2013–2018), highlighting six core themes: (a) intelligent tutoring, (b) dropout prediction, (c) performance prediction, (d) adaptive learning, (e) learning analytics, and (f) automation. Similarly, Ilić et al. (2023) analyzed 305 studies (2010–2021) and identified four AI roles: (a) learner modeling, (b) learning analytics, (c) adaptive assessment, and (d) personalized learning. In terms of AI techniques, Dogan et al. (2023) examined 276 publications (1999–2022) and found three thematic clusters: AI in teaching, student behavior prediction, and personalized learning. They also highlighted the consistent use of AI methods such as neural networks and support vector machines. Ouyang et al. (2022) reviewed 32 studies (2010–2020) and confirmed four key AI functions in online higher education: (a) predicting learning outcomes, (b) recommending resources, (c) automating assessment, and (d) enhancing learning experiences.

Recent studies have explored AI applications in specific subfields of online education. Rather than forming isolated lines of inquiry, these studies have collectively extended core functional themes like prediction, personalization, and assessment into different learning contexts. For example, research in MOOCs has focused primarily on dropout prediction and learner analytics (e.g., Alghamdi et al., 2025), while studies in blended learning have examined AI's role in supporting asynchronous and individualized learning (Park & Doo, 2024). In language learning contexts, particularly English as a foreign language, AI tools like ChatGPT have been studied for their influence on sustained engagement and learning support (Tram et al., 2024).

Similarly, intelligent tutoring systems integrated with emerging technologies like augmented and virtual reality have enhanced immersive and personalized learning experiences (Lampropoulos, 2025). Research on AI-driven assessment has evaluated automated essay scoring and adaptive testing, highlighting both technical progress and functional limitations (Gardner et al., 2021). Studies on AI in learning management systems and broader digital platforms have emphasized adaptive learning, intelligent proctoring, and

outcome optimization (Daniel et al., 2024; Qazi et al., 2024).

These subfield-specific reviews illustrate how similar AI functions—prediction, personalization, adaptive feedback, and automated assessment—have been repeatedly implemented across various educational settings. However, these reviews were typically organized around specific contexts, tasks, or technologies, making it difficult to model cross-context thematic structures. It is unclear how recurring themes intersect, overlap, or evolve at the broader field level.

Methodologically, most existing reviews have relied on qualitative synthesis, manual coding, or small datasets. Conceptually, they described themes but did not quantitatively model how these themes related to one another or evolved over time. Given the rapid growth of AIOL publications, such approaches may have been insufficient for capturing large-scale thematic dynamics.

Compared to traditional qualitative reviews, a quantitative, ML-driven overview that applies automated analysis to a large corpus allows research themes to emerge from data rather than being imposed a priori. This approach has been successfully applied in recent studies, such as Chen et al. (2025), who used topic modeling to identify research themes related to metacognition in education, and Liu et al. (2021), who conducted a topic modeling and trend analysis of MOOCs publications. These studies demonstrated the advantages of ML-driven methods in revealing thematic structures and relationships that are difficult to capture through qualitative reviews. Accordingly, a quantitative, ML-driven overview of AIOL research is both timely and methodologically appropriate.

To this end, this study adopted a topic-based bibliometric approach using BERTopic, a transformer-based topic modeling technique, to analyze 1,048 AIOL publications from 2010 to 2023. Unlike traditional probabilistic topic models like latent Dirichlet allocation (LDA), which rely on bag-of-words representations, BERTopic performs topic discovery using contextual document embeddings derived from pre-trained transformer models (e.g., BERT). Clustering occurs in the embedding space, with bag-of-words representations used only in the final stage to construct interpretable topic representations via c-TF-IDF weighting. BERTopic provides insights into structural patterns across a corpus, enabling researchers to map trends, forecast developments, and make informed decisions about research directions and educational initiatives.

By systematically modeling topic structures and temporal shifts, this study identified current hot topics, emerging directions, and allowed for comparison between data-driven themes and those synthesized in prior reviews. This approach directly addressed the fragmentation noted in earlier literature and offered an integrated, quantitative mapping of the AIOL research landscape.

This study was guided by three research questions:

1. What are the current hot topics in AIOL research?
2. In what ways have AIOL topics evolved over time?
3. What emerging directions are shaping future AIOL research?

Data Collection and Selection

This study followed the PRISMA guidelines (Figure 1) to collect AIOL articles from the Science Citation Index (SCI) and Social Sciences Citation Index (SSCI) databases on the Web of Science platform, using search terms that consisted of three components combined with AND operators (Moher et al., 2010). AI-related terms (e.g., artificial intelligence, machine learning, and neural network*) were adapted and expanded based on prior studies such as Chen et al. (2023) and Chu et al. (2022) to ensure comprehensive coverage of major AI-related terminologies while maintaining search precision. Online learning-related keywords (e.g., online learning, e-learning, MOOC*, and massively open online course*) were informed by Liu et al. (2021). In the search query, the asterisk (*) is used as a wildcard character in the search strategy to capture different word endings and variations of the root term. The full list of search phrases is shown in Figure 1. The data search returned 4,195 journal articles. The first and last authors excluded studies that were duplicated, not authored in English, retrospective analyses, or unrelated to AIOL according to the criteria presented in Table 1, resulting in 1,048 studies published between 2010 and 2023 for analysis. The titles, keywords, and abstracts of the selected papers were consolidated into a single document for topic modeling.

Figure 1

Data Search and Collection Process

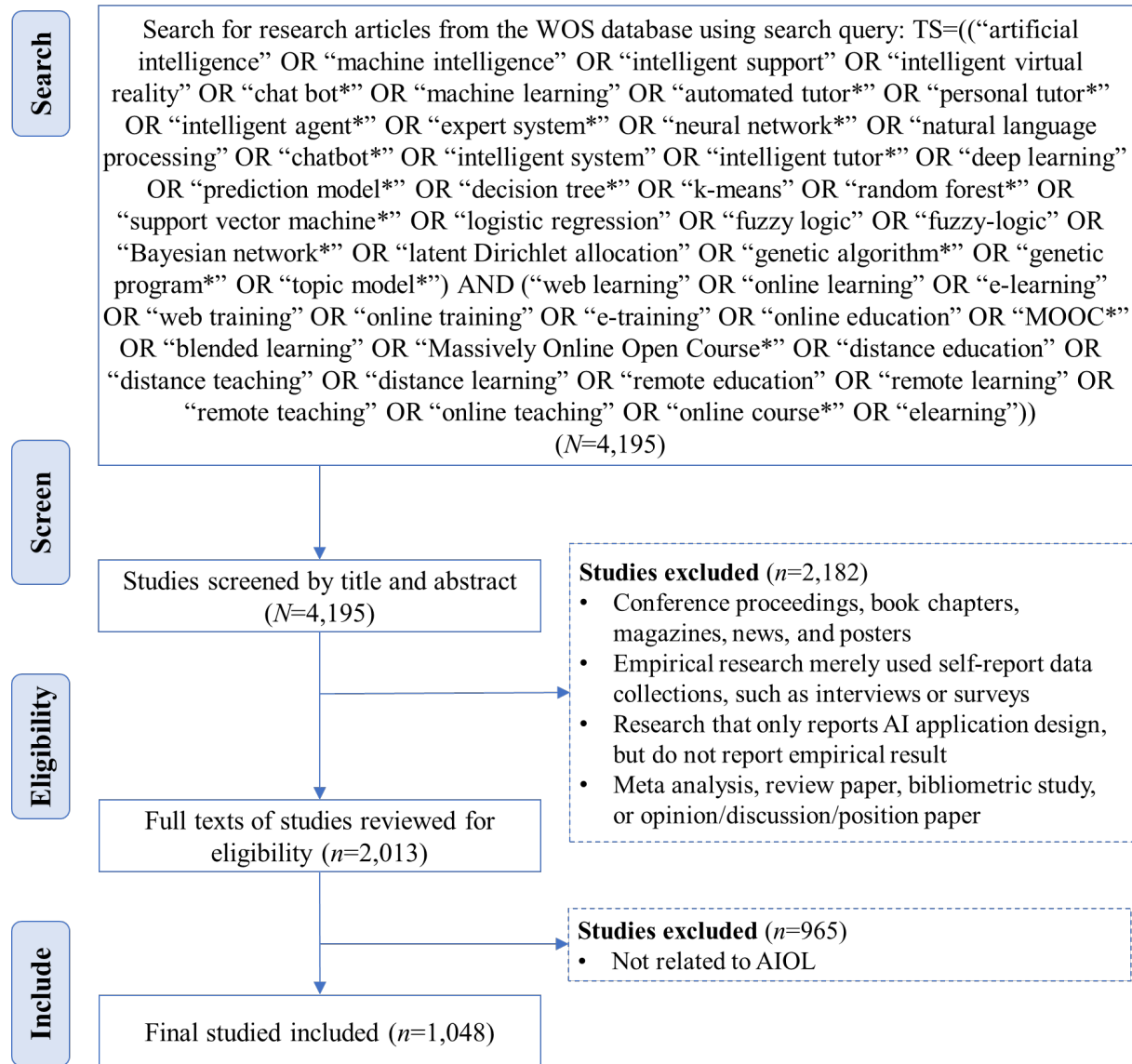


Table 1

Inclusion and Exclusion Criteria

Criteria	Code	Meaning
Inclusion	I1	Conducted in online education
	I2	Reported the actual AI-supported instruction and learning processes
	I3	Published in peer-reviewed journals
	I4	Empirical research that demonstrated AI's actual effects on education
Exclusion	E1	Conference proceedings, book chapters, magazines, news, or posters
	E2	Only used self-report data collections (e.g., interviews or surveys)
	E3	Only reported AI application design without empirical results
	E4	Meta-analysis, review papers, bibliometric study, or opinion/discussion/position papers

BERTopic Modeling Analysis

This study employed BERTopic, a framework that integrates contextual document embeddings with clustering techniques for topic modeling. Unlike traditional probabilistic models such as latent Dirichlet allocation, which rely on bag-of-words representations, BERTopic clusters documents in an embedding space derived from pre-trained transformer models (e.g., BERT). Topic discovery occurs in the embedding space, while bag-of-words representations are used at the topic representation stage to extract interpretable keywords via class-based term frequency-inverse document frequency (c-TF-IDF). Thus, topic formation via BERTopic was driven by embeddings rather than word frequency statistics. The modeling process consisted of five steps as described below: (a) document embedding, (b) dimensionality reduction, (c) clustering, (d) cluster-level bag-of-words generation, and (e) topic representation. The workflow is shown in Figure 2.

In step 1, embed documents, the abstracts of AIOL research articles were transformed into numerical representations using the sentence-transformers framework with the all-MiniLM-L6-v2 English language model. This model mapped sentences and paragraphs into a 384-dimensional dense vector space. These embeddings represented the contextual semantic relationships between documents and served as the basis for clustering.

Step 2, reducing dimensionality, was performed using uniform manifold approximation and projection (UMAP), a non-linear technique that preserved both local and global data structure (McInnes et al., 2018). Three parameters were considered: (a) `n_neighbors` (number of nearest neighbors for approximating local structure); (b) `n_components` (number of dimensions in the reduced embedding space); and (c) `min_dist`

(minimum distance between data points in the low-dimensional space). Default BERTopic parameter settings were retained unless otherwise specified.

In step 3, cluster documents. After dimensionality reduction, density-based clustering and outlier detection were conducted using hierarchical density-based spatial clustering of applications with noise (HDBSCAN). This method identified clusters of varying shapes and outliers to improve topic representation by reducing noise and enhancing accuracy. Clustering occurred in the embedding space, so topic allocation reflected semantic proximity rather than keyword overlap. The HDBSCAN algorithm included two key parameters: `min_cluster_size` (which controlled the minimum number of samples per cluster) and `min_samples` (which influenced noise point identification by setting a threshold for the minimum number of samples to avoid classification as noise).

Step 4 was to generate bag-of-words. After clustering, all documents within each cluster were aggregated into a single document. Word frequencies were computed to create a cluster-level bag-of-words representation. Initially, the default stopwords list provided by BERTopic was used. However, as the initial modeling results showed limited interpretability, the stopwords list was expanded by incorporating widely used English stopwords from the natural language toolkit (NLTK). The final stopwords list was a merged set of BERTopic's default stopwords and the extended NLTK list, with duplicates removed. The resulting vectors were L1-normalized to account for differences in cluster size. It is important to note that the bag-of-words representation was used only to enhance topic interpretability; it did not determine cluster formation, which was based exclusively on contextual embeddings.

Step 5, represent topics using class-based term frequency-inverse document frequency (c-TF-IDF). The c-TF-IDF score for a word x in cluster c was calculated as: $W_{x,c} = d = tf_{x,c} \times \log(\frac{1+A}{f_x})$, where $tf_{x,c}$ denoted the word frequency of f_x in cluster c , A represented the average number of words per cluster, and f_x denoted the total word frequency across all clusters. Words within each cluster were ranked according to their c-TF-IDF scores. In this study, the top 20 words ($N = 20$) with the highest c-TF-IDF values were selected to represent each topic. It is important to note that the computationally derived topics do not aim to replicate author-provided keywords. Instead, they represent emergent semantic structures at the corpus level.

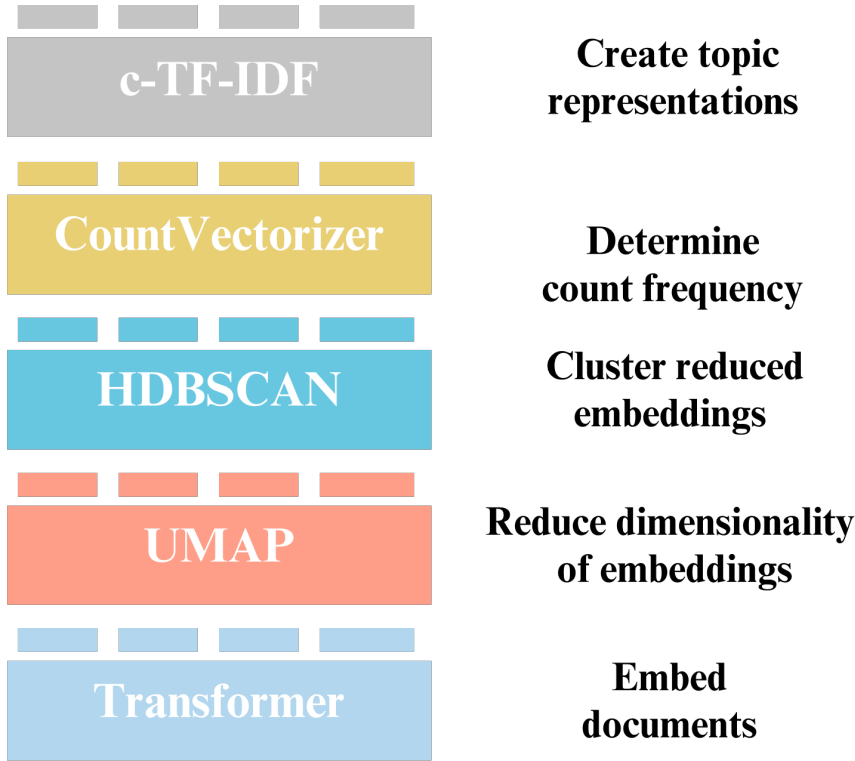
During topic modeling, comparisons across different models with varying numbers of topics were conducted by two domain experts. These experts evaluated the models based on representative terms and articles according to four criteria adapted from Chen et al. (2025).

- Representative terms within each topic should form a meaningful and coherent topic.
- Top representative articles should be closely related to the topic.
- Minimal overlap should exist between topics within the same model.
- All important AIOL topics should be included.

Based on this evaluation, a total of 11 distinct and meaningful topics were identified.

Figure 2

Workflow For BERTopic Modeling



Results

Research Topics and Trends

Table 2 shows the BERTopic modeling results for the 1,048 AIOL research articles, from which a total of 11 distinct and meaningful topics were extracted. The largest topic was machine learning for online predictive analytics, while the smallest significant topic was optimization algorithm-based personalized online learning. In Table 2, the topic proportion was calculated as $P_k = (\sum_d \theta_{d,k} / D)$, where P_k represented the prevalence of the k th topic, and $\theta_{d,k}$ was the probability of document d being associated with the k th topic.

Table 2

AIOL Research Topics

Topic	Proportion	<i>S</i>	<i>p</i>	Trend
Machine learning for online predictive analytics	17.50%	37	0.04874	↑↑
MOOC forum and review analysis	10.44%	41	0.02854	↑↑
Intelligent adaptive online learning systems	12.77%	-79	1.954e-05	↓↓↓↓
Emotion recognition and deep learning in online learning	9.22%	29	0.1253	↑
AI-enhanced online language education	10.22%	13	0.5112	↑
Online learning in COVID-19	7.47%	51	0.006196	↑↑↑
Self-regulated online learning	10.34%	37	0.04874	↑↑
AI chatbots-assisted online learning	6.27%	31	0.1005	↑
AI-assisted online music and art education	5.72%	33	0.0798	↑
Electroencephalogram (EEG)-based online learning analytics	3.69%	37	0.04874	↑↑
Optimization algorithm-based personalized online learning	6.37%	-65	0.0004589	↓↓↓↓

Note: ↑(↓): increasing (decreasing) trend but not significant ($p > 0.05$)

↑↑(↓↓), ↑↑↑(↓↓↓), and ↑↑↑↑(↓↓↓↓): significantly increasing (decreasing) trend ($p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively).

Figure 3 illustrates the proportion of each topic in the dataset, highlighting their relative prevalence to emphasize the dominant themes in the AIOL research field. Figure 4 presents the temporal evolution of each topic, identifying emerging, declining, or stable research directions over time. Significant topics included machine learning for online predictive analytics, MOOC forum and review analysis, online learning in COVID-19, self-regulated online learning, and EEG-based online learning analytics, all of which have seen notable increases in research interest. In contrast, two topics, namely intelligent adaptive online learning systems and optimization algorithm-based personalized online learning, have shown a significant downward trend.

Figure 3

Distribution of Topics Among Articles

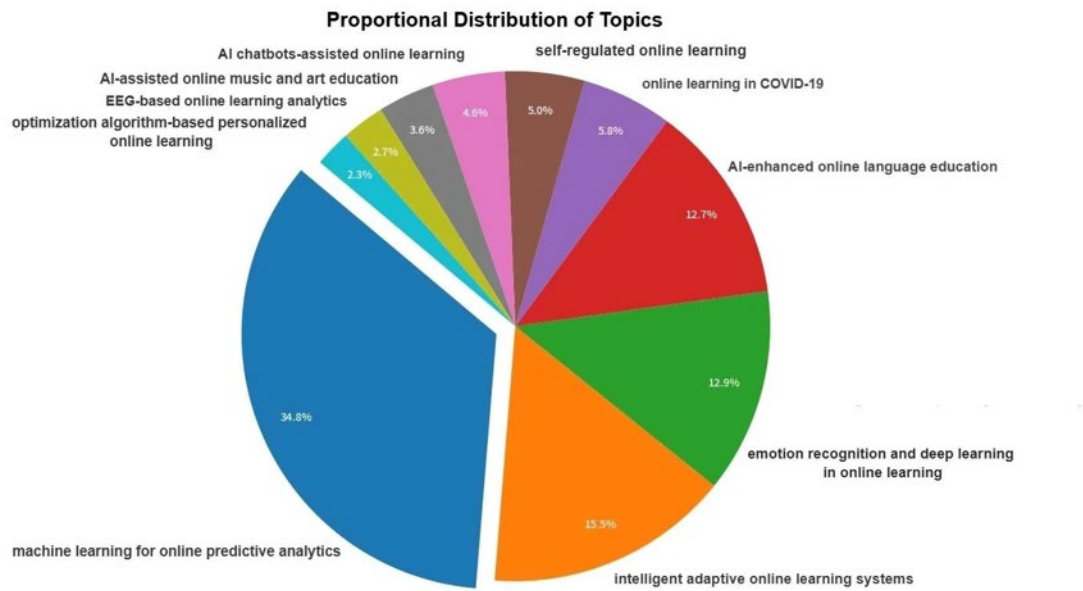


Figure 4

Topic Trends

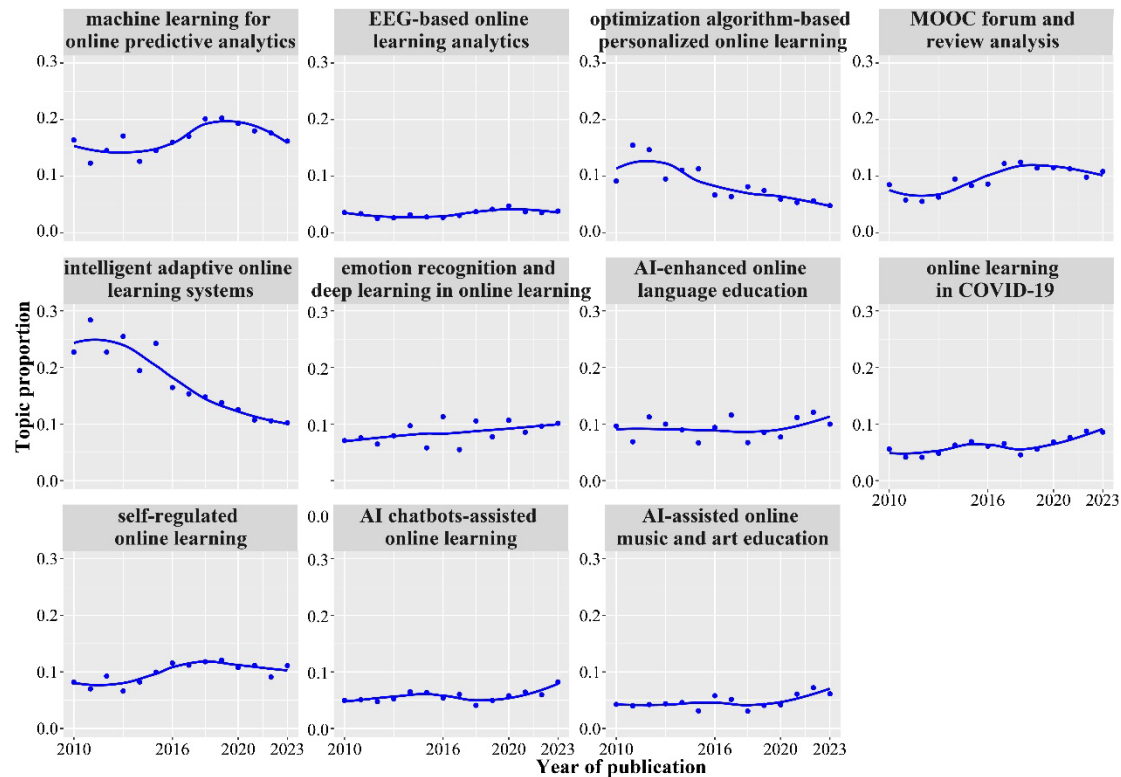
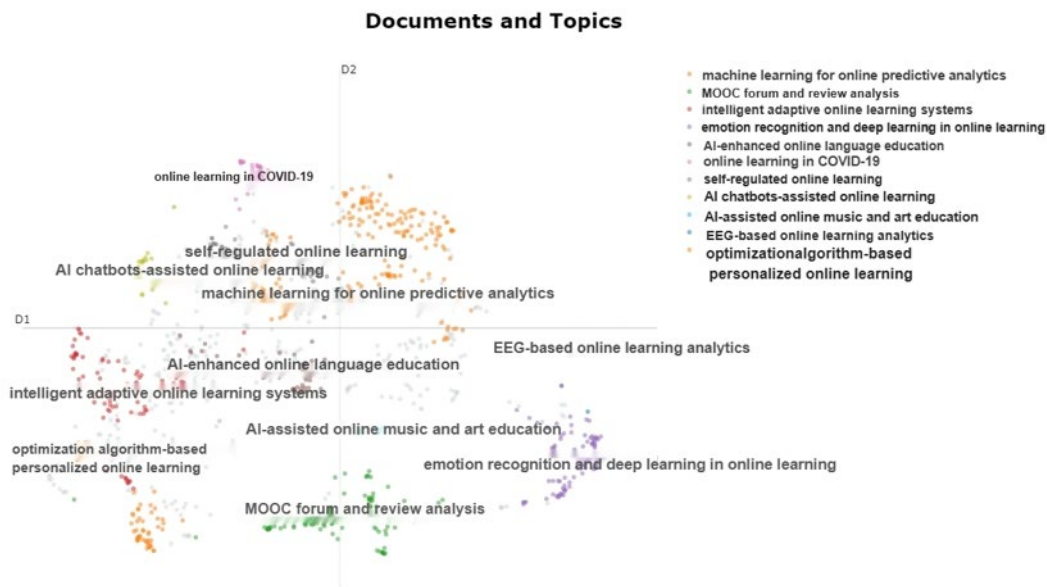


Figure 5 shows the distribution of AIOL research articles. Each dot represents an article, and the color indicates the topic to which it belongs, providing a deeper understanding of the relationships between topics in a two-dimensional space. The axes, D1 and D2, were derived from the UMAP dimensionality reduction technique, which projected the high-dimensional topic space into two principal components. These components captured the most important patterns and relationships in the data, allowing for a visualization of how articles were distributed based on their topic associations. Articles on machine learning for online predictive analytics (orange) were concentrated, indicating content similarity, while those on MOOC forum and review analysis (green) formed a distinct cluster. Articles on online learning in COVID-19 and self-regulated online learning were closely grouped, suggesting content overlap. In contrast, articles on EEG-based online learning analytics were more dispersed, reflecting diversity in research methods or applications.

Figure 5

Distribution of AIOL Articles

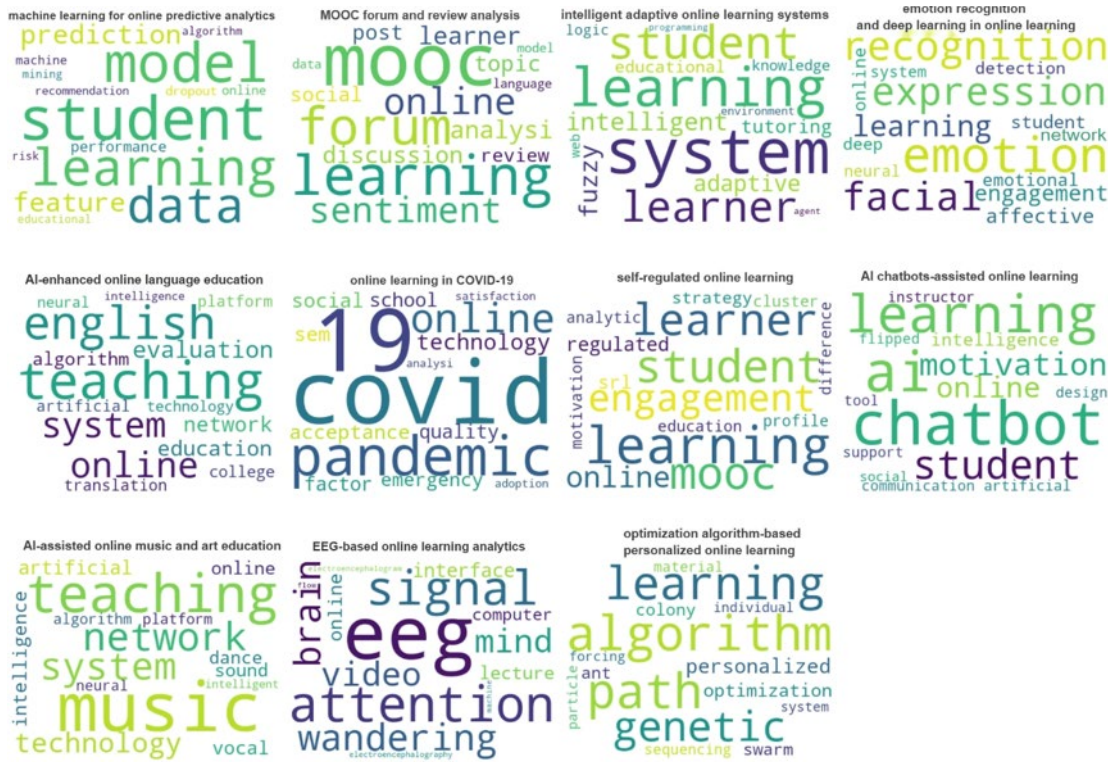


Terms Used Most Frequently

Figure 6 depicts the terms used most frequently, providing an intuitive overview of the key concepts within the dataset. The term student appeared most frequently, with other commonly occurring terms including e-learning, model, education, use, and learning. A key finding from the analysis was the prevalence of terms related to AI-supported online learner engagement, such as MOOC, method, system, model, machine learning, and neural network. These terms reflected a focus on online learning platforms, instructional methods, learning systems, and AI applications. Figure 7 presents word clouds for each topic to complement the BERTopic analysis by visually highlighting the most salient terms within each identified cluster. While

Figure 7

Specific Topic Word Clouds

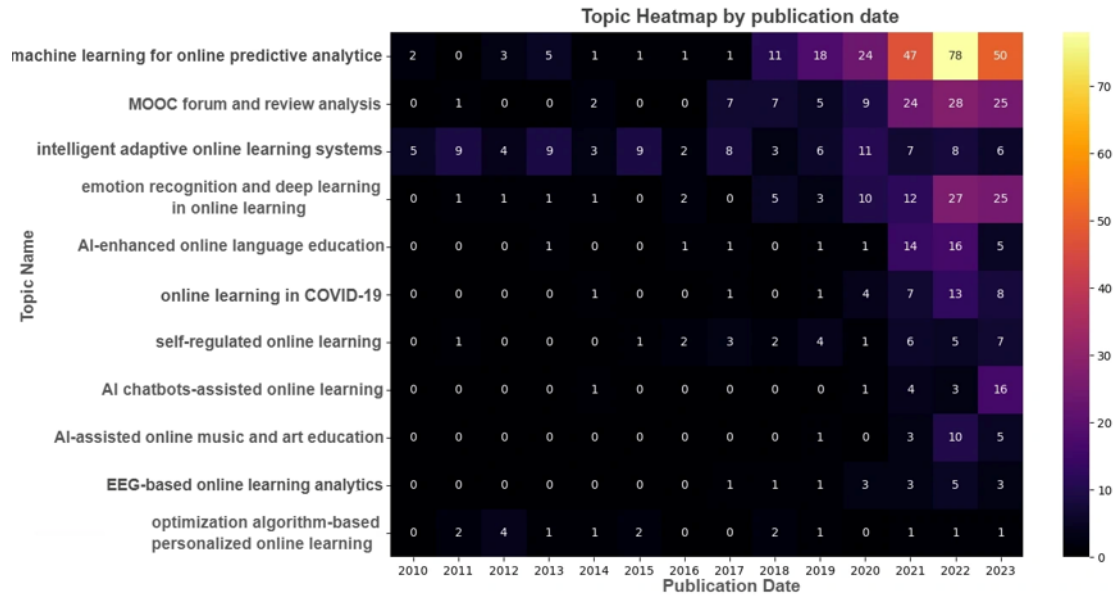


Topic Heat and Intensity

Figure 8 shows a heatmap of the number of articles for each topic by publication year, highlighting periods of heightened research activity. Overall, AIOL research has shown a continuous upward trend, with an increasing number of articles published, particularly since 2017. Six out of the eleven topics reached their peak in volume by 2022. It should be noted that the total number of articles presented in the figure does not equal the total analyzed (1,048) due to how the BERTopic algorithm handles clustering. When using the HDBSCAN algorithm, points that do not belong to sufficiently dense clusters are assigned the label -1, which represents noise points. Following Grootendorst (2022), this study excluded articles labeled as -1.

Figure 8

Heat Map of AIOL Topics

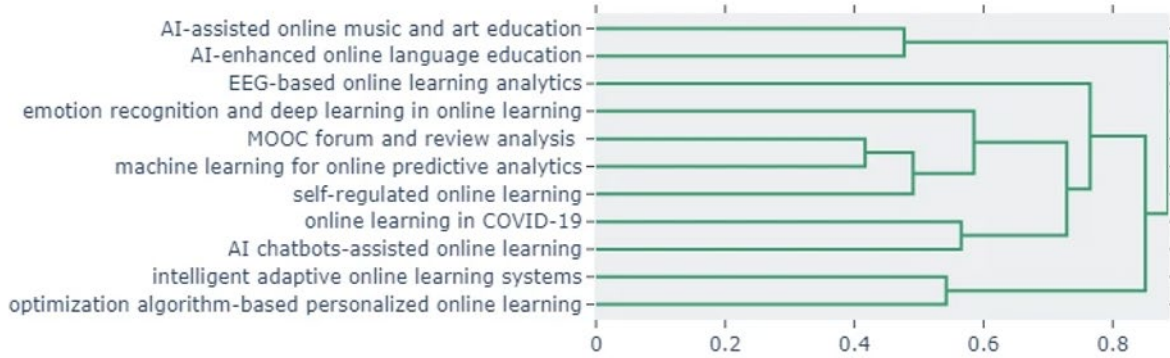


Topic Clustering

Figure 9 presents the hierarchical clustering analysis of AIOL topics, visualizing their relationships and similarities through a dendrogram. This analysis revealed how different topics were related and grouped, highlighting potential overlaps and shared themes. It provided deeper insights into the underlying structure of the topic distribution over time. Results showed that machine learning for online predictive analytics clustered with MOOC forum and review analysis and self-regulated online learning, which indicated high content relevance. Intelligent adaptive online learning systems tightly clustered with optimization algorithm-based personalized online learning, which reflected similarities in personalized learning technologies. AI-assisted online music and art education grouped with AI-enhanced online language education and EEG-based online learning analytics, which suggested common AI applications. Lastly, online learning in COVID-19 clustered with AI chatbots-assisted online learning, which highlighted the rise of chatbot-enhanced education during COVID-19.

Figure 9

Hierarchical Clustering Analysis Results



Developmental Stages of AIOL Research

Figure 10 synthesizes the annual article counts to depict the developmental stages of AIOL research, from emerging to mature phases. Results indicated that AIOL research is still in a phase of rapid development. The whole period of AIOL research can be divided into two stages: a stable development period (2010–2016) and a rapid development period (2017 to the present). To analyze the topic differences and evolution during these two developmental stages, this study modeled the articles from each period separately; the results are summarized in Table 3.

Figure 10

Developmental Stages of AIOL Research

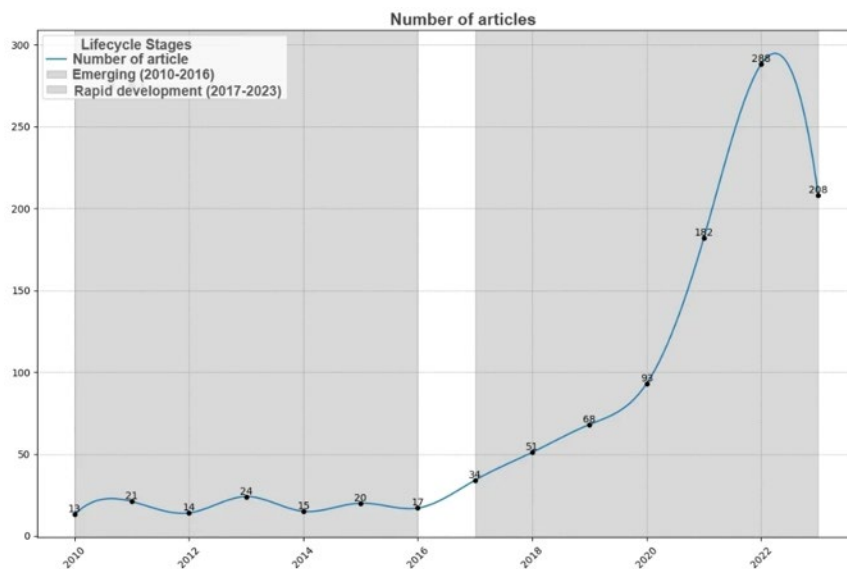


Table 3

Topics From the Two Periods of AIOL Research

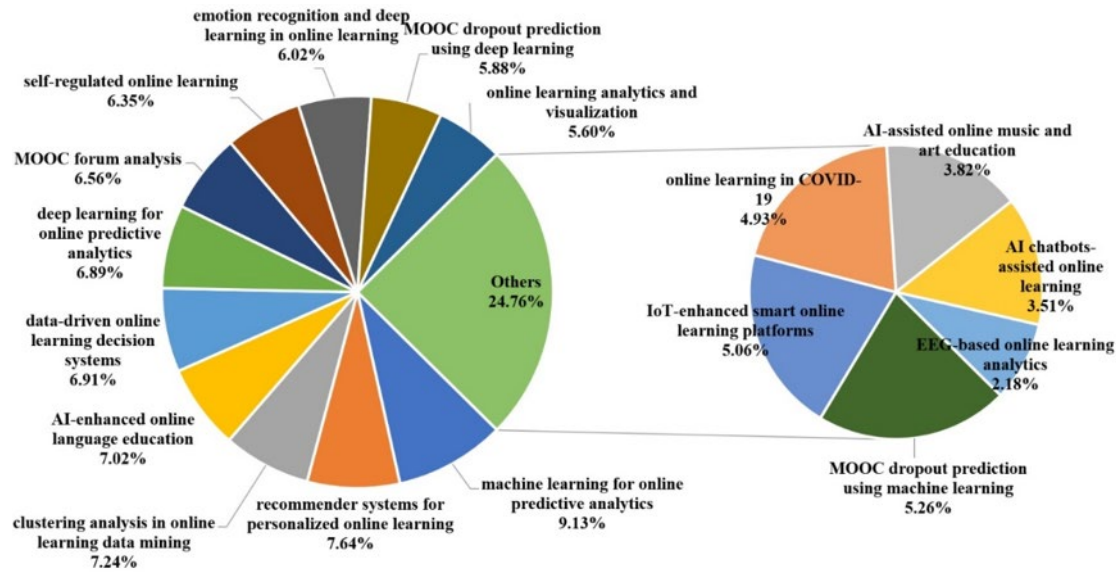
Stage	Topic name
Stable development (2010–2016)	Intelligent online learning systems
	Machine learning for online education
	Personalized online learning
Rapid development (2017 to the present)	Recommender systems for personalized online learning
	MOOC forum analysis
	Machine learning for online predictive analytics
	Emotion recognition and deep learning in online learning
	AI-enhanced online language education
	Online learning in COVID-19
	Self-regulated online learning
	Clustering analysis in online learning data mining
	AI chatbots-assisted online learning
	AI-assisted online music and art education
	MOOC dropout prediction using machine learning
	MOOC dropout prediction using deep learning
	EEG-based online learning analytics
	Online learning analytics and visualization
	Data-driven online learning decision systems
	Deep learning for online predictive analytics
	IoT-enhanced smart online learning platforms

Research Topics and Trends in the Rapid Development Period

Figure 11 summarizes the AIOL research topics in the rapid development period, with the top three topics being machine learning for online predictive analytics, recommender systems for personalized online learning, and clustering analysis in online learning data mining.

Figure 11

Topics in the Rapid Development Period

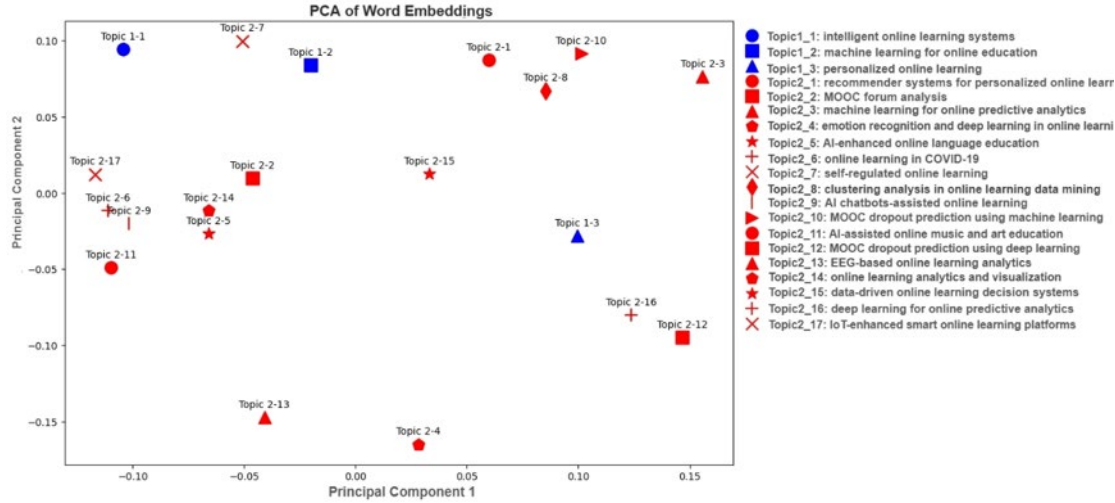


Evolution of AIOL Research Topics

This study used word embedding vectors generated by the all-MiniLM-L6-v2 model to represent topic vectors. The embeddings of top topic words were averaged to obtain the semantic representation of the topic; subsequently, principal component analysis was adopted for dimensionality reduction to visualize the distribution of topics (Figure 12), with the Euclidean distance between the reduced points representing the path of topic evolution.

Figure 12

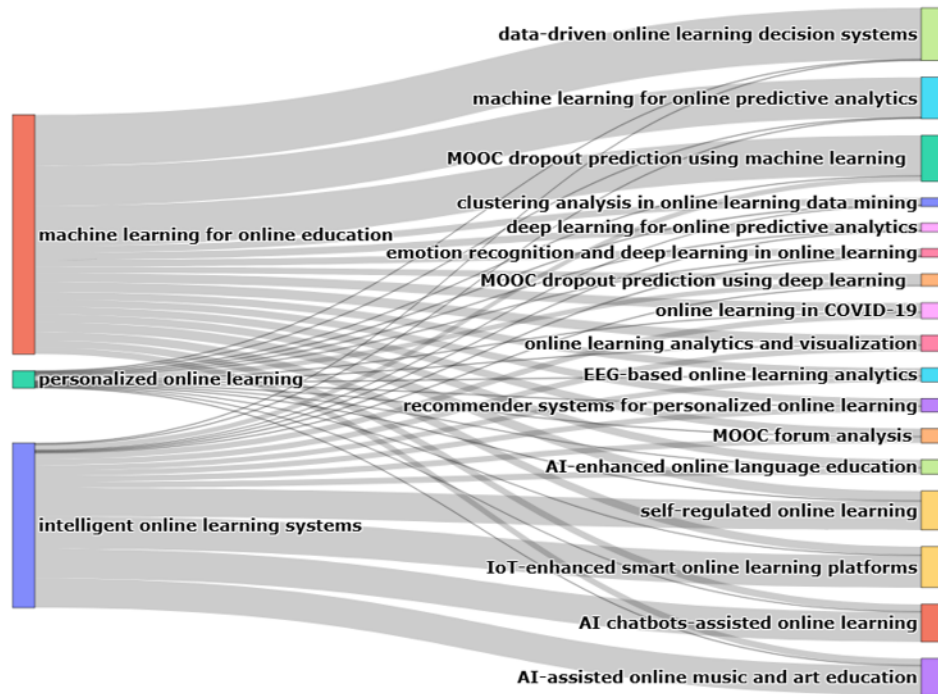
Distribution of Topics



The evolutionary relationship between topics from the two periods was determined by calculating the inverse of the Euclidean distance between them, with a larger value indicating a stronger evolutionary relationship between two topics. To intuitively represent the evolutionary path of AIOL research, this study calculated the inverse of the Euclidean distances between all topics and used them as evolutionary weights to create a Sankey diagram (Figure 13), where the thickness of the connections between topics from different periods indicated their evolutionary weight, with thicker connections representing closer evolutionary relationships. One of the most prominent patterns was the strong expansion of machine learning for online education into a wide variety of specialized subtopics in the later period, including (a) data-driven decision systems, (b) predictive analytics, (c) MOOC dropout prediction, (d) clustering and data mining approaches, and (e) deep learning-based analytics. The thick flows indicated that the early foundational focus on machine-learning methods acted as a platform for more intricate applications aimed at understanding learner behavior, forecasting outcomes, and improving educational decision-making.

Figure 13

Evolution of AIOL Topics in the Two Stages



Discussion

Analysis Results

The BERTopic analysis (Table 2) revealed recurring themes in AIOL research, notably machine learning for online predictive analytics and self-regulated online learning. Trend tests from 2010 to 2023 showed sustained growth in AI's role in supporting learner performance and self-regulation. Figures 3 and 4 highlighted increasing interest in interactive platforms, pandemic-driven online learning, and multimodal data analysis. In contrast, the decline of intelligent adaptive online learning systems and optimization algorithm-based personalized learning suggested a shift from system-centric approaches to data-driven, learner-centered AI support. These findings reflected a transition toward more applied, learner-focused, and technologically advanced research.

In online environments, predictive analytics using machine learning models learner engagement, predicts outcomes, and assesses dropout risks, addressing challenges in the absence of face-to-face interaction (Sjølie et al., 2022). SRL, essential in flexible, learner-centered environments, has remained a critical area despite a slight decline in recent years (Figure 4), as confirmed by trend analysis (Kim & Kumi-Yeboah,

2025). AI tools like intelligent feedback systems and personalized trackers have helped foster motivation, self-efficacy, and metacognitive awareness (Jin et al., 2023; Shi & Zhang, 2025).

Emerging themes such as MOOC forum and review analysis, AI-enhanced online language education, and AI chatbots-assisted learning indicated a shift from traditional content delivery to interactive learning. Figures 6 and 7 highlight AIOL's focus on computational modeling and data-intensive analytics, with studies analyzing forum interactions and AI-driven language tools, such as chatbots, for personalized learning (Ruyang et al., 2025; Sharma et al., 2024). The rise of emotion recognition also pointed to integrating multimodal data, addressing learners' cognitive and emotional states.

The growing focus on online learning in COVID-19 reflected the shift to remote education, with research on adaptive learning systems and AI supporting large-scale remote learning during the pandemic (Karakaya, 2021). In contrast, the decline of intelligent adaptive systems reinforced the shift toward emotion-aware, data-driven AI support. The adoption of large language models and conversational agents illustrated a move toward learner-responsive AI capable of addressing diverse needs (Bahari & Liu, 2025).

Compared to dominant themes synthesized in prior reviews, the data-driven topics identified in this study showed both convergence and extension. Previous reviews (e.g., Alghamdi et al., 2025; Daniel et al., 2024; Gardner et al., 2021; Lampropoulos, 2025; Qazi et al., 2024;) consistently highlighted predictive analytics, adaptive learning, learner modeling, and automated assessment as central areas of AIOL research. Our findings largely confirmed these conclusions, with several prominent topics aligning closely with these established categories. However, the topic modeling approach further refined these broad themes by revealing more nuanced substructures. For instance, while earlier reviews like Gardner et al. (2021) often grouped assessment-related applications together, our analysis differentiated among automated grading, AI-generated feedback, and adaptive testing systems, indicating increasing specialization in the field. Additionally, emerging themes related to generative AI technologies, which were less prominent in earlier reviews, pointed to a shift in research focus from predictive modeling toward interactive and generative AI applications in online learning contexts.

The analysis highlighted the interdisciplinary nature of AIOL research, linking fields like computer science, pedagogy, psychology, and linguistics. Hierarchical clustering (Figure 9) showed strong connections across areas like AI-assisted online music education and EEG-based learning analytics. Developmental trends (Figure 10) revealed AIOL's rapid evolution, with core themes consolidating while new areas emerge (Table 3). This underscored the need for greater collaboration across disciplines, as AI integration into online learning requires a holistic understanding of both technological and human factors.

Technological Evolutions

The emerging trends identified reflected a leap forward in both analytical capability and pedagogical potential of AI through three major stages. Early AI systems were largely rule-based or expert systems that demonstrated high transparency and interpretability. However, they offered limited adaptability. This is because these systems relied mainly on predefined rules and simple decision trees to provide basic guidance and personalization, often in the form of automated quizzes or instructional content suggestions; however, they lacked the ability to learn from data and adapt over time. According to Figures 4 and 9, topics such as

machine learning for online predictive analytics have dominated research since 2017, reflecting increasing adoption of data-driven approaches for modeling learner behavior, engagement, and performance.

Advances in machine learning have enabled data-driven predictions of learner behavior, engagement, and performance, improving retention and outcomes in large-scale online courses. Recent applications of deep learning and multimodal data analysis allow AIOL systems to process diverse data (e.g., text, video, audio, and physiological signals), enabling emotion recognition, cognitive load monitoring, and automated feedback. The integration of AI with immersive technologies like AR and VR have created more interactive, personalized, and dynamic learning experiences.

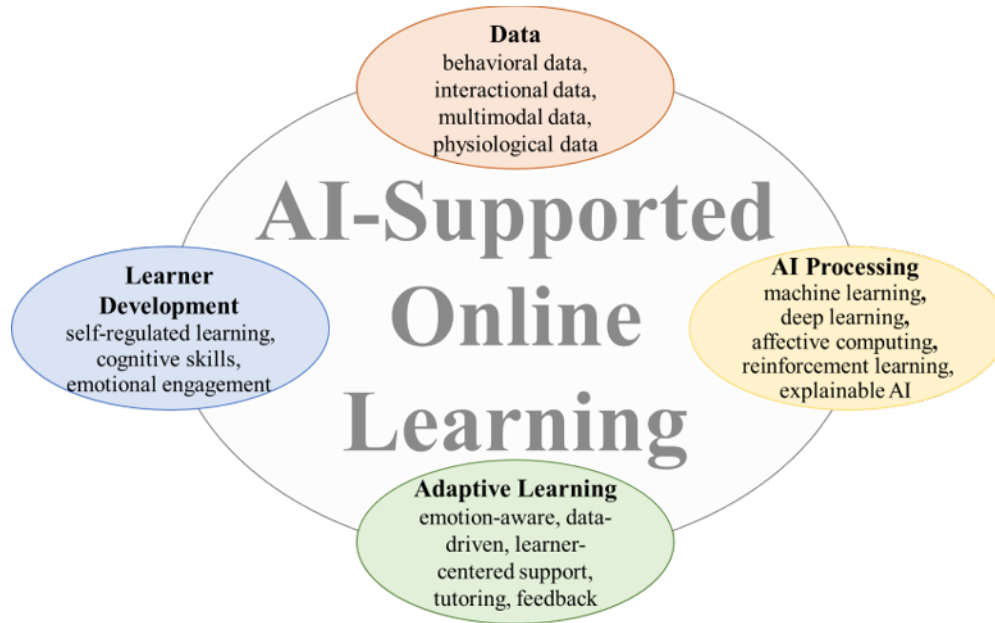
According to Figures 8 and 9, AIOL research has shown a clear shift from traditional behavioral data modeling (e.g., clickstreams and quiz scores) toward multimodal learning analytics that integrate both cognitive and affective dimensions to support real-time adaptation and personalization (Becerra et al., 2025). This form of adaptation emphasizes learner-centered AIOL systems that provide emotion-aware and data-driven support to enhance both academic achievement and emotional well-being, aligning with the observed growth in topics related to emotion recognition, EEG-based analytics, and SRL.

Conceptual Model of AIOL

Based on the empirical findings from the BERTopic analysis, this study proposed a conceptual model to guide future AIOL research and AI-powered online learning practices (see Figure 14). The model is a synthesis of the dominant and emerging thematic trends identified in the results, rather than an independently theorized framework. It integrates four interrelated components, namely data, AI processing, adaptive learning, and learner development.

Figure 14

Conceptual Model of AIOL



The emphasis on multimodal and physiological data within the data component was directly informed by the observed growth of topics such as emotion recognition, deep learning, and EEG-based online learning analytics. These findings indicated a shift from simply behavioral tracking to richer representations of learners' cognitive and affective states. Learner data, including (a) behavioral data (e.g., quiz scores and clicks); (b) interactional data (e.g., forum participation); (c) multimodal data (e.g., video, audio, eye-tracking); and (d) physiological data (e.g., heart rate or brainwave patterns) have been integrated into AI systems to capture not only what learners do but also how they think and feel.

Building on these data sources, the AI system applies machine learning for predictive analytics and deep learning to process complex inputs. The prominence and sustained relevance of predictive analytics in our topic modeling results provided the empirical basis for positioning prediction as a core function of the AI processing component. In addition, affective computing techniques have been incorporated to recognize learners' emotional states, reflecting the significant upward trend in emotion-aware research themes identified in the analysis.

Explainable AI (XAI) was included in the model as a cross-cutting design principle rather than as a distinct empirical topic. Although XAI did not emerge as an independent theme in the modeling results, its inclusion responds to the growing reliance on AI-driven predictions and adaptive decisions identified in our findings. In this context, explainability served as a transparency mechanism to enhance trust, accountability, and informed learner engagement.

The insights generated by the AI are used to deliver adaptive support. Consistent with the declining trend of traditional system-centric adaptive systems observed in the results, adaptation in this model has been

conceptualized as data-driven and emotion-aware support, rather than rule-based pathway optimization. Examples include feedback aligned with learners' emotional states, performance-informed resource recommendations, and real-time scaffolding.

Finally, these adaptive supports aim to foster learner autonomy and self-regulation. The statistically significant long-term growth of self-regulated online learning identified in the trend analysis provided the empirical foundation for positioning learner development as the ultimate outcome of the model. Through continuous feedback loops, both the AI system and the learner co-adapt over time.

Overall, the proposed model represents a synthesis of empirically identified thematic trajectories and illustrates how data-driven insights can inform the design of future AI-supported online learning systems.

Implications for Future AIOL Research and Pedagogy

Based on the findings, this study has provided several implications for future AIOL research.

- Improve multimodal learning analytics to integrate behavioral, interactional, affective, and physiological data for learner modeling.
- Develop transparent and explainable AI models to enhance instructor trust in AI-informed decision-making.
- Strengthen interdisciplinary collaboration across education, psychology, computer science, and cognitive science to advance AIOL.
- Expand AI research beyond academic subjects to skill-based domains such as arts and sports, where motor, creative, and embodied skills require AI-driven assessment and feedback.
- Investigate ethical and privacy protection challenges in multimodal and physiological data collection and analytics.
- Develop real-time adaptive systems based on reinforcement learning and multimodal data to improve personalization.
- Examine the long-term impact of AI feedback and dashboards on SRL abilities such as metacognition, motivation, and autonomy.

This study also provided implications for future AIOL pedagogical practice.

- Support instructors in interpreting AI outputs through training on data literacy and explainable AI tools.
- Use AI-informed adaptive decisions to tailor difficulty, pacing, feedback, and recommendations to personalize individual learning.
- Strengthen SRL by integrating AI-based prompts, reflective tools, and intelligent coaching into course design.
- Exploit emotion detection and sentiment analysis for affective support and timely pedagogical interventions.

- Improve feedback quality and efficacy through AI chatbots and autograders that offer immediate clarification and guidance.
- Promote student engagement in online learning communities based on AI tutors and automated identification of confusion.
- Use AI-enhanced multimodal tools (e.g., VR/AR, motion capture) to support practical and performance-based learning in arts and sports.
- Ensure transparency with students about data collection and usage, and establish norms for privacy protection and algorithmic fairness.
- Improve pedagogical design that emphasizes human-AI collaboration where instructors retain agency while AI augments decisions.

Limitations, Reflections, and Future Work

This study had some limitations. First, we focused on SCI/SSCI journal publications on AI-supported instructional and learning processes in online education, excluding position papers, opinions, and design-oriented works, which often lack empirical evidence. We also excluded survey- and interview-based studies to maintain focus on observable learning processes and AI-driven instructional mechanisms grounded in behavioral or system-generated data. Including self-report studies, which examine perceptions and attitudes toward AI, would have broadened the scope to perception-oriented research, outside the study's process-centered focus. This may explain the underrepresentation of social constructivist or community-oriented research, where outcomes are harder to capture. Future research could separately analyze self-report studies to compare perception- and process-oriented research trajectories, offering a more comprehensive understanding of AIOL scholarship. Additionally, the study covered publications from 2010 to 2023, a period selected for bibliometric analysis, ensuring stable indexing records and citation data. Including more recent publications, still undergoing updates, would have affected the stability and comparability of longitudinal analyses. Future research could incorporate newer publications once their bibliographic and citation data stabilize. For instance, Dağhan and Gündüz (2022) reviewed studies from 2000 to 2018. Chen et al. (2025) also noted that including recent publications could impact the accuracy of trend identification due to fluctuating citation counts. Lastly, search-term selection involved trade-offs between recall and precision. Future studies may expand datasets using advanced screening strategies to manage larger retrieval sets and broaden coverage.

From a methodological perspective, this study did not use keyword-based analyses due to limitations in the availability and precision of keywords, as some articles lack author keywords or use predefined lists that may not fully capture the study's scope. Instead, embedding-based topic modeling was used to analyze contextual relationships across titles, keywords, and abstracts, which better summarized core research focuses, suitable for large-scale bibliometric studies (e.g., Chen et al., 2025; Liu et al., 2021). During modeling, each document was assigned to a single dominant topic based on embedding similarity. While this approach provided a macro-level analysis, it may have oversimplified the multidimensional nature of scholarly articles. Future studies could explore multiple topic assignments per article to better capture topic overlap. As with any unsupervised method, the topics identified may be sensitive to hyperparameters (e.g., `min_cluster_size`) and domain-specific terminology. To address this, we used careful preprocessing,

parameter tuning, and expert validation, following established procedures in previous studies (e.g., Chen et al., 2025; Liu et al., 2021). Additionally, BERTopic's data-driven nature means that research themes can only emerge as distinct topics when they form coherent clusters. Although socially oriented studies on AI-supported collaborative learning and community development exist, they did not emerge as dominant topics in this dataset, reflecting the current focus on individual-level analytics in AIOL. Future research may explore more socially oriented AI-supported learning designs. Finally, while some identified domains (e.g., MOOCs, language learning, COVID-19 education) have been popular in broader distance learning literature, the topics in this study specifically reflected how AI methods are integrated into these contexts. Future studies could conduct comparative analyses with broader online learning corpora to distinguish AI-specific trends from general pedagogical developments.

Conclusion

This topic-based bibliometric study traced the evolution of AIOL from 2010 to 2023 using BERTopic, a topic modeling framework that integrates contextual embeddings with clustering techniques. Analyzing a large corpus of publications, the study identified major thematic patterns and shifts in research focus. Unlike traditional probabilistic models like LDA, which use bag-of-words representations, BERTopic operates in a semantic embedding space, allowing for more accurate document-level thematic identification. The findings showed progression from early AI applications to more sophisticated ML and DL approaches, with emerging hotspots in predictive analytics, sentiment recognition, adaptive learning systems, and EEG monitoring, highlighting a focus on data-driven personalization and learner modeling. Current research directions include enhancing adaptive learning systems, supporting SRL, and integrating multimodal data to capture learners' cognitive and emotional states through DL and affective computing. While the study did not assess the effectiveness of these approaches, it identified areas of sustained research interest. It also underscored the importance of developing adaptive, transparent, and interpretable AI algorithms and suggested that future research could explore AI applications beyond traditional academic subjects, such as the arts and sports. By combining bibliometric analysis with BERTopic, this study offered a macro-level overview of AIOL research trends, illustrating how embedding-based topic modeling complements qualitative reviews. The findings should be interpreted within the methodological limits of computational text analysis.

Conflict of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Acknowledgement

This work was fully supported by the National Natural Science Foundation of China (No. 62307010), the Fund for Innovative Technology-in-Education (FITE) for UGC Funded Inter-institutional Collaborative Activities Project (120045) entitled "Advancing Digital Competency for University Teachers and Students in the Era of Generative Artificial Intelligence", and the Teaching Development Grant (102729) entitled "Integrate GPT-based Intelligent Feedback and Tutoring for Data Science Courses" of Lingnan University, Hong Kong.

References

- Alghamdi, S., Soh, B., & Li, A. (2025). A comprehensive review of dropout prediction methods based on multivariate analysed features of MOOC platforms. *Multimodal Technologies and Interaction*, 9(1), 3. <https://doi.org/10.3390/mti9010003>
- Bahari, A., & Liu, Y. (2025). AI integration in EFL teacher development: A mixed-methods evaluation of digital competency, professional trajectories, and pedagogical innovation within adaptive learning ecosystems. *Interactive Learning Environments*, 1–17. <https://doi.org/10.1080/10494820.2025.2591251>
- Becerra, A., Cobos, R., & Lang, C. (2025). Enhancing online learning by integrating biosensors and multimodal learning analytics for detecting and predicting student behaviour: A review. *Behaviour & Information Technology*, 45(9), 1885–1910. <https://doi.org/10.1080/0144929X.2025.2562322>
- Chen, X., Cheng, G., Zou, D., Zhong, B., & Xie, H. (2023). Artificial intelligent robots for precision education: A topic modeling-based bibliometric analysis. *Educational Technology & Society*, 26(1), 171–186. [https://doi.org/10.30191/ETS.202301_26\(1\).0013](https://doi.org/10.30191/ETS.202301_26(1).0013)
- Chen, X., Li, X., Zou, D., Xie, H., & Wang, F. (2025). Metacognition research in education: Topic modeling and bibliometrics. *Educational Technology Research and Development*, 73, 1399–1427. <https://doi.org/10.1007/s11423-025-10451-8>
- Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 1, 100002. <https://doi.org/10.1016/j.caeai.2020.100002>
- Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. *Educational Technology & Society*, 25(1), 28–47. <https://www.jstor.org/stable/48647028>
- Chu, H. C., Hwang, G. H., Tu, Y. F., & Yang, K. H. (2022). Roles and research trends of artificial intelligence in higher education: A systematic review of the top 50 most-cited articles. *Australasian Journal of Educational Technology*, 38(3), 22–42. <https://doi.org/10.14742/ajet.7526>
- Dağhan, G., & Gündüz, A. Y. (2022). Research trends in educational technology journals between 2000 and 2018: A Web scraping study. *Education and Information Technologies*, 27(4), 5179–5214. <https://doi.org/10.1007/s10639-021-10762-2>
- Daniel, K., Msambwa, M. M., Antony, F., & Wan, X. (2024). Motivate students for better academic achievement: A systematic review of blended innovative teaching and its impact on learning. *Computer Applications in Engineering Education*, 32(4), e22733. <https://doi.org/10.1002/cae.22733>

- Dogan, M. E., Goru Dogan, T., & Bozkurt, A. (2023). The use of artificial intelligence (AI) in online learning and distance education processes: A systematic review of empirical studies. *Applied Sciences*, 13(5), 3056. <https://doi.org/10.3390/app13053056>
- Gardner, J., O'Leary, M., & Yuan, L. (2021). Artificial intelligence in educational assessment: 'Breakthrough? Or buncombe and ballyhoo?' *Journal of Computer Assisted Learning*, 37(5), 1207–1216. <https://doi.org/10.1111/jcal.12577>
- Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794. <https://doi.org/10.48550/arXiv.2203.05794>
- Ilić, M., Mikić, V., Kopanja, L., & Vesin, B. (2023). Intelligent techniques in e-learning: A literature review. *Artificial Intelligence Review*, 56(12), 14907–14953. <https://doi.org/10.1007/s10462-023-10508-1>
- Jin, S. H., Im, K., Yoo, M., Roll, I., & Seo, K. (2023). Supporting students' self-regulated learning in online learning using artificial intelligence applications. *International Journal of Educational Technology in Higher Education*, 20(1), 37. <https://doi.org/10.1186/s41239-023-00406-5>
- Karakaya, K. (2021). Design considerations in emergency remote teaching during the COVID-19 pandemic: A human-centered approach. *Educational Technology Research and Development*, 69(1), 295–299. <https://doi.org/10.1007/s11423-020-09884-0>
- Kim, Y., & Kumi-Yeboah, A. (2025). Strategies to promote self-regulated learning for online diverse students. *Distance Education*, 46(2), 272–300. <https://doi.org/10.1080/01587919.2024.2358855>
- Lampropoulos, G. (2025). Augmented reality, virtual reality, and intelligent tutoring systems in education and training: A systematic literature review. *Applied Sciences*, 15(6), 3223. <https://doi.org/10.3390/app15063223>
- Liu, C., Zou, D., Chen, X., Xie, H., & Chan, W. H. (2021). A bibliometric review on latent topics and trends of the empirical MOOC literature (2008–2019). *Asia Pacific Education Review*, 22(3), 515–534. <https://doi.org/10.1007/s12564-021-09692-y>
- McInnes, L., Healy, J., Saul, N., & Großberger, L. (2018). UMAP: Uniform manifold approximation and projection. *Journal of Open Source Software*, 3(29), 861. <https://doi.org/10.21105/joss.00861>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, 8(5), 336–341. <https://doi.org/10.1016/j.ijsu.2010.02.007>
- Munir, H., Vogel, B., & Jacobsson, A. (2022). Artificial intelligence and machine learning approaches in digital education: A systematic revision. *Information*, 13(4), 203. <https://doi.org/10.3390/info13040203>

- Ouyang, F., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies*, 27(6), 7893–7925. <https://doi.org/10.1007/s10639-022-10925-9>
- Park, Y., & Doo, M. Y. (2024). Role of AI in blended learning: A systematic literature review. *International Review of Research in Open and Distributed Learning*, 25(1), 164–196. <https://doi.org/10.19173/irrodl.v25i1.7566>
- Qazi, S., Kadri, M. B., Naveed, M., Khawaja, B. A., Khan, S. Z., Alam, M. M., & Su'ud, M. M. (2024). AI-driven learning management systems: Modern developments, challenges and future trends during the age of ChatGPT. *Computers, Materials & Continua*, 80(2), 3289–3314. <https://doi.org/10.32604/cmc.2024.048893>
- Ruyang, L., Lu, T., Mengyao, X., & Hedi, Y. (2025). Contemporary trends of AI-powered foreign language teaching: A systematic literature review from connectivism perspective. *Innovation in Language Learning and Teaching*, 1–21. <https://doi.org/10.1080/17501229.2025.2586144>
- Sharma, P., Akgun, M., & Li, Q. (2024). Understanding student interaction and cognitive engagement in online discussions using social network and discourse analyses. *Educational Technology Research and Development*, 72(5), 2631–2654. <https://doi.org/10.1007/s11423-023-10261-w>
- Shi, S., & Zhang, H. (2025). EFL students' motivation predicted by their self-efficacy and resilience in artificial intelligence (AI)-based context: From a self-determination theory perspective. *Learning and Motivation*, 91, 102151. <https://doi.org/10.1016/j.lmot.2025.102151>
- Sjølie, E., Espenes, T. C., & Buø, R. (2022). Social interaction and agency in self-organizing student teams during their transition from face-to-face to online learning. *Computers & Education*, 189, 104580. <https://doi.org/10.1016/j.compedu.2022.104580>
- Torres-Vergara, C. J., Alfaro-García, V. G., Merigó, J. M., Atif, A., & McGreal, R. (2025). Twenty-five years of the International Review of Research in Open and Distributed Learning: A bibliometric analysis. *International Review of Research in Open and Distributed Learning*, 26(2), 286–306. <https://doi.org/10.19173/irrodl.v26i2.8662>
- Tram, N. H. M., Nguyen, T. T., & Tran, C. D. (2024). ChatGPT as a tool for self-learning English among EFL learners: a multi-methods study. *System*, 127, 103528. <https://doi.org/10.1016/j.system.2024.103528>



August – 2026

Does Interaction Quality Enhance Motivation Among Online Postsecondary Students? Exploring the Mediating Role of Anxiety

Moatasim A. Barri

Taibah University, Medina, Saudi Arabia

Abstract

The literature has highlighted the importance of engaging students in online learning through multiple forms of interaction: student-content, student-instructor, and student-student interactions. While the quality of interaction plays a significant role in either enhancing or undermining students' motivation to learn, anxiety has remained a persistent challenge. Limited research has examined the indirect relationships between each of the three forms of interaction and motivation through the lens of anxiety. This study aimed to investigate the mediating role of anxiety in the relationships between the forms of interaction quality and student motivation in university-level online courses. Data were collected from 140 students via an online survey, and multiple procedures ensured instrument validity. Structural equation modeling was used to evaluate research hypotheses. The research model revealed a good fit with the collected data based on the recommended goodness-of-fit indices. Findings revealed that anxiety fully mediated the relationships between student-content and student-instructor interactions and motivation, and partially mediated the relationship between student-student interaction and motivation. These results suggested that instructors should carefully consider the quality of all three types of interaction when designing and delivering online courses, with the goal of reducing students' anxiety and supporting their motivation to learn.

Keywords: online learning, anxiety, motivation, interaction quality, postsecondary education

Introduction

The demand for online postsecondary education has grown rapidly (Xia et al., 2022) due to increasing student numbers and constraints on space, time, and scheduling (Lu et al., 2024). In response, many institutions have shifted from face-to-face to online delivery, especially for theoretical courses (Xia et al., 2022). Online learning may be synchronous (Banna et al., 2015; Majewska & Zvobgo, 2023), asynchronous (Lu et al., 2024), or blended (Yamagata-Lynch, 2014), offering students flexibility to learn from anywhere (Martin & Bolliger, 2022). However, online learning places greater demands on instructors (Conklin & Dikkers, 2021), who are expected to implement strategies tailored to the online environment across multiple areas, including teaching (Mahmood, 2021), assessment and evaluation (Martin et al., 2019), facilitation (Martin, Polly, & Ritzhaupt, 2020; Conklin & Dikkers, 2021), and interaction (Abrami et al., 2011; Banna et al., 2015; Martin & Bolliger, 2018). A comprehensive review by Baldwin et al. (2018) indicated that interaction was one of the key criteria in many rubrics used to design, deliver, and evaluate high-quality online courses.

Interaction has been defined as a reciprocal event involving at least two objects and actions that mutually influence each other (Wagner, 1994) and it has long played a critical role in shaping the quality of the online learning experience. Bernard et al. (2009) identified three key forms of interaction: student-content (SC), student-instructor (SI), and student-student (SS), as originally conceptualized by Moore (1989). The quality of these interactions predicts student satisfaction (Wang et al., 2024), engagement (Limbu & McKinley, 2025), and learning outcomes (Quadir et al., 2019). To optimize SC, SI, and SS interactions, online courses should include both synchronous and asynchronous components (Banna et al., 2015; Martin, Wang, & Sadaf, 2020; Yamagata-Lynch, 2014).

Motivation has been defined as an internal state that drives students to act, pursue goals, and stay engaged (Ormrod, 2008). Research has shown that postsecondary online learners often report low motivation due to limited or poor SI and SS interactions (Esra & Sevilen, 2021; Lin et al., 2017; Xia et al., 2022). Low-quality SC interaction, resulting from (a) unclear instructions; (b) ineffective feedback; (c) a mismatch between students' expectations and course content; (d) poor course organization (Esra & Sevilen, 2021); and (e) dissatisfaction with the format and quality of course content and delivery (Xia et al., 2022) has also contributed to reduced motivation. Overall, studies indicated that student interaction positively predicted motivation (Çebi, 2023). Peer interaction further enhanced motivation directly and indirectly by reducing technological and communication barriers (Kulal et al., 2025). Additionally, intrinsic motivation has been positively related to SC, SI, and SS interactions (Jung et al., 2025).

Anxiety has been defined as a feeling of concern, uneasiness, and apprehension about uncertain situations (Ormrod, 2008) and has been the most commonly reported negative emotion among online postsecondary students (Xu & Wang, 2023). These students experienced higher anxiety than those in lower educational levels (Xu & Wang, 2023) or traditional classrooms (Li et al., 2024). Anxiety has been positively linked to low motivation in online learning (Peng et al., 2024). Research has also shown that low-quality interaction contributed to elevated anxiety: discomfort with instructor capacity and communication increased general and technology anxieties (Ritzhaupt et al., 2022), and separation from instructors and peers, along with technological challenges, positively predicted online learning anxiety (Peng et al., 2024).

Recent studies provided important insights into the relationships between interaction and different forms of anxiety. Lu et al. (2024) found that SI and SS interactions indirectly reduced test anxiety through emotional engagement, while SI interaction directly reduced test anxiety. Wut and Wong (2024) showed that interaction quality directly reduced performance and technology anxieties and indirectly alleviated psychological distress by reducing these anxieties in online learning environments. Additionally, an interactive online educational environment that enhanced student interaction with instructors, peers, and content had a significant impact on reducing general anxiety among postsecondary students (Wang, 2024).

Existing research has primarily examined interaction, motivation, and anxiety separately or has focused only on the direct relationships among them. However, very limited work has explored whether general anxiety serves as a connecting pathway between the quality of SC, SI, and SS interactions and student motivation in online postsecondary education. To address this gap, the present study examined the indirect relationships between interaction quality and motivation through general anxiety, offering a more comprehensive understanding of how different forms of interaction may support motivational outcomes in online learning environments.

Theoretical Framework and Hypotheses Development

The Quality of Interaction Through Self-Determination Theory

Interaction is a central element of online course design (Baldwin et al., 2018; Martin et al., 2021) and includes engagement with content, communication with instructors, and collaboration with peers (Bernard et al., 2009; Majewska & Zvobgo, 2023; Moore, 1989). High-quality interaction is typically supported through a combination of asynchronous and synchronous components (Yamagata-Lynch, 2014; Martin, Wang, & Sadaf, 2020; Martin & Bolliger, 2022). Asynchronous elements provide students with flexibility and autonomy in managing their learning (Jacobi, 2018), whereas synchronous elements promote real-time interaction and relatedness (Banna et al., 2015; Conklin & Dikkers, 2021; Jacobi, 2018; Majewska & Zvobgo, 2023). Together, these elements enhance students' engagement and a sense of competence (Martin, Wang, & Sadaf, 2020).

The current study used self-determination theory (SDT), developed by Deci and Ryan (1985), as a framework to evaluate the quality of interaction in terms of three psychological needs—autonomy, competence, and relatedness. SDT emphasizes that motivation strengthens when these needs are fulfilled (Ryan & Deci, 2017). SDT was selected as the primary framework because it provides an explicit mechanism linking interaction quality to students' basic psychological needs, which are closely associated with anxiety and motivation. Online course designs that support choice, mastery, and social connection are associated with engagement, intrinsic motivation, retention, and learning outcomes (Martin et al., 2018; Hsu et al., 2019; Shank et al., 2025).

The first psychological need in SDT is autonomy, which refers to experiencing control, independence, and ownership over one's actions (Ryan & Deci, 2017). Autonomy is fostered when courses provide flexibility through (a) asynchronous learning materials that allow self-paced engagement, and (b) asynchronous communication channels that enable students to connect with instructors and peers at their convenience (Jacobi, 2018; Martin et al., 2018). The second basic psychological need in SDT is competence, which refers

to feeling effective and capable of achieving desired outcomes. It is fostered when courses include (a) clear instructions and expectations for performance and communication, (b) timely and clear instructor feedback, and (c) collaborative activities that provide an optimal level of challenge (Jacobi, 2018; Martin et al., 2018; Shank et al., 2025). The third basic psychological need in SDT is relatedness, which refers to feeling connected to others and experiencing a sense of belonging. Relatedness is fostered when courses include synchronous learning opportunities and communication channels that enable meaningful interaction with instructors or peers (Jacobi, 2018; Martin et al., 2018; Shank et al., 2025).

When these psychological needs were supported, students were more likely to engage cognitively, emotionally, and behaviorally (He et al., 2025; Huang & Wang, 2023), interact with instructors, peers, content, and the platform (He et al., 2025), feel competent and connected (Hsu et al., 2019), and perform optimally (Huang & Wang, 2023). In contrast, in courses lacking autonomy-, competence-, and relatedness-supportive features, students reported lower motivation (Esra & Sevilen, 2021; Hartnett, 2015; Lin et al., 2017; Xia et al., 2022) and higher anxiety (Ritzhaupt et al., 2022; Peng et al., 2024). SDT therefore provided a coherent framework for examining how interaction quality relates to anxiety and motivation in online learning.

The Quality of Student-Content Interaction

Student-content interaction involves engagement with course materials, both asynchronously (e.g., readings, recorded lectures, assignments, discussion forums, projects, assessments) and synchronously (e.g., real-time presentations and demonstrations; Moore, 1989; Bernard et al., 2009). SC interaction is associated with active engagement and improved learning outcomes (Zimmerman, 2012; Nieuwoudt, 2018; Quadir et al., 2019).

In this study, SC interaction quality was conceptualized through SDT needs: (a) autonomy, reflected in asynchronous materials that enable self-paced engagement (Jacobi, 2018); (b) competence, reflected in clear instructions and expectations (Martin & Bolliger, 2018; Martin et al., 2018); and (c) relatedness, reflected in synchronous opportunities for interaction (Jacobi, 2018).

The Quality of Student-Instructor Interaction

Student-instructor interaction involves communication and feedback that can occur synchronously (e.g., phone calls, live chats, virtual sessions) or asynchronously (e.g., e-mail, forums; Abrami et al., 2011; Banna et al., 2015). Students often prefer synchronous channels for direct communication (Conklin & Dikkers, 2021), while e-mail and forums have been valued for personal contact, rapport building, discussion, updates, and feedback delivered in multiple formats (Conklin & Dikkers, 2021; Glazier, 2016; Martin & Bolliger, 2018). SI interaction has been identified as the most important interaction type (Martin & Bolliger, 2018), positively associated with SC interaction quality (Majewska & Zvobgo, 2023), and linked to engagement, satisfaction, perceived learning, retention, and grades (Alqurashi, 2018; Glazier, 2016; Martin & Bolliger, 2018). Optimal SI interaction includes multiple communication channels and timely feedback to maintain connection and reduce isolation (Conklin & Dikkers, 2021; Martin & Bolliger, 2018).

In this study, SI interaction quality was conceptualized through SDT needs: (a) autonomy, supported by flexible asynchronous communication channels (Glazier, 2016; Conklin & Dikkers, 2021); (b) competence,

supported by timely and effective feedback (Jacobi, 2018; Martin et al., 2018); and (c) relatedness, supported by synchronous tools enabling real-time connection (Jacobi, 2018; Conklin & Dikkers, 2021; Majewska & Zvobgo, 2023).

The Quality of Student-Student Interaction

Student-student interaction refers to communication between peers in pairs or small groups, with or without instructor presence (Moore, 1989), and may occur synchronously (e.g., live sessions, chats) or asynchronously (e.g., forums, e-mail; Abrami et al., 2011). Effective online courses integrate both formats (Yamagata-Lynch, 2014; Martin & Bolliger, 2018), as students value synchronous peer interaction (Banna et al., 2015). SS interaction is associated with active engagement (Martin & Bolliger, 2018), improved learning outcomes (Quadir et al., 2019), and higher achievement, as supported by meta-analytic evidence (Bernard et al., 2009).

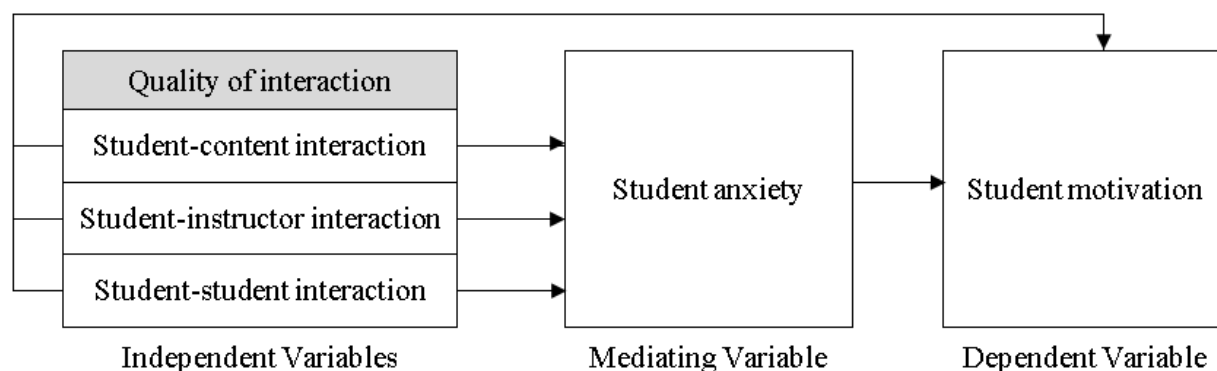
In this study, SS interaction quality was conceptualized through SDT needs: (a) autonomy, supported by flexible asynchronous peer communication channels (Jacobi, 2018; Martin et al., 2018); (b) competence, supported by optimally challenging collaborative activities (Martin et al., 2018; Shank et al., 2025); and (c) relatedness, supported by synchronous tools that foster real-time peer connection (Jacobi, 2018; Shank et al., 2025).

Research Model and Hypotheses Development

Figure 1 shows the research model, which included student anxiety as the mediator, interaction quality (SC, SI, SS) as the independent variables, and student motivation as the dependent variable.

Figure 1

Research Model



The Direct Relationship Between Anxiety and Motivation

In this study, student anxiety referred to the extent to which students experienced concern while engaging in online learning environments due to a lack of autonomy, competence, and relatedness. Student motivation referred to the extent to which students experienced interest and willingness to engage in online learning based on perceived autonomy, competence, and relatedness. The literature has shown that online

learning can enhance students' motivation when the quality of interaction is high (Çebi, 2023; Jung et al., 2025; Kulal et al., 2025), but it can also heighten anxiety due to limited social interaction, technical challenges, isolation, weak two-way communication, or discomfort with instructor support (Peng et al., 2024; Ritzhaupt et al., 2022). Research has shown that online postsecondary students report higher anxiety (Li et al., 2024) and lower intrinsic and extrinsic motivation (Stark, 2019) than students in face-to-face settings. Anxiety among online learners has also been positively associated with lower motivation (Peng et al., 2024), suggesting that elevated anxiety can reduce students' willingness to engage. Given these patterns, this study examined the direct relationship between anxiety and motivation as part of testing anxiety's mediating role between the quality of each of SC, SI, and SS interactions and motivation. Accordingly, the following hypothesis was proposed:

H1: Student anxiety is negatively associated with their motivation toward online learning.

The Direct Relationships Between Interaction Quality and Anxiety

Literature on online postsecondary education has shown that interaction quality plays a key role in shaping student anxiety. Poor instructor communication or competence has been linked to higher anxiety (Ritzhaupt et al., 2022), with increased anxiety when students felt separated from instructors and peers or when they faced technological challenges (Peng et al., 2024). In contrast, high-quality SI interaction has been negatively associated with test anxiety (Lu et al., 2024), and strong interaction can reduce both performance- and technology-related anxieties (Wut & Wong, 2024). Interactive online environments that promote engagement with instructors, peers, and content has also helped alleviate anxiety (Wang, 2024). Accordingly, the current study examined the direct relationships between the quality of SC, SI, and SS interactions and student anxiety as a foundation for testing anxiety's mediating role. The following hypotheses were proposed:

H2: The quality of SC interaction is negatively associated with student anxiety toward online learning.

H4: The quality of SI interaction is negatively associated with student anxiety toward online learning.

H6: The quality of SS interaction is negatively associated with student anxiety toward online learning.

The Direct Relationships Between Interaction Quality and Motivation

Literature on online postsecondary education has shown that interaction quality is significantly associated with student motivation. Low levels of motivation have been linked to limited real-time SI and SS interaction, insufficient feedback and instructions, poor course organization, and mismatches between student expectations and content (Esra & Sevilen, 2021; Lin et al., 2017; Xia et al., 2022). Conversely, high-quality interaction with instructors, peers, or content has been shown to enhance motivation (Çebi, 2023; Jung et al., 2025; Kulal et al., 2025). The current study examined the direct relationships between SC, SI, and SS interaction quality and motivation as part of testing anxiety's mediating role. Accordingly, the following hypotheses were proposed:

H3: The quality of SC interaction is positively associated with student motivation toward online learning.

H5: The quality of SI interaction is positively associated with student motivation toward online learning.

H7: The quality of SS interaction is positively associated with student motivation toward online learning.

The Indirect Relationships Between Interaction Quality and Motivation Through Anxiety

In online postsecondary education, interaction quality has been shown to indirectly affect psychological outcomes through anxiety. For example, Wut and Wong (2024) found that interaction quality reduced psychological distress via performance and technology anxiety. Lu et al. (2024) showed that high-quality SI and SS interactions indirectly alleviated test anxiety through emotional engagement. Kulal et al. (2025) found that peer interaction indirectly enhanced motivation by reducing technology and communication barriers. However, there has been limited research on the indirect effects of SC, SI, and SS interaction quality on motivation through anxiety. Accordingly, the following hypotheses were proposed:

H8: Students' anxiety mediates the relationship between the quality of SC interaction and their motivation.

H9: Students' anxiety mediates the relationship between the quality of SI interaction and their motivation.

H10: Students' anxiety mediates the relationship between the quality of SS interaction and their motivation.

Method

Population, Sampling, and Setting

The study population comprised undergraduate students in online courses at public universities in Saudi Arabia. Due to time and access constraints, a convenience sample was drawn from a single university in Medina.

Sample and Participants

A total of 140 undergraduate students completed the online survey, with no missing data or outliers. Most participants were male (84%), and nearly half were enrolled in community college (47.1%). Other disciplines included humanities (12.1%), business (9.3%), computer science (9.3%), medical science (7.9%), engineering (5%), science (4.3%), education (2.9%), and law (2.1%). Regarding technology use, 52.1% identified as intermediate users, 27.9% as experts, and 20% as novices. Most accessed the platform via smartphones (53.6%) or computers (40.7%), while a few used tablets (5.7%). Device availability and quality varied: 43.6% had no private access, 28.6% had outdated devices, 26.5% had unworkable devices, and 18.6% had out-of-order devices. Internet access was poor for 59.3%, expensive for 37.1%, and unavailable for 10.7%. Most participants reported sufficient knowledge and skills to use their devices (70%) and the learning management system (LMS; 73.6%). Multiple responses were allowed for device status and Internet access items; therefore, percentages may exceed 100%.

Procedure and Data Collection

Ethical approval was obtained from the institutional review board of the participating public university. Undergraduate students received an email invitation containing the survey link, study purpose, and confidentiality assurances. Participation was voluntary, and participants provided electronic informed consent prior to completing the survey. Responses were collected anonymously.

The Research Instrument

The research instrument was an online survey with 23 items divided into two sections. The first section included eight demographic questions on gender, academic discipline, technology use, device type and status, Internet access, and knowledge/skills for using the device and LMS. The second section had 15 Likert-scale items, developed by the author from the literature on student anxiety, motivation, and interaction in online postsecondary education. These items were grouped into five dimensions—SC interaction, SI interaction, SS interaction, student anxiety (SA), and student motivation (SM)—each with three items (see Table 1). Items were randomly ordered, and participants responded on a five-point scale from 1 (*strongly disagree*) to 5 (*strongly agree*).

Table 1

Dimensions and Likert-Scale Items of the Research Instrument

SDT need	Item	Variable
Student-content interaction		
		Online courses provide . . .
Autonomy	SC1	asynchronous learning materials that I can engage with at my own pace (e.g., readings, recorded lectures, presentation slides, Web-based resources)
Competence	SC2	clear instructions and expectations for my performance, participation, and communication
Relatedness	SC3	synchronous learning materials that foster interaction with my instructor and classmates
Student-instructor interaction		
		Online courses provide . . .
Autonomy	SI1	asynchronous communication channels for connecting with the instructor (e.g., e-mail, instructor forum)
Competence	SI2	clear instructor feedback on my assignments and inquiries
Relatedness	SI3	synchronous communication channels for connecting with the instructor (e.g., virtual classrooms, chats, phone calls)
Student-student interaction		
		Online courses provide . . .
Autonomy	SS1	asynchronous communication channels for connecting with peers (e.g., student e-mail lists, discussion forums)
Competence	SS2	collaborative learning activities that offer an optimal level of challenge
Relatedness	SS3	synchronous communication channels for connecting with peers (e.g., virtual student sessions, group chats)
Student motivation		
		Online learning . . .
Autonomy	SM1	allows me to study at my own pace
Competence	SM2	provides support I need to succeed
Relatedness	SM3	helps me feel connected to others

SDT need	Item	Variable
		Student anxiety
		Online learning makes me feel anxious due to a lack of . . .
Autonomy	SA1	control and independence in my learning
Competence	SA2	confidence in my ability to succeed
Relatedness	SA3	opportunities to interact with peers or instructors

Validating the Research Instrument

Content Validity of the Research Instrument. Content validation was conducted in three stages. First, the 23-item English instrument (15 Likert-scale, 8 demographic) was developed based on literature on student anxiety, motivation, and interaction in online postsecondary education. Second, three experts reviewed the instrument, leading to modifications for clarity, reduced ambiguity, and better alignment with constructs. Third, the instrument was translated into Arabic and back-translated into English by a bilingual translator, confirming the accuracy and validity of the Arabic version.

Construct Validity of the Research Instrument. Exploratory factor analysis was used to identify dimensions of the research instrument. A sample of 140 students was suitable for conducting reliable factor analysis (Sapnas & Zeller, 2002), and data suitability was confirmed with KMO = .774 ($p < .001$) and Bartlett's $\chi^2 = 622.42$ ($df = 105, p < .001$). Principal components extraction with varimax rotation revealed five dimensions as shown in Table 2, with all 15 Likert-scale items loading above .50. The final instrument included 23 items: 15 Likert-scale and 8 demographics.

Table 2

Factor Analysis Results of the Research Instrument

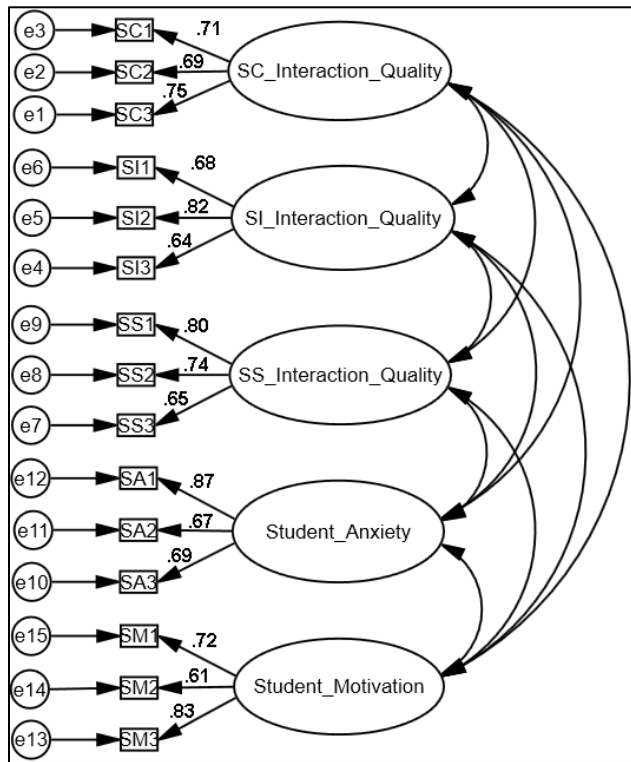
Dimension and item	Factor loading					Eigenvalue	Variance(%)
	1	2	3	4	5		
Quality of SC interaction						4.51	30.06
SC1	0.723						
SC2	0.655						
SC3	0.778						
Quality of SI interaction						1.89	12.60
SI1		0.731					
SI2		0.835					
SI3		0.686					
Student motivation						1.27	8.47
SM1			0.680				
SM2			0.672				
SM3			0.857				

Dimension and item	Factor loading					Eigenvalue	Variance(%)
	1	2	3	4	5		
Quality of SS interaction						1.12	7.44
SS1				0.688			
SS2				0.581			
SS3				0.840			
Student anxiety						1.06	7.04
SA1					0.722		
SA2					0.825		
SA3					0.662		
Total variance (%) =							65.61

A confirmatory factor analysis supported the five-dimension instrument identified in the exploratory factor analysis and showed good fit (CFI = .96, TLI = .95, RMSEA = .04, SRMR = .06). All standardized loadings (see Figure 2) exceeded .60, supporting construct validity.

Figure 2

Confirmatory Factor Analysis Model with Standardized Loadings



Convergent Validity of the Research Instrument. Convergent validity was assessed following Fornell and Larcker (1981) and Hair et al. (2010), with average variance extracted (AVE) $\geq .50$, composite reliability (CR) $> .70$, and factor loadings $> .60$. All dimensions met these criteria (see Figure 2 and Table 3), indicating adequate convergent validity.

Discriminant Validity of the Research Instrument. Discriminant validity was assessed following Fornell and Larcker (1981), with the square root of the average variance extracted ($\sqrt{\text{AVE}}$) for each dimension exceeding its inter-dimension correlations (see Table 3), indicating adequate discriminant validity.

Reliability of the Research Instrument. Cronbach's alpha (α) and composite reliability (CR) were used to assess internal consistency. Following Hair et al. (2017), CR values $\geq .60$ are acceptable; in this study, CR ranged from .76 to .79, indicating satisfactory reliability. Cronbach's alpha ranged from .66 to .73. Although two constructs fell slightly below the .70 threshold (Nunnally, 1978), lower values are common in short scales with few items.

Table 3

Results of Convergent and Discriminant Validity and Composite Reliability of the Research Instrument

Dimension	α	CR	AVE	$\sqrt{\text{AVE}}$	Inter-dimension correlation				
					SC	SI	SS	SA	SM
SC	0.67	0.76	0.51	0.71	1	.39	.44	-.44	.33
SI	0.70	0.76	0.52	0.72	.39	1	.27	-.40	.13
SS	0.72	0.78	0.54	0.73	.44	.27	1	-.39	.46
SA	0.73	0.79	0.56	0.75	-.44	-.40	-.39	1	-.35
SM	0.66	0.77	0.53	0.73	.33	.13	.46	-.35	1

Common Method Bias. Harman's single-factor test indicated that no single factor accounted for the majority of the variance, with the single factor explaining approximately 26.1% of the total variance. In addition, a common latent factor analysis revealed minimal changes ($< .20$) in standardized loadings, suggesting that common method bias was unlikely to be a serious concern.

Goodness-of-Fit, Assumption Testing, and Statistical Power

Prior to analysis, goodness-of-fit measures—including chi-square/degree of freedom (χ^2/df), goodness-of-fit index (GFI), adjusted goodness-of-fit index (AGFI), normalized fit index (NFI), incremental fit index (IFI), comparative fit index (CFI), and standardized root mean square residual (SRMR)—were used to assess model fit. All values met recommended criteria (see Table 4), indicating a reasonable fit.

Table 4

Goodness-of-Fit Measures and Values for the Research Model

Goodness-of-fit measure	Recommended criterion	Research model value
χ^2 / df	Low as 2 or high as 5 (Marsh & Hocevar, 1985)	2.83
GFI	> 0.95 (Byrne, 1994)	0.96
AGFI	> 0.90 (Byrne, 1994)	0.91
NFI	> 0.90 (Byrne, 1994)	0.92
IFI	> 0.90 (Bollen, 1989)	0.91
CFI	> 0.90 (Bentler, 1990)	0.93
SRMR	< 0.08 (Hu & Bentler, 1999)	0.05

Structural equation modeling (SEM) assumptions were examined prior to analysis. Tolerance and variance inflation factor (VIF) values indicated no multicollinearity, as tolerances did not approach zero (Tabachnick & Fidell, 2007) and VIFs did not exceed 5 or 10 (James et al., 2013) (see Tables 5 and 6). Histograms and normal Q-Q plots supported the assumption of normality. Scatterplots of standardized predicted values against standardized residuals for the anxiety and motivation models (Figures 3 and 4) showed random dispersion around zero with no systematic curvature and approximately constant variance, supporting the assumptions of linearity and homoscedasticity.

The research hypotheses were tested using structural equation modeling (SEM) in AMOS. Direct and indirect relationships were tested using 2,000 bootstrap samples with bias-corrected 95% confidence intervals. Using Cohen's (1988) multiple regression framework, the statistical power exceeded .99 for both student motivation ($R^2 = .26$, four predictors, $N = 140$) and student anxiety ($R^2 = .29$, three predictors) at $\alpha = .05$.

Table 5

*Multicollinearity Diagnostics for the Model
Predicting Student Anxiety*

Independent variable	Tolerance	VIF
SC	0.73	1.38
SI	0.83	1.20
SS	0.80	1.26

Table 6

*Multicollinearity Diagnostics for the Model
Predicting Student Motivation*

Independent variable	Tolerance	VIF
SC	0.68	1.47
SI	0.78	1.29
SS	0.76	1.32
SA	0.71	1.41

Figure 3

Scatterplot of Standardized Predicted Values and Standardized Residuals for the Model Predicting Student Anxiety

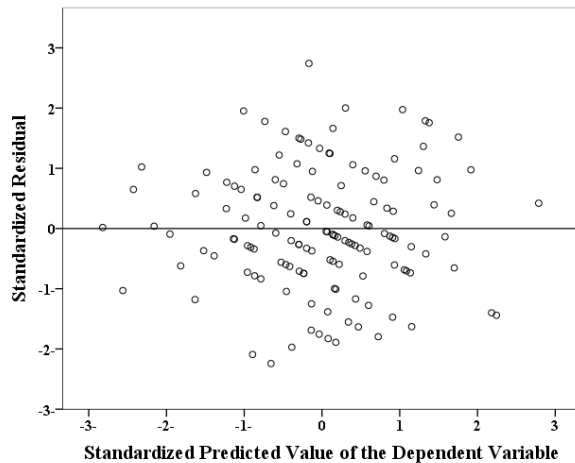
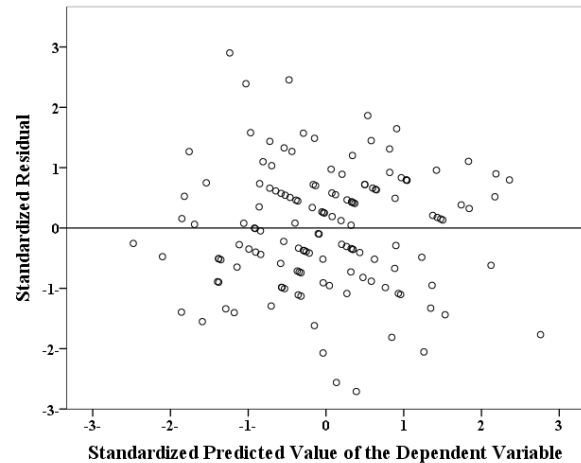


Figure 4

Scatterplot of Standardized Predicted Values and Standardized Residuals for the Model Predicting Student Motivation



Results

Table 7 presents the results of the mediation analysis that examined students' perceived anxiety in the relationships between the quality of SC, SI, and SS interactions, as well as students' perceived motivation in online learning. SC and SI interaction qualities were significantly associated with anxiety (SC → anxiety: $\beta = -0.25$, $p = .002$; SI → anxiety: $\beta = -0.24$, $p = .002$), and anxiety was significantly associated with motivation ($\beta = -0.20$, $p = .022$). However, the direct relationships between SC and SI interaction qualities and motivation were not significant (SC → motivation: $\beta = 0.13$, $p = .153$; SI → motivation: $\beta = 0.09$, $p = .261$). Their indirect relationships with motivation through anxiety were significant (SC → anxiety → motivation: $\beta = 0.05$, $p = .021$, 95% CI [.01, .16]; SI → anxiety → motivation: $\beta = 0.05$, $p = .015$, 95% CI [.01, .13]), indicating full mediation. SS interaction quality was significantly associated with anxiety (SS → anxiety: $\beta = -0.21$, $p = .009$), which in turn was significantly associated with motivation. The direct association between SS interaction quality and motivation remained significant (SS → motivation: $\beta = 0.35$, $p = .008$). The indirect association via anxiety was also significant (SS → anxiety → motivation: $\beta = 0.04$, $p = .037$, 95% CI [.01, .13]), indicating partial mediation. The model explained 26% of the variance in motivation and 29% in anxiety, indicating moderate explanatory power.

Table 7

Results of the Structural Path Analysis and Mediation Testing of the Research Model

Hypothesis	Path type	β	p	95% CI		Hypothesis supported?	Mediation type
				LL	UL		
H1: anxiety→ motivation	Direct	-.20	.022	-.38	-.02	Yes	-
H2: SC→ anxiety	Direct	-.25	.002	-.45	-.05	Yes	-
H3: SC→ motivation	Direct	.13	.153	-.04	.30	No	-
H8: SC→ anxiety→ motivation	Indirect	.05	.021	.01	.16	Yes	Full
H4: SI→ anxiety	Direct	-.24	.002	-.40	-.10	Yes	-
H5: SI→ motivation	Direct	.09	.261	-.26	.08	No	-
H9: SI→ anxiety→ motivation	Indirect	.05	.015	.01	.13	Yes	Full
H6: SS→ anxiety	Direct	-.21	.009	-.40	-.02	Yes	-
H7: SS→ motivation	Direct	.35	.008	.15	.53	Yes	-
H10: SS→ anxiety→ motivation	Indirect	.04	.037	.01	.13	Yes	Partial

Note. R^2 for student motivation = .26; R^2 for student anxiety = .29; LL = lower limit; UL = upper limit.

Discussion

This study examined how interaction quality—conceptualized through SDT principles of autonomy, competence, and relatedness—was associated with student motivation in online learning through anxiety. It investigated whether anxiety statistically mediated the relationships between each of SC, SI, and SS interaction quality and student motivation.

The findings indicated that anxiety fully mediated the relationship between SC interaction and motivation. The significant negative association between SC interaction quality and anxiety aligned with prior research in which well-designed, interactive learning environments were associated with lower levels of reported anxiety (Wang, 2024). Within SDT, clear expectations and structured content may be associated with greater perceived competence (Martin et al., 2018), while asynchronous materials may be positively associated with autonomy through flexible engagement (Jacobi, 2018). Incorporating appropriate synchronous materials may correspond with stronger perceptions of relatedness (Jacobi, 2018). The absence of a significant direct path from SC interaction quality to motivation suggested that the association between SC interaction quality and motivation may have operated primarily through students' emotional experiences rather than through a direct motivational pathway.

A similar pattern emerged for SI interaction quality, where anxiety fully mediated its association with motivation. Supportive and responsive instructor communication may be associated with lower uncertainty and isolation, which are commonly linked to anxiety in online learning contexts. Timely and constructive feedback has been positively associated with perceived competence (Jacobi, 2018; Martin et al., 2018; Shank et al., 2025). Synchronous communication channels—such as virtual classrooms, chats, or phone calls—have been linked to perceived relatedness through real-time interaction (Jacobi, 2018; Martin et al.,

2018; Conklin & Dikkers, 2021; Majewska & Zvobgo, 2023), while asynchronous communication channels such as e-mail or forums have been positively associated with perceived autonomy, as they allowed students to engage on their own terms (Conklin & Dikkers, 2021; Glazier, 2016). These elements have also been linked to lower anxiety and a more supportive learning environment in which SI interaction quality was positively associated with student motivation. The non-significant direct path indicated that the association between SI interaction quality and motivation was largely explained by anxiety. This pattern suggested that instructor interaction may be associated with motivation primarily through students' emotional regulation rather than through a direct motivational pathway.

The findings showed that anxiety partially mediated the relationship between SS interaction and student motivation, indicating that SS interaction quality was associated with motivation both directly and indirectly through anxiety. From an SDT perspective, asynchronous communication tools, such as discussion forums and peer e-mail lists, have been positively associated with perceived autonomy by allowing students to engage at their own pace (Jacobi, 2018; Martin et al., 2018), while synchronous communication tools—such as virtual student sessions and group chats—have been linked to perceived relatedness through real-time interaction (Jacobi, 2018; Shank et al., 2025). Collaborative problem-solving activities are associated with perceived competence (Martin et al., 2018; Shank et al., 2025). Although higher-quality SS interaction was associated with lower anxiety, the significant direct path suggested that peer interaction may relate to motivation through additional social mechanisms—such as belonging or social presence—that extend beyond anxiety. SS interaction quality may be particularly connected to relatedness, which may be associated with motivation independently of anxiety.

Taken together, the findings suggested that different forms of interaction may operate through distinct psychological pathways within the SDT framework. SC and SI interactions appear to be closely linked to emotional regulation processes, whereas SS interaction may be more strongly connected to relatedness-based social mechanisms that extend beyond anxiety. This distinction advances our theoretical understanding by demonstrating that interaction types are not functionally equivalent in their associations with motivation.

Conclusion

The findings showed that anxiety differentially mediated the relationships between types of interaction quality and motivation. Anxiety fully mediated the relationships between SC and SI interactions and motivation, indicating these interactions related to motivation primarily through anxiety. For SS interaction, anxiety partially mediated the relationship, suggesting that peer interaction related to motivation through additional emotional or social factors beyond anxiety.

This study advanced online learning research by demonstrating that SC, SI, and SS interaction quality were not functionally equivalent in their associations with anxiety and motivation. While prior SDT-informed research has often treated interaction quality as a unified construct, the present findings distinguished full mediation for SC and SI interaction quality from partial mediation for SS interaction quality. This distinction suggested that SC and SI interaction quality may be linked to motivation primarily through emotional regulation processes, whereas SS interaction quality may operate through additional social mechanisms, such as belonging or social presence. By differentiating these pathways, the study refined our

theoretical understanding of interaction quality within SDT and contributed to a more nuanced model of motivational processes in online learning contexts.

The findings of this study offer practical implications for instructional designers and instructors. For SC interaction quality, the full mediation pattern suggests that online courses may include asynchronous learning materials that support autonomy, explicit guidance on performance, participation, and communication to support competence, as well as synchronous learning activities that support relatedness. For SI interaction quality, the full mediation pattern suggests that instructors may maintain asynchronous communication channels (e.g., e-mail, forums) to support autonomy, synchronous communication channels (e.g., virtual classrooms, phone calls) to support relatedness, and timely, constructive feedback on assignments and inquiries to support competence. The partial mediation observed for SS interaction suggested that peer activities and communication channels should be designed to support not only anxiety-related processes but also broader social dimensions such as belonging and relatedness.

This study had several methodological limitations. The use of a convenience sample from a single public university in Saudi Arabia limited the generalizability of the findings to other cultural or educational settings. The sample was predominantly male students enrolled in virtual courses, which may have limited the generalizability of the results. The study data were self-reported, which may have introduced response biases. The researcher did not control instructors' course design, which may have contributed to variability in students' perceptions of interaction quality, anxiety, and motivation. This study was limited by the exclusive use of an SDT-informed framework to measure the three forms of interaction quality. Each construct in the instrument was measured using only three items, which may not have fully captured its breadth. This study employed a cross-sectional design, which precluded causal inference. Finally, focusing solely on anxiety as a mediator limited consideration of other relevant mediators or covariates.

Future research could expand on the findings of this study in several ways. First, replicating the findings across institutions, digital platforms, academic levels, cultural contexts, and more gender-balanced samples would help assess the generalizability of the observed mediation patterns. Second, incorporating qualitative methods could complement self-reported data and help reduce potential response bias. Third, future studies could control course design features to improve internal validity when examining associations among interaction quality, student anxiety, and motivation. Fourth, this study adopted an SDT-informed framework; future research could extend this work by comparing alternative theoretical models, such as social-cognitive or control-value theories. Fifth, future research could use longer and more comprehensive measurement scales to capture the constructs more fully and strengthen measurement validity. Sixth, experimental or quasi-experimental designs could be employed to examine the effect of SDT-informed interaction quality on student anxiety and motivation. Finally, future research could examine additional emotional, social, cognitive, demographic, or contextual mediators or covariates to further clarify the mechanisms linking interaction quality and student motivation.

Declaration Regarding Availability of Data and Material

The data collected and analyzed during the current study are available from the author upon reasonable request.

References

- Abrami, P. C., Bernard, R. M., Bures, E. M., Borokhovski, E., & Tamim, R. M. (2011). Interaction in distance education and online learning: Using evidence and theory to improve practice. *Journal of Computing in Higher Education*, 23, 82–103. <https://doi.org/10.1007/s12528-011-9043-x>
- Alqurashi, E. (2018). Predicting student satisfaction and perceived learning within online learning environments. *Distance Education*, 40(1), 133–148. <https://doi.org/10.1080/01587919.2018.1553562>
- Baldwin, S., Ching, Y.-H., & Hsu, Y.-C. (2018). Online course design in higher education: A review of national and statewide evaluation instruments. *TechTrends*, 62(1), 46–57. <https://doi.org/10.1007/s11528-017-0215-z>
- Banna, J., Lin, M.-F. G., Stewart, M., & Fialkowski, M. K. (2015). Interaction matters: Strategies to promote engaged learning in an online introductory nutrition course. *Journal of Online Learning and Teaching*, 11(2), 249–261. https://jolt.merlot.org/Vol11no2/Banna_0615.pdf
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>
- Bernard, R. M., Abrami, P. C., Borokhovski, E., Wade, C. A., Tamim, R. M., Surkes, M. A., & Bethel, E. C. (2009). A meta-analysis of three types of interaction treatments in distance education. *Review of Educational Research*, 79(3), 1243–1289. <https://doi.org/10.3102/0034654309333844>
- Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research*, 17(3), 303–316. <https://doi.org/10.1177/0049124189017003004>
- Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows*. Sage Publications.
- Çebi, A. (2023). How e-learning readiness and motivation affect student interactions in distance learning? *Education and Information Technologies*, 28, 2941–2960. <https://doi.org/10.1007/s10639-022-11312-0>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203771587>
- Conklin, S., & Dikkers, A. G. (2021). Instructor social presence and connectedness in a quick shift from face-to-face to online instruction. *Online Learning*, 25(1), 135–150. <https://doi.org/10.24059/olj.v25i1.2482>
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Springer. <https://doi.org/10.1007/978-1-4899-2271-7>

- Esra, M., & Sevilen, Ç. (2021). Factors influencing EFL students' motivation in online learning: A qualitative case study. *Journal of Educational Technology and Online Learning*, 4(1), 11–22. <https://dergipark.org.tr/en/pub/jetol/issue/60134/817680>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.2307/3151312>
- Glazier, R. A. (2016). Building rapport to improve retention and success in online classes. *Journal of Political Science Education*, 12(4), 437–456. <https://doi.org/10.1080/15512169.2016.1155994>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: A global perspective*. Pearson.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage Publications.
- Hartnett, M. K. (2015). Influences that undermine learners' perceptions of autonomy, competence and relatedness in an online context. *Australasian Journal of Educational Technology*, 31(1). <https://doi.org/10.14742/ajet.1526>
- He, J., Wang, Q., & Lee, H. (2025). Enhancing online learning engagement: Teacher support, psychological needs satisfaction, and interaction. *BMC Psychology*, 13, 696. <https://doi.org/10.1186/s40359-025-03016-0>
- Hsu, H.-C. K., Wang, C. V., & Levesque-Bristol, C. (2019). Reexamining the impact of self-determination theory on learning outcomes in the online learning environment. *Education and Information Technology*, 24, 2159–2174. <https://doi.org/10.1007/s10639-019-09863-w>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Huang, Y., & Wang, S. (2023). How to motivate student engagement in emergency online learning? Evidence from the COVID-19 situation. *Higher Education*, 85, 1101–1123. <https://doi.org/10.1007/s10734-022-00880-2>
- Jacobi, L. (2018). What motivates students in the online communication classroom? An exploration of self-determination theory. *Journal of Educators Online*, 15(2). <https://doi.org/10.9743/jeo.2018.15.2.1>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning with applications in R*. Springer. <https://doi.org/10.1007/978-1-4614-7138-7>

- Jung, E., Samuel, A., & Ma, T. (2025). Beyond motivation: The critical role of learner interaction in MOOCs. *American Journal of Distance Education*, 0(0) 1–22.
<https://doi.org/10.1080/08923647.2025.2474295>
- Kulal, A., N, A., & Dinesh, S. (2025). Online peer interactions and student motivation: A perception study. *Behaviour & Information Technology*, 44(18), 1–21.
<https://doi.org/10.1080/0144929X.2025.2477756>
- Li, Z., Guan, P., Li, J., & Wang, J. (2024). Comparing online and offline Chinese EFL learners' anxiety and emotional engagement. *Acta Psychologica*, 242, Article e104114,
<https://doi.org/10.1016/j.actpsy.2023.104114>
- Limbu, Y. B., & McKinley, C. (2025). Factors associated with student engagement in online learning during COVID-19 pandemic: A systematic review. *Online Learning*, 29(1), 293–325.
<https://doi.org/10.24059/olj.v29i1.4221>
- Lin, C.-H., Zhang, Y., & Zheng, B. (2017). The roles of learning strategies and motivation in online language learning: A structural equation modeling analysis. *Computers & Education*, 113, 75–85.
<https://doi.org/10.1016/j.compedu.2017.05.014>
- Lu, K., Zhu, J., Pang, F., & Liu, Z. (2024). Understanding college students' test anxiety in asynchronous online courses: the mediating role of emotional engagement. *International Journal of Educational Technology in Higher Education*, 21, Article 50. <https://doi.org/10.1186/s41239-024-00482-1>
- Mahmood, S. (2021). Instructional strategies for online teaching in COVID-19 pandemic. *Human Behavior and Emerging Technologies*, 3(1), 199–203. <https://doi.org/10.1002/hbe2.218>
- Majewska, I., & Zvobgo, V. (2023). Students' satisfaction with the quality of synchronous online learning under the COVID-19 pandemic: Perceptions from liberal arts and science undergraduates. *Online Learning*, 27(1), 313–335. <https://doi.org/10.24059/olj.v27i1.3201>
- Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin*, 97(3), 562–582. <https://doi.org/10.1037/0033-2909.97.3.562>
- Martin, F., & Bolliger, D. U. (2018). Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment. *Online Learning*, 22(1), 205–222.
<https://doi.org/10.24059/olj.v22i1.1092>
- Martin, F., & Bolliger, D. U. (2022). Designing online learning in higher education. In O. Zawacki-Richter & I. Jung (Eds.), *Handbook of open, distance and digital education* (pp. 1–22). Springer.
https://doi.org/10.1007/978-981-19-0351-9_72-1

- Martin, F., Bolliger, D. U., & Flowers, C. (2021). Design matters: Development and validation of the online course design elements (OCDE) instrument. *The International Review of Research in Open and Distributed Learning*, 22(2), 46–71. <https://doi.org/10.19173/irrodl.v22i2.5187>
- Martin, F., Polly, D., & Ritzhaupt, A. (2020, September 8). Bichronous online learning: Blending synchronous and synchronous online learning. *EDUCAUSE Review*. <https://er.educause.edu/articles/2020/9/bichronous-online-learning-blending-asynchronous-and-synchronous-online-learning>
- Martin, F., Ritzhaupt, A., Kumar, S., & Budhrani, K. (2019). Award-winning faculty online teaching practices: Course design, assessment and evaluation, and facilitation. *The Internet and Higher Education*, 42, 34–43. <https://doi.org/10.1016/j.iheduc.2019.04.001>
- Martin, F., Wang, C., & Sadaf, A. (2020). Facilitation matters: Instructor perception of helpfulness of facilitation strategies in online courses. *Online Learning*, 24(1), 28–49. <https://doi.org/10.24059/olj.v24i1.1980>
- Martin, N., Kelly, N., & Terry, P. (2018). A framework for self-determination in massive open online courses: Design for autonomy, competence, and relatedness. *Australasian Journal of Educational Technology*, 34(2). <https://doi.org/10.14742/ajet.3722>
- Moore, M. G. (1989). Editorial: Three types of interaction. *American Journal of Distance Education*, 3(2), 1–7. <https://doi.org/10.1080/08923648909526659>
- Nieuwoudt, J. (2018). Exploring online interaction and online learner participation in an online science subject through the lens of the interaction equivalence theorem. *Student Success*, 9(4), 53–62. <https://doi.org/10.5204/ssj.v9i4.520>
- Nunnally, J. C. (1978). *Psychometric theory*. McGraw-Hill.
- Ormrod, J. E. (2008). *Human learning* (5th ed.). Pearson Prentice Hall.
- Peng, R., Wang, S., & Liu, N. (2024). Factors influencing Chinese EFL students' online learning anxiety in the post-COVID-19 era. *Heliyon*, 10(4), Article e26112. <https://doi.org/10.1016/j.heliyon.2024.e26112>
- Quadir, B., Yang, J. C., & Chen, N. S. (2019). The effects of interaction types on learning outcomes in a blog-based interactive learning environment. *Interactive Learning Environments*, 30(2), 293–306. <https://doi.org/10.1080/10494820.2019.1652835>
- Ritzhaupt, A. D., Rehman, M., Wilson, M. L., & Ruggles, K. (2022). Exploring the factors associated with undergraduate students' online learning anxiety: Development of the online learner anxiety scale. *Online Learning*, 26(3), 111–131. <https://doi.org/10.24059/olj.v26i3.2810>

- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. The Guilford Press. <https://doi.org/10.1521/978.14625/28806>
- Sapnas, K. G. & Zeller, R. A. (2002). Minimizing sample size when using exploratory factor analysis for measurement. *Journal of Nursing Measurement*, 10(2), 135–153. <https://doi.org/10.1891/jnum.10.2.135.52552>
- Shank, E., Tang, H., & Morris, W. (2025). Motivation in online course design using self-determination theory: An action research study in a secondary mathematics course. *Educational Technology Research and Development*, 73, 415–441. <https://doi.org/10.1007/s11423-024-10410-9>
- Stark, E. (2019). Examining the role of motivation and learning strategies in student success in online versus face-to-face courses. *Online Learning*, 23(3), 234–251. <https://doi.org/10.24059/olj.v23i3.1556>
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Pearson Education.
- Wagner, E. D. (1994). In support of a functional definition of interaction. *American Journal of Distance Education*, 8(2), 6–29. <https://doi.org/10.1080/08923649409526852>
- Wang, X., Hassan, A. B., Shwu Pyng, H., Ye, H., & Aminuddin, A. A. (2024). The impact of perceived quality of online interaction on satisfaction of international student interactions in higher education. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2023.2293454>
- Wang, Y. (2024). An interactive online educational environment to reduce anxiety, improve emotional well-being, and critical thinking for college students. *Acta Psychologica*, 248, Article e104347. <https://doi.org/10.1016/j.actpsy.2024.104347>
- Wut, T. M., & Wong, S M. H. (2024). Does online class student interactions mitigate distress and anxiety problems during the post-pandemic period? Perspective from the conservation of resources theory. *Learning Environment Research*, 27, 995–1011. <https://doi.org/10.1007/s10984-024-09516-x>
- Xia, Y., Hu, Y., Wu, C., Yang, L., & Lei, M. (2022). Challenges of online learning amid the COVID-19: College students' perspective. *Frontiers in Psychology*, 13, Article e1037311. <https://doi.org/10.3389/fpsyg.2022.1037311>
- Xu, T., & Wang, H. (2023). High prevalence of anxiety, depression, and stress among remote learning students during the COVID-19 pandemic: Evidence from a meta-analysis. *Frontiers in Psychology*, 13, Article e1103925. <https://doi.org/10.3389/fpsyg.2022.1103925>
- Yamagata-Lynch, L. C. (2014). Blending online asynchronous and synchronous learning. *The International Review of Research in Open and Distributed Learning*, 15(2). <https://doi.org/10.19173/irrodl.v15i2.1778>

Zimmerman, T. D. (2012). Exploring learner to content interaction as a success factor in online courses.
The International Review of Research in Open and Distributed Learning, 13(4), 152–165.
<https://doi.org/10.19173/irrodl.v13i4.1302>



August – 2026

Building a Framework to Guide the Design and Development of Open Educational Practices for the Digital Training of University Professors

Francesc M. Esteve-Mon¹, Anna Sánchez-Caballé¹, María Ángeles Llopis Nebot¹, Virginia Viñoles Cosentino¹, and Jose Cela-Ranilla²

¹Universitat Jaume I, Spain; ²Universidad de Valladolid, Spain

Abstract

In the context of widespread digitalization in higher education, enhancing teacher digital competence (TDC) has become a critical priority. However, institutional efforts to advance TDC, including the development of open educational resources (OER), are often hindered by the prevalence of training models that lack pedagogical depth and rely on transmissive methods. Current literature on open educational practices (OEP) has highlighted the transformative potential of openness in teaching, yet it also suggests a disconnect between the availability of resources and their actual adoption. Building upon this gap, this article describes the creation of a theoretical framework intended to guide the design of OER-based training that effectively integrates OEP for faculty development. The method employed was a qualitative, exploratory approach, constituting the first phase of a design-based research (DBR) project. The framework was constructed in two stages: first, defining its conceptual foundations through a rigorous literature review, and second, validating and refining these elements through input from experts in educational technology. The primary result is a validated theoretical framework for OER-based digital training. The conclusions emphasize that this framework is distinct from others because it serves as a meta-reflective element, focusing directly on the teacher's own professional development and practice, rather than only on student-facing teaching processes. While the framework can guide institutional digital transformation, its initial design was developed within a specific university context, following the DBR approach. Future research will involve implementing, revising, and validating the framework in other contexts.

Keywords: teacher digital competence, open educational resources, OER, theoretical framework, open educational practices, OEP

Introduction

Integrating technologies into higher education is a process that has been going on for several decades but has accelerated significantly in recent years. This critical turning point is a response to various institutional digital transformation initiatives that seek to take advantage of the opportunities and challenges posed by emerging technologies and to redefine educational paradigms (Esteve-Mon et al., 2023).

However, the COVID-19 health crisis critically highlighted a dual need. On the one hand, it is necessary to have a robust technological infrastructure, and on the other, more pressingly, there is the need to develop adequate digital training for the entire university community, especially faculty (Crawford et al., 2020). In this context, teacher digital competence (TDC) has been consolidated as a key aspect, as highlighted by various experts (Falloon, 2020). While frameworks such as DigCompEdu (Redecker & Punie, 2017) have been fundamental in defining and even assessing TDC, they can be excessively fragmented and instrumental, offering little guidance on the pedagogical processes for TDC development (Hidalgo, 2024).

At the European level, and in response to the COVID-19 pandemic, various administrations have promoted policies to strengthen the digitalization of higher education, with the support of NextGenerationEU funds. Specifically, in Spain, the context of this research, thanks to the Recovery, Transformation, and Resilience Plan, various projects have emerged to improve TDC through the production of open educational resources (OER; Sánchez-Caballé et al., 2025).

However, these initiatives often remain limited to creating materials for enhancing digital competence, maintaining overly transmissive training models (Viñoles-Cosentino et al., 2022). Such models generally focus on the creation, exchange, and reuse of OER, failing to leverage the full potential of open pedagogies to propose open educational practices (OEP) with more constructivist, collaborative, situated, and open approaches (Cronin, 2017; Koseoglu et al., 2020).

This paper describes the process of creating a theoretical framework that serves as a reference for the design of OER-based training actions, enabling the full potential of OEP to be harnessed for the digital training of university professors. In other words, this framework offers the elements to shift from a vision excessively focused on the linear transmission of knowledge and object repositories toward embedding these into teacher training processes and professional development through OEP.

Theoretical Framework

Digital Training of University Professors

The professional development of university teaching staff has become a central concern in higher education in recent decades, reflecting the growing need to prepare faculty for increasingly complex and changing technological and pedagogical environments (Buils et al., 2024). Traditionally, the training of university teaching staff has focused primarily on research, giving it greater weight and neglecting pedagogical training or relegating it to informal or mandatory spaces. However, in recent decades, awareness has grown regarding the importance of supporting and training faculty not only in their disciplinary knowledge but also in their pedagogical and teaching practices (Buils et al., 2024; Más-Torelló, 2011; Rué Domingo, 2015).

This shift towards greater recognition of teaching as a complex and situated practice has led to the development of various initiatives and institutional frameworks aimed at improving teaching and learning processes in the university setting (Sánchez-Tarazaga et al., 2022). Considering the impact that digitalization has had in recent decades, special emphasis has been placed on strategies that enable teachers to develop the skills necessary to teach in a world increasingly mediated by digital technology (Basilotta-Gómez-Pablos et al., 2022; Castañeda et al., 2024; Fernández-Batanero et al., 2022). The impact of the COVID-19 pandemic has further highlighted the need for the comprehensive development of TDC (Bartolic et al., 2022; Crawford et al., 2020; García-Peñalvo & Corell, 2020). The sudden shift from in-person to emergency remote learning highlighted gaps in pedagogical integration and emphasized the importance of preparing teachers not only to use digital tools but also to adapt teaching strategies to ever-evolving contexts (Castañeda et al., 2023; Crawford et al., 2020; García-Peñalvo & Corell, 2020).

In this regard, various initiatives have been implemented at the European level. The European Commission, for example, has emphasized the need to integrate pedagogical and technological innovation into higher education through policies such as the *Digital Education Action Plan 2021–2027*, which prioritizes the professional development of teachers at all educational levels and promotes digital transformation as an institutional objective (European Commission, 2020). Likewise, the European Commission's Joint Research Center has developed frameworks such as DigCompEdu to promote the development of teachers' digital skills (Redecker & Punie, 2017).

In the Spanish context, this growing interest in teaching quality has been reinforced by initiatives such as the NextGenerationEU Recovery Plan, which has prioritized improving the digital skills of teaching staff as part of broader modernization and digitalization efforts (Castañeda et al., 2023). These funds have supported national plans such as Unidigital, aimed at modernizing university campuses and fostering digital and pedagogical innovation. These initiatives reflect the recognition of the digital transformation of higher education not only as a technological challenge, but also as a pedagogical and cultural one, requiring coordinated efforts at multiple levels of the university system (Viñoles Cosentino et al., 2025).

However, despite these advances, several studies highlight persistent gaps and challenges in the professional development of university professors, particularly regarding the integration of digital technologies into pedagogical practices (Castañeda et al., 2018; García-Peñalvo & Corell, 2020; Mercader & Gairín, 2020). Although universities increasingly offer training in digital tools and platforms and aim to develop the digital competences of their faculty, these initiatives are still often fragmented or disconnected from broader institutional strategies, which limits their impact on sustainable change (Oliveira et al., 2021; Viñoles Cosentino, 2025). In this sense, it is necessary to go beyond purely technical training to address the pedagogical and critical dimensions of digital competence, which are essential for promoting transformative and student-centered learning environments. More holistic, collaborative, and contextualized approaches to teacher development are needed (Biesta et al., 2015; Castañeda et al., 2018).

Scientific literature highlights the need to align teacher professional development with broader institutional policies and strategies, fostering not only technical skills but also pedagogical agency and reflexivity (Castañeda et al., 2018; Esteve-Mon et al., 2023; Stenalt et al., 2023). In this sense, competency development should be understood as a shared responsibility within universities,

encompassing institutional frameworks, training programs, and the development of supportive environments that foster experimentation and critical engagement (Inamorato dos Santos et al., 2023; Viñoles-Cosentino et al., 2022).

Strategies such as the use of OER and OEP can offer scaffolding opportunities to support the development of teachers' digital competence within a broader framework that considers professional growth and institutional transformation.

From OER to OEP

OER are a key innovation in education, facilitating free access to teaching, learning, and research materials. The OER concept was introduced by UNESCO, referring to materials in the public domain or under open licenses that allow their use, adaptation, and redistribution with minimal restrictions (UNESCO, 2002, 2012). The distinctive feature of OER is their openness, supported by the five freedoms proposed by Wiley and Hilton (2018): retain, reuse, revise, remix, and redistribute. These freedoms foster participatory, collaborative, and personalized learning. UNESCO (2019) has actively promoted the incorporation of OER into educational systems in recent years, considering that their use can contribute to building more inclusive knowledge-based societies and meeting the Sustainable Development Goals (SDGs) and the 2030 Agenda. Furthermore, OER serve as “boundary objects” that enable faculty to collaborate, remix content, and, consequently, reflect on their own pedagogy (McBride & Abramovich, 2022; Wiley & Hilton, 2018).

Linked to OER are OEP, which include both the use and creation of open resources, as well as participatory teaching methodologies and approaches (Blomgren, 2018). Indeed, some authors contend that OER have focused excessively on content (Cronin, 2017), resulting in a limited impact on learning; consequently, they necessitate additional pedagogical practices (Tlili et al., 2023; Zou et al., 2025). Cronin (2017) defined OEP as a broad set of practices ranging from the use of OER to the adoption of open pedagogies. In higher education, UNESCO has highlighted the transformative role of OER and OEP, promoting teaching models focused on student autonomy, pedagogical innovation, and personalized learning (Inamorato dos Santos, 2019; Weller, 2013).

From a constructivist perspective, following Paskevicius (2017), OEP are articulated around four essential components: (a) designing of personalized learning; (b) creating resources, which enhances creativity and performance; (c) implementing of innovative teaching-learning activities; and (d) using collaborative assessment strategies that involve students in the generation of new resources. However, several authors highlight the need to carefully evaluate the pedagogical quality of these resources (Bournissen et al., 2019; Lemos de Carvalho et al., 2019). Deng et al. (2020) asserted that for meaningful learning to take place with OEP, it is necessary to develop motivating resources, adopt teaching methods that encourage students' intellectual commitment, and implement integrated technological and human strategies for evaluation and feedback.

Research on open education in the university context has focused primarily on its adoption at the institutional level, identifying the key factors that facilitate and hinder its practice. Following Stagg (2014, p. 4), faculty members do not appear to be familiar with the use of these types of resources due to various reasons such as a lack of investment in open education or a lack of time to explore OER. This is also due to other reasons related to the need to develop a culture of open creation, such as a lack of awareness of intellectual property issues, the level of TDC, and the need to adapt current teaching

practices to open practices (Koseoglu & Bozkurt, 2018).

Al Albri et al. (2022) proposed design principles that promote the integration of OER use and creation in higher education. Essentially, they recommended that open resources be integrated throughout the course using constructivist pedagogical models that place students at the center of the process and establish an intertwined connection between the knowledge and skills learned during the course.

Moving toward a more open model in higher education entails a profound transformation in the way we understand learning, integrating key elements such as technology, pedagogy, and the values of equity and collaboration, advancing toward a more inclusive, innovative, and sustainable university.

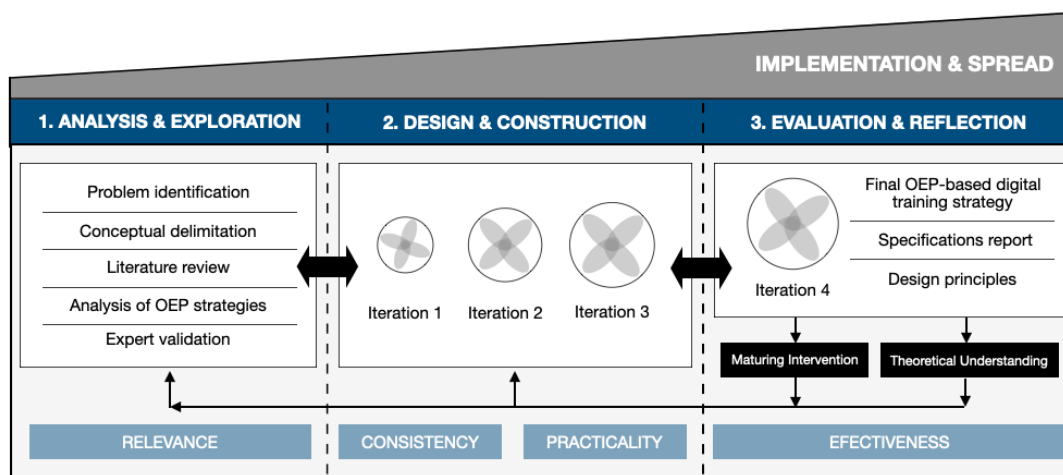
Context and Approach of the Research

This research is part of an R & D project aimed at promoting the development of OER to improve the digital competence of teaching staff at Jaume I University (Ref: CIGE/2023/18). This university, founded in the 1990s, offers various digitalization and teaching services to support the use of digital technologies in the educational context. Since 2020, it has launched a new digital transformation plan. As a result of this plan and the European NextGeneration grants for digitalization resulting from the COVID-19 pandemic, the university coordinated a regional initiative for the development of OER that fosters digital competence.

The methodological approach adopted follows design-based research (DBR), a systematic methodology characterized by iterative cycles of design, implementation, analysis, and redesign, with a dual focus on refining both theory and practice (Bakker, 2018; Hoadley & Campos, 2022; Wang & Hannafin, 2005). The objective of this approach is not only to provide solutions to a complex educational problem but also to generate a series of design principles that can be applied to similar contexts (Plomp & Nieveen, 2009; Van den Akker et al., 2006). Following the model of McKenney and Reeves (2013), the project is structured around three main phases: (a) analysis and exploration, (b) design and construction, and (c) evaluation and reflection. These phases are iterative, and each cycle generates insights that inform subsequent refinements (Figure 1).

Figure 1

Main Phases of DBR Process in the Research Project



Note. DBR = design-based research; OEP = open educational practices. See McKenny & Reeves (2013) for further discussion about DBR.

This article reports on the first phase—analysis and exploration—which focuses on understanding the problem, reviewing existing knowledge, and developing a preliminary conceptual framework that will underpin the subsequent design and implementation cycles.

Method

This paper uses a fundamentally qualitative approach with an exploratory focus that pursues a deep and exhaustive description of the phenomenon it analyzes—in this case, constructing a theoretical framework that serves as a reference for the design of OER-based training actions. This approach was selected due to its suitability for capturing the complexity, contextual nature, and multidimensionality of open pedagogical practices in digital learning environments.

This exploratory process was carried out in two phases: the first phase was to define the basic framework based on scientific literature, and the second phase was to contrast and validate it with input from experts. This two-phase design enabled an iterative refinement of the framework, combining theoretical grounding with expert-informed validation to enhance its coherence, relevance, and analytical robustness.

The initial framework was developed through a structured mapping and synthesis process. Three key studies (Al Abri et al., 2022; Paskevicius, 2017; Stagg, 2014) were selected for their complementary perspectives on open pedagogies. Key elements were extracted, compared in a matrix to identify overlaps and gaps, and iteratively synthesized by the research team into six general categories, with consensus ensured through researcher triangulation.

Participants and Instruments

In the first phase, which focused on the initial conceptualization of the reference model based on existing literature (Booth et al., 2022), a set of key dimensions of open pedagogy was used to structure the analysis: (a) pedagogical principles; (b) learning outcomes; (c) teaching and learning activities; (d) teaching and learning resources; (e) assessment systems; and (f) institutional support.

For the second phase of testing and validation, two focus groups were held with technical and

educational experts in teacher training, open pedagogies, and educational technology. A total of 13 experts (54% men and 46% women) from six different universities in Spain and Portugal participated. Participants were selected through purposive sampling, using specific criteria relevant to the research, such as the type of institution they attended, their roles, and the diversity of their professional profiles.

Procedure

Before the sessions, participants received an information sheet outlining the study objectives, framework, and discussion topics. The sessions were organized into three blocks, aligned with the components of the framework, and were moderated using open-ended questions to encourage critical evaluation and suggestions. With informed consent, the discussions were audio-recorded and transcribed verbatim for analysis.

The sessions were conducted via Google Meet and lasted between 90 and 120 minutes each. They were led by a moderator following a semi-structured approach, which fostered an environment of active participation.

Data Analysis

The sessions were recorded and transcribed, with prior informed consent from the participants and authorization from the university's ethics committee (Ref.: CEISH/91/2004). An inductive approach was adopted to analyze the data, using a system of categories designed from the research objectives and the conceptual framework generated in the preliminary phase, presented below. The analytical strategy used allowed the framework to be progressively refined through close engagement with the empirical material, ensuring alignment between the data, the analytical categories, and the study's objectives.

Qualitative software Atlas.ti (Version 25; <https://atlasti.com/>) was used for coding and analysis, and strategies such as triangulation of researchers and verification by various members of the research team were carried out to ensure the quality of the analysis. Following Cisterna (2005), triangulation in qualitative research was understood as a process of cross-checking, comparing, and dialectically interpreting information obtained from different sources, instruments, and analytical perspectives, with the aim of constructing a deeper, more coherent, and well-grounded understanding of the phenomenon under study. Specifically, credibility was strengthened through researcher triangulation and iterative discussions of coding decisions; dependability and confirmability were supported by maintaining a transparent audit trail of analytical procedures; and transferability was addressed by providing detailed descriptions of the analytical process and context.

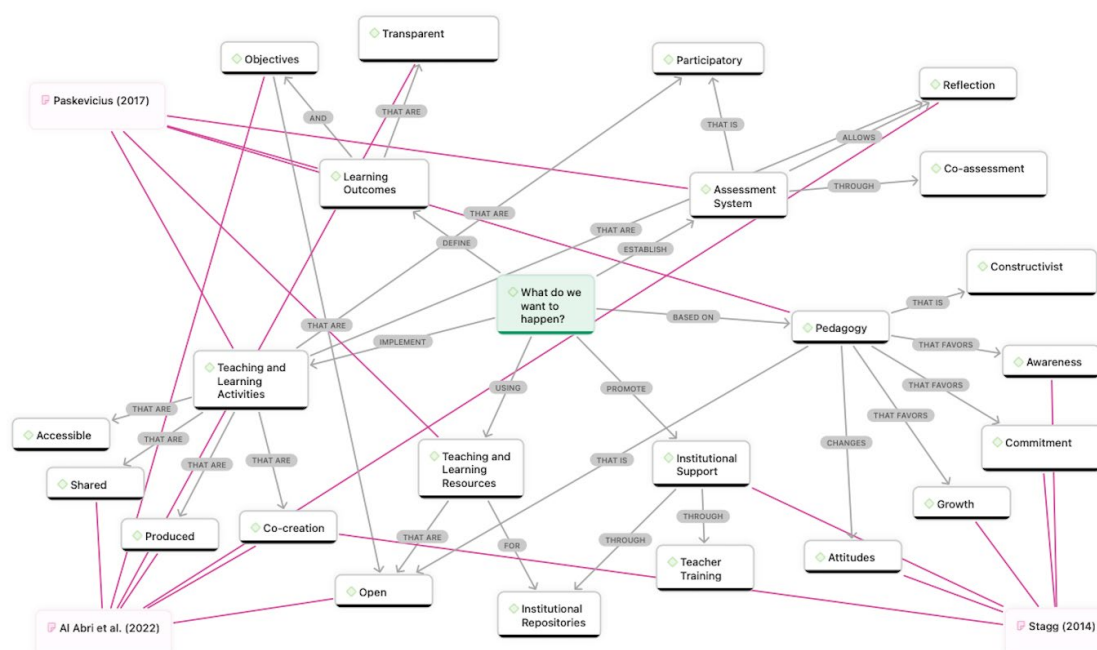
Results

The results are presented in two phases, as defined in the section "Participants and Instruments." The first phase identifies the conceptual foundations of the framework, and the second delves into the basic elements of these foundations.

Phase 1: Conceptual Foundation of the Framework

As illustrated in Figure 2, this framework, grounded in the literature (Al Abri et al., 2022; Paskevicius, 2017; Stagg, 2014), identifies the main elements and interrelationships that are expected when working with open pedagogies. The selection and organization of these elements respond to their recurrent presence in prior research and to their conceptual relevance for explaining how open pedagogical practices are designed, implemented, and sustained in digital learning contexts. These elements are: (a) pedagogical principles; (b) learning outcomes; (c) teaching and learning activities; (d) teaching and learning resources; (e) assessment systems; and (f) institutional support.

Figure 2



Note. The lines in pink represent a relationship that is grounded in the cited literature, while the lines in grey indicate the interrelationships between the main elements of the framework.

Based on these foundations and their interrelations, the first version of the framework was generated with a checklist-like organization, which took as its structural reference the vision of teachers, students (in this case, teachers who participate as students or trainees in the courses), and the educational institution itself.

Phase 2: Empirical Enrichment of the Framework

This phase describes the in-depth analysis conducted by the experts, in which they contrasted the theoretically defined foundations and validated their relevance in the debate generated in the focus groups. Through this process, the experts examined the extent to which the proposed components of the OEP framework were reflected in practice, discussed their coherence with prior research, and identified areas of convergence and refinement. The result of this expert dialogue is presented, organized around three elements: (a) pedagogical principles and learning outcomes; (b) teaching-learning and assessment activities; and (c) institutional resources and support.

Pedagogical Principles and Learning Outcomes

One element that emerged as central is the need for a pedagogical model focused on transferability and applicability, in which learning activities are meaningfully and relevantly connected to the real-life context of the preservice teachers' practice. Participants emphasized the importance of ensuring that content and activities are not perceived as mere formalities but as relevant proposals for their professional development.

Another key principle is the emphasis on collaborative knowledge building. The experts emphasized the

importance of fostering peer-to-peer learning dynamics, in which participants share practices, experiences, and reflections to build collective understanding, as seen in the following excerpt from our focus group discussions.

I think all of this is much more effective collaboratively, right? Especially when we're thinking about these open educational resources. Subjects aren't managed individually; rather, it's desirable for them to be managed by teams of university professors.

Likewise, critical reflection and the integration of OEP are seen as tools for more conscious and committed teaching, although the importance of addressing potential resistance, such as beliefs, was emphasized.

One of the things I see as important when promoting these open educational resources is that educators address the issue of beliefs We see that the main obstacles are often beliefs.

In this regard, a clear alignment between expected learning outcomes, activities, and assessment is proposed to ensure didactic coherence and transparency, as we will discuss in more detail in the following section.

Teaching-Learning and Assessment Activities

One of the most reiterated points was the importance of designing activities in training courses that are transferable to participants' different areas of knowledge, that are also editable and reusable in their own subjects, and that, in turn, are based on real-life cases, avoiding fictitious situations without direct application. In this regard, it is important not to use and share these resources only among teachers, but rather to use and co-create them with students.

Create open educational resources with students and have university faculty develop strategies so that students in our undergraduate and master's programs can co-create this content. [The challenge will be] how to involve students in these processes.

Emphasis was also placed on the design of collective activities, such as forums or guided debates, that allow for the exchange of perspectives and the building of knowledge among peers.

Regarding evaluation in these courses, the importance of formative assessment focused on developmental feedback, rather than immediate numerical grades, was emphasized. A model of initial, continuous, and final assessment was proposed to facilitate more precise monitoring of the learning process.

Likewise, peer assessment was recognized and highlighted as a relevant practice, both for learning and for creating communities of practice.

I believe that incorporating into one's own teaching practice a component of openness, transparency, sharing, co-creation, or that openness ... would be very useful. And so, when you design, you're thinking that you're making something you want to share, not something you want to lock away.

Resources and Institutional Support

Among the most frequently expressed ideas in relation to this third element was the necessary active commitment that educational institutions must demonstrate in the promotion and sustainability of OER. Along these lines, various initiatives were mentioned, such as calls for projects or workshops.

At the institutional level, one idea is to promote it by incentivizing requests for certain projects or initiatives, or through events to disseminate open educational resources.

It was also highlighted that these practices will only be sustainable through institutional support, and to achieve this, it is essential that they be recognized within academic career structures and adequately incentivized, enabling their gradual integration into faculty members' practice.

Ideally, it should end up being an intrinsic element, meaning I do it because I'm convinced I have to, but maybe even at the beginning, because that external reward mechanism is a bit necessary.

Furthermore, the role and involvement of university services such as libraries and resource centers, teaching innovation units, and educational technology were highlighted in supporting faculty in the use and creation of OER. Emphasis was also placed on the need for stable and secure infrastructures for resource curation, publication, and reuse, such as institutional repositories and interoperable metadata systems.

Ensuring the quality of shared resources was another reiterated aspect of this element. To achieve this, it was suggested to work through ex ante evaluation systems, such as rubrics or indicators, and other systems, such as peer review or public scoring systems, were proposed.

Finally, it was noted that artificial intelligence is an emerging tool for the co-creation and adaptation of materials, as well as for content curation, which must be taken into account now in an active, ethical, and critical manner.

Discussion and Conclusions

This article examines current issues in the digital training of university professors. Despite the institutional drive to improve TDC, transmissive training models persist, failing to fully harness the potential of open pedagogy. In response, this paper describes the process of developing a theoretical framework that underpins the design of training initiatives that overcome current limitations by effectively integrating the potential of OEP.

First, the conceptual foundations of the framework were identified through a literature review (Al Abri et al., 2022; Paskevicius, 2017; Stagg, 2014), and then, the core elements of these foundations were further explored through expert input. The elements that comprise the framework for OER-based digital training are detailed in Table 1, following the design and validation process.

Table 1

Elements of the Framework for OER-Based Digital Training

Core elements according to type of stakeholder
Teachers Transparency in objectives, learning outcomes, and assessment. Clearly communicate course objectives, expected learning outcomes, and assessment criteria. Open learning outcomes. Define learning outcomes that allow students to explore and construct knowledge in an open and collaborative manner. Co-creation of learning outcomes. Include students in the process of defining learning outcomes so they are committed to the course objectives. Pedagogical flexibility. Adapt teaching strategies to students' needs and interests, maintaining an open and flexible approach. Professional collaboration and peer assessment. Promote collaboration between teachers and other professionals to share materials and resources, as well as implement peer assessment processes. Participatory and reflective assessment. Design assessment mechanisms in which students actively participate and reflect on their own learning process. Focus on constructivist pedagogy. Focus on a constructivist pedagogy that encourages collaborative work, knowledge creation, and active student production of materials. Share open materials. Provide access to teaching materials under open licenses, encouraging students to use, adapt, and share them. Participation in communities of practice. Collaborate with other educators through communities of practice to share experiences and resources on OEP.
Students (or teachers participating as students or trainees) Co-creation of learning outcomes. Actively participate in defining learning outcomes to align them with one's interests and goals. Collaborative production. Working together with other students to produce knowledge, teaching materials, or projects that can be shared openly and freely. Share and peer-review open materials. Publish one's work and materials under open licenses and participate in peer-review processes with peers. Active and reflective participation in assessment. Contribute to the assessment of one's own learning process and that of peers through self-reflection and open dialogue with the teacher. Awareness of the use of open resources. Learn to locate, use, and create OER, taking into account open licenses and the sustainability of collaborative learning. Motivation. Taking the course should be motivating for the student in such a way that it fosters engagement in the learning process.
Institution Institutionalizing support for OEP. Develop policies that promote OEP, including the use of OER and co-creation practices. Institutional commitment to OEP. Promote a clear institutional vision of commitment to OEP, highlighting the importance of adopting open practices for educational innovation.

Promoting and recognizing open source work. Recognize and value the open source work of faculty and students, such as the production of open source materials and participation in communities of practice.

Provide tools and training in OER and TDC. Offer ongoing training and technological tools to facilitate the creation, use, and distribution of open materials, as well as for the development of TDC.

Sustainable perspective. Ensure that OEP-related initiatives are sustainable in the long term, both economically and technologically.

Institutional repository of open materials. Create and maintain repositories of open educational materials that can be used, modified, and shared by the entire educational community.

Promote a change in teacher attitudes. Promote changes in teachers' attitudes toward OEP by offering incentives and support for the transformation toward an open pedagogical model.

This framework conceives openness as a transversal attribute of professional practices rather than a mere characteristic of resources (Cronin, 2017). It shifts the focus from the availability of materials to the transformation of the dynamics of pedagogical design, collaboration, agency, and teaching reflection (Allman et al., 2024; Koseoglu & Bozkurt, 2018). Consistent with this framework, previous literature has also highlighted that implementing OEP in professional development programs encompasses not only the creation and use of OER, but also collaborative processes, participatory pedagogies, and teaching practices that transform the relationship between teaching and learning (Cronin & MacLaren, 2018).

Although there are different reference frameworks and indicators for promoting OER and OEP (Brandenburger, 2022; UNESCO, 2019), the present work presents a differential element: this framework directly informs teaching practice and digital training, rather than focusing exclusively on student-centered teaching and learning processes. By doing so, this study effectively bridges the previously identified gap between the mere technical availability of resources and the actual implementation of open pedagogy. This contribution advances current research by operationalizing OEP as a catalyst for professional identity and long-term pedagogical change.

Its meta-reflective approach is a significant contribution, as it positions teachers as learners, authors, and agents of pedagogical transformation (de Almeida & Amiel, 2023). Furthermore, it can serve as a reference framework for the institutional promotion of digital transformation processes (Esteve-Mon et al., 2023).

From a theoretical perspective, this work proposes a broader understanding of OEP by placing them in the context of teacher professional development, moving beyond their limited consideration as mere teaching strategy. In terms of training design, the framework suggests that digital training programs be structured around processes of co-creation, open production, peer review, and systematic reflection on practice, rather than focusing exclusively on proficiency in technological tools (Cronin & MacLaren, 2018).

While this framework is grounded in international literature and expert validation, its design was developed in a specific context and focused on a single university, which is essential in DBR studies. However, this contextualization implies that the framework reflects certain organizational cultures, levels of digital maturity, and openness policies specific to that environment. This contextualization, although necessary to ensure its initial relevance, is a limitation to its transferability. Consistent with

the DBR methodology outlined, future work will advance to the subsequent phases of design and evaluation. This focus on implementation and revision will not only allow for refining and consolidating the proposal through successive iterative cycles but will also strengthen its validation and adaptation to other university contexts. The potential for broader institutional change in open and digital learning ecosystems will be explored, contributing to a deeper understanding of how OEP can sustain long-term educational digital transformation processes.

Acknowledgments

This work was supported by the Department of Education, Culture and Universities of the Valencian Government (Spain) under Grant CIGE/2023/18. The authors used Gemini Pro and the services of a professional translator to improve the readability and language of the manuscript. The authors confirm that they have written and reviewed the manuscript and take full responsibility for its content.

References

- Al Albri, M. H., Bannan, B., & Dabbagh, N. (2022). The design and development of an open educational resources intervention in a college course that manifests in open educational practices: A design-based research study. *Journal of Computing in Higher Education*, 34, 154–188. <https://doi.org/10.1007/s12528-021-09285-z>
- Allman, B., Kimmons, R., Dickson-Deane, C., Bozkurt, A., Warr, M., Stefaniak, J., & Dash, M., & Bondah, F. E. (2024). EdTechnica: A vision of an educational publishing community of practice that is accessible, flexible, and just. *International Journal of Educational Technology in Higher Education*, 21, Article 37. <https://doi.org/10.1186/s41239-024-00466-1>
- Bakker, A. (2018). *Design research in education: A practical guide for early career researchers*. Taylor & Francis.
- Bartolic, S. K., Boud, D., Agapito, J., Verpoorten, D., Williams, S., Lutze-Mann, L., Matzat, U., Moreno, M. M., Polly, P., Tai, J., Marsh, H. L., Lin, L., Burgess, J. L., Habtu, S., Rodrigo, M. M., Roth, M., Heap, T., & Guppy, N. (2022). A multi-institutional assessment of changes in higher education teaching and learning in the face of COVID-19. *Educational Review*, 74(3), 517–533. <https://doi.org/10.1080/00131911.2021.1955830>
- Basilotta-Gómez-Pablos, V., Matarranz, M., Casado-Aranda, L.-A., & Otto, A. (2022). Teachers' digital competencies in higher education: A systematic literature review. *International Journal of Educational Technology in Higher Education*, 19(1), Article 8. <https://doi.org/10.1186/s41239-021-00312-8>
- Biesta, G., Priestley, M., & Robinson, S. (2015). The role of beliefs in teacher agency. *Teachers and Teaching*, 21(6), 624–640. <https://doi.org/10.1080/13540602.2015.1044325>
- Blomgren, C. (2018). OER awareness and use: The affinity between higher education and K–12. *The International Review of Research in Open and Distributed Learning*, 19(2). <https://doi.org/10.19173/irrodl.v19i2.3431>
- Booth, A., Sutton A., Clowes, M., & Martyn-St James, M. (2022). *Systematic approaches to a successful literature review*. SAGE.
- Bournissen, J., Tumino, M. C., & Carrión, F. (2019). MOOC: Evaluación y medición de la calidad percibida [MOOC: Evaluation and measurement of perceived quality]. *International Journal of Educational Research and Innovation (IJERI)*, 11, 18–32. <https://www.upo.es/revistas/index.php/IJERI/article/view/2899/2946>
- Brandenburger, B. (2022). A multidimensional and analytical perspective on Open Educational Practices in the 21st century. *Frontiers in Education*, 7, Article 990675. <https://doi.org/10.3389/feduc.2022.990675>
- Buils, S., Viñoles-Cosentino, V., Esteve-Mon, F. M., & Sánchez-Tarazaga, L. (2024). Digital training in university teaching induction programmes in Spain: A comparative analysis based on

- DigComp and DigCompEdu. *Educación XX1*, 27(2), 37-64.
<https://doi.org/10.5944/educxx1.38652>
- Castañeda, L., Esteve, F. & Adell, J. (2018). ¿Por qué es necesario repensar la competencia docente para el mundo digital? [Why rethinking teaching competence for the digital world?]. *RED. Revista de Educación a Distancia*, 56(6). <http://dx.doi.org/10.6018/red/56/6>
- Castañeda, L., Vanaclocha, N., Velasco Pérez, J. R., Ruiz Martínez, P. M., Hartillo Hermoso, M. I., Pereira González, E., & Ruiz Martínez, A. (2023). *Marco de competencia digital docente universitario. Creación y validación* [Framework of digital competence for university teachers. Creation and validation]. Proyecto UNIDIGITAL DigCompEdu-FYA.
<http://hdl.handle.net/10201/136836>
- Castañeda, L., Viñoles, V., Concannon, F., Pedersen, A., Al-Hmiedat, P., & Lobato, N. (2024). The CUTE CANVAS: developing a design tool for planning strategic actions for institutional of digital competencies. *Journal of Decision Systems*, 33(1), 30–52.
<https://doi.org/10.1080/12460125.2023.2167274>
- Cisterna, F. (2005). Categorización y triangulación como procesos de validación del conocimiento en investigación cualitativa [Categorization and triangulation as knowledge validation processes in qualitative research]. *Theoria*, 14(1), 61–71.
- Crawford, J., Butler-Henderson, K., Rudolph, J., Malkawi, B., Glowatz, M., Burton, R., Magni, P. A., & Lam, S. (2020). COVID-19: 20 countries' higher education intra-period digital pedagogy responses. *Journal of Applied Teaching and Learning (JALT)*, 3(1), 9–28.
<https://doi.org/10.37074/jalt.2020.3.1.7>
- Cronin, C. (2017). Openness and praxis: Exploring the use of open educational practices in higher education. *The International Review of Research in Open and Distributed Learning*, 18(5), 15–34. <https://doi.org/10.19173/irrodl.v18i5.3096>
- Cronin, C., & MacLaren, I. (2018). Conceptualising OEP: A review of theoretical and empirical literature in Open Educational Practices. *Open Praxis*, 10(2), p. 127–143.
<https://doi.org/10.5944/openpraxis.10.2.825>
- de Almeida Sousa, J., & Amiel, T. (2023). Prácticas Educativas Abiertas desde la perspectiva de educadores abiertos: Aportaciones a la formación del profesorado [Open educational practices from the perspective of open educators: Contributions to teacher professional development]. *EduTec, Revista Electrónica De Tecnología Educativa*, 85, 85–102.
<https://doi.org/10.21556/edutec.2023.85.2853>
- Deng, R., Benckendorff, P., & Gannaway, D. (2020). Linking learner factors, teaching context, and engagement patterns with MOOC learning outcomes. *Journal of Computer Assisted Learning*, 36(5), 688–708. <https://doi.org/10.1111/jcal.12437>

- Esteve-Mon, F. M., Postigo, A. Y., & Castañeda, L. (2023). A strategic approach of the crucial elements for the implementation of digital tools and processes in Higher Education. *Higher Education Quarterly*, 77(3), 558–573. <https://doi.org/10.1111/hequ.12411>
- European Commission. (2020). *Digital education action plan 2021–2027. Resetting education and training for the digital age*. https://ec.europa.eu/education/sites/education/files/document-library-docs/deap-communication-sept2020_en.pdf
- Falloon, G. (2020). From digital literacy to digital competence: The teacher digital competency (TDC) framework. *Educational Technology Research and Development*, 68, 2449–2472. <https://doi.org/10.1007/s11423-020-09767-4>
- Fernández-Batanero, J. M., Montenegro-Rueda, M., Fernández-Cerero, J., & García-Martínez, I. (2022). Digital competences for teacher professional development. Systematic review. *European Journal of Teacher Education*, 45(4), 513–531. <https://doi.org/10.1080/02619768.2020.1827389>
- García-Peñalvo, F. J., & Corell, A. (2020). La COVID-19: ¿enzima de la transformación digital de la docencia o reflejo de una crisis metodológica y competencial en la educación superior? [COVID-19: The enzyme of the digital transformation of teaching or the reflection of a methodological and competence crisis in higher education?]. *Campus Virtuales*, 9(2), 83–98. <http://uajournals.com/ojs/index.php/campusvirtuales/article/view/740>
- Hidalgo, M. (2024). Análisis del concepto de Competencia Digital Docente: una revisión sistemática de la literatura [Analysis of the concept of Digital Teaching Competence: A systematic literature review]. *RELATEC: Revista Latinoamericana de Tecnología Educativa*, 23(1), 25–41. <https://doi.org/10.17398/1695-288X.23.1.25>
- Hoadley, C., & Campos, F. C. (2022). Design-based research: What it is and why it matters to studying online learning. *Educational Psychologist*, 57(3), 207–220. <https://doi.org/10.1080/00461520.2022.2079128>
- Inamorato dos Santos, A. (2019). *Practical guidelines on open education for academics: Modernising higher education via open educational practices*. European Union. <https://doi.org/10.2760/55923>
- Inamorato dos Santos, A., Chinkes, E., Carvalho, M. A. G., Solórzano, C. M. V., & Marroni, L. S. (2023). The digital competence of academics in higher education: Is the glass half empty or half full? *International Journal of Educational Technology in Higher Education*, 20, Article 9. <https://doi.org/10.1186/s41239-022-00376-0>
- Koseoglu, S., & Bozkurt, A. (2018). An exploratory literature review on open educational practices. *Distance Education*, 39(4), 441–461. <https://doi.org/10.1080/01587919.2018.1520042>
- Koseoglu, S., Bozkurt, A., & Havemann, L. (2020). Critical questions for open educational practices. *Distance Education*, 41(2), 153–155. <http://dx.doi.org/10.1080/01587919.2020.1775341>

- Lemos de Carvalho, G., Jr., Cebrián Robles, D., Cebrián De La Serna, M., & Raposo Rivas, M. (2019). Comparative study SPOC vs. MOOC for socio-technical contents from usability and user satisfaction. *Turkish Online Journal of Distance Education*, 20(2) 4–20.
<https://doi.org/10.17718/tojde.557726>
- Mas Torelló, Ò. (2011). El profesor universitario: Sus competencias y formación [The university teacher: His competences and his training]. *Profesorado. Revista de Currículum y Formación de Profesorado*, 15(3), 195–211.
<http://www.redalyc.org/articulo.oa?id=56722230013>
- McBride, M. & Abramovich, S. (2022). Crossing the boundaries through OER adoption: Considering open educational resources (OER) as boundary objects in higher education. *Library & Information Science Research*, 44(2), Article 101154.
<https://doi.org/10.1016/j.lisr.2022.101154>
- McKenney, S., & Reeves, T. C. (2013). Systematic review of design-based research progress: Is a little knowledge a dangerous thing? *Educational Researcher*, 42(2), 97–100.
<https://doi.org/10.3102/0013189X12463781>
- Mercader, C., & Gairín, J. (2020). University teachers' perception of barriers to the use of digital technologies: The importance of academic discipline. *International Journal of Educational Technology in Higher Education*, 17, Article 4. <https://doi.org/10.1186/s41239-020-0182-x>
- Oliveira, G., Grenha Teixeira, J., Torres, A. & Morais, C. (2021). An exploratory study on the emergency remote education experience of higher education students and teachers during the COVID-19 pandemic. *British Journal of Educational Technology*, 52(4), 1357–1376.
<https://doi.org/10.1111/bjet.13112>
- Paskevicius, M. (2017). Conceptualizing open educational practices through the lens of constructive alignment. *Open Praxis*, 9(2), 125–140. <https://doi.org/10.5944/openpraxis.9.2.519>
- Plomp, T., & Nieveen, N. (Eds.) (2009). *An introduction to educational design research*. Netherlands Institute for Curriculum Development (SLO).
https://ris.utwente.nl/ws/portalfiles/portal/14472302/Introduction_20to_20education_20design_20research.pdf
- Redecker, C., & Punie Y. (2017). *European framework for the digital competence of educators: DigCompEdu*. European Union. <https://doi.org/10.2760/159770>
- Rué Domingo, J. (2015). El desarrollo profesional docente en Educación Superior: agenda, referentes y propuestas para su adopción [The professional development in teaching in higher education: Agenda, its frame and proposals for its adoption]. *REDU. Revista de Docencia Universitaria*, 13(3), 217–236. <https://doi.org/10.4995/redu.2015.5461>
- Sánchez-Caballé, A., Llopis-Nebot, M.A., & Esteve-Mon, F.M. (2025). Analysis of Digital Competence and Digital Teaching Competence Training Platform Designed by Five Valencian Universities. In E. Trepulè (Ed.), *EDEN Annual Conference Proceedings* (pp. 150-151).

- Sánchez-Tarazaga, L., Ruiz-Bernardo, P., Viñoles Cosentino, V., & Esteve-Mon, F. M. (2022). University teaching induction programmes: A systematic literature review. *Professional Development in Education*, 50(2), 279–295. <https://doi.org/10.1080/19415257.2022.2147577>
- Stagg, A. (2014). OER adoption: A continuum for practice. *RUSC. Universities and Knowledge Society Journal*, 11(3), 151–164. <http://dx.doi.org/10.7238/rusc.v11i3.2102>
- Stenalt, M. H., Johnson, M. W., & Aagaard, J. (2023). Reclaiming the teacher perspective in digital education – An analysis of university teachers’ agency. *Interactive Learning Environments*, 32(8), 4109–4121. <https://doi.org/10.1080/10494820.2023.2194930>.
- Tlili, A., Garzón, J., Salha, S., Huang, R., Xu, L., Burgos, D., Denden, M., Farrel, O., Farrow, R., Bozkurt, A., Amiel, T., McGreal, R., López-Serrano, A & Wiley, D. (2023). Are open educational resources (OER) and practices (OEP) effective in improving learning achievement? A meta-analysis and research synthesis. *International Journal of Educational Technology in Higher Education*, 20, Article 54. <https://doi.org/10.1186/s41239-023-00424-3>
- UNESCO. (2002). *Forum on the impact of open courseware for higher education in developing countries. Final report*. <https://r.issu.edu.do/n2>
- UNESCO. (2012). *Declaración de París de 2012 sobre los REA* [2012 Paris OER declaration]. Congreso Mundial sobre los Recursos Educativos Abiertos (REA). https://unesdoc.unesco.org/ark:/48223/pf0000246687_spa
- UNESCO. (2019). *Recommendation on open educational resources (OER)*. <https://r.issu.edu.do/mWf>
- Van den Akker, J. J. H., Gravemeijer, K., McKenney, S., & Nieveen, N. (2006). *Educational design research*. Routledge. <https://doi.org/10.4324/9780203088364>
- Viñoles Cosentino, V., Esteve-Mon, F. M., & Adell-Segura, J. (2025). Design and development of a digital transformation plan for a Spanish university based on an EDR process. *UTE. Teaching & Technology, Universitas Tarraconsensis*, 2025(2). <https://doi.org/10.17345/ute.2025.4161>
- Viñoles-Cosentino, V., Sánchez-Caballé, A. y Esteve-Mon, F.M. (2022). Desarrollo de la competencia digital docente en contextos universitarios. Una revisión sistemática. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 20(2). <http://doi.org/10.15366/reice2022.20.2.001>
- Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, 53(4), 5–23. <https://doi.org/10.1007/BF02504682>
- Weller, M. (2013). The battle for open—A perspective. *Journal of Interactive Media in Education*, 2013(3), Article 15. <https://doi.org/10.5334/2013-15>

Wiley, D., & Hilton, J. L., III (2018). Defining OER-enabled pedagogy. *The International Review of Research in Open and Distributed Learning*, 19(4).

<https://doi.org/10.19173/irrodl.v19i4.3601>

Zou, R., Jiang, L., & Wider, W. (2025). Bibliometric insights into the open education landscape. *The International Review of Research in Open and Distributed Learning*, 26(1), 283–309.

<https://doi.org/10.19173/irrodl.v26i1.7953>



August – 2026

Structure and Content of Interactions: Social Presence in Asynchronous Online Whole-Class vs. Small-Group Discussions

K. Bret Staudt Willet¹, Chufeng Bai¹, Jaesung Hur¹, and Mete Akcaoglu²

¹Florida State University; ²Georgia Southern University

Abstract

This study investigates how the format of asynchronous online class discussions (whole class vs. small group) affects the structure and content of student interactions. Data were collected from 8 weeks of discussion forums in a graduate-level, asynchronous, online course at a comprehensive university in the southeastern United States. Social network analysis (SNA) metrics, such as response rate, reciprocity, transitivity, diameter, density, in-degree, and closeness centrality, were used to evaluate interaction structure. Linguistic Inquiry and Word Count (LIWC) software was used to analyze the linguistic features of the interaction content. Findings revealed differences between the formats: whole-class discussions exhibited greater diameter and variability in response rate, emphasizing formal and analytical communication. In contrast, small-group discussions displayed higher reciprocity and transitivity, as well as more support and socioemotional engagement. Findings underscore the importance of designing and facilitating collaborative online learning experiences by strategically employing discussion formats to meet specific learning objectives. Implications for practice include incorporating both whole-class and small-group discussions to leverage the strengths of each format, such as creativity and critical thinking in larger groups and mutuality and support in small groups. Future research should examine these effects across different educational levels and subject areas, considering instructional contexts and personal learner characteristics.

Keywords: online learning, class discussion, group size, social network analysis, natural language processing

Introduction

Online courses are increasingly viewed as an important part of modern higher education. With the advancement of technology and changing times, more students are participating in online courses. According to the National Center for Education Statistics, more than half of students in higher education in the United States (approximately 10 million students, or 54.2%) participated in online courses in fall 2022 (2023), twice as many as a decade earlier, with five million students in fall 2013 (2015).

Online learning presents both advantages and challenges. A primary advantage is its flexibility and accessibility, allowing students from various regions to access quality education and teaching materials without the constraints imposed by physical distance and time (Ho et al., 2023). In addition, online learning supports the integration of diverse teaching tools, including multimedia, simulations, and interactive modules, thereby enriching the educational experience. However, online education also presents persistent challenges, including technical support, assessment integrity, and the necessity for additional teacher training (Palau et al., 2021). Furthermore, learners' experiences of social presence are often diminished in online learning. Online learning often features less robust communication and interaction than face-to-face instruction, which may result in decreased motivation, participation, and satisfaction (Sung & Mayer, 2012). Less engagement can lead to some learners feeling less connected and becoming less actively involved in their online educational journey (Ho et al., 2023).

Discussion forums are one of the most common course activities used to increase interaction and social presence in asynchronous online courses. Beyond their social benefits, discussion forums also help students understand course objectives, provide real-world applications, and can be used to promote student engagement (Wickershaw & Dooley, 2006). Although instructor feedback can help increase student engagement, peer feedback is equally important (Salter & Conneely, 2015).

Group discussions can enhance students' social skills and their awareness of societal issues (Akcaoglu & Lee, 2016; Steponavičius et al., 2023), but the size of these groups influences student engagement. For instance, larger discussion groups have been associated with lower student participation, poorer individual performance, and weaker team cohesion; conversely, smaller groups tend to yield better outcomes (Zhan et al., 2022). Recent research in this area has tended to focus on the impact of group size differences on learners' knowledge and skills (Elaish et al., 2023), critical thinking (Yang et al., 2022), and engagement levels (Luo et al., 2023; Zhan et al., 2022).

In our study, we extended research on online discussion group size to include the nature of the discussions themselves. Specifically, we compared the discussion group size between the format of whole-class discussions versus those in small groups. The purpose of this study was to investigate how the format of asynchronous online class discussions affects the structure and content of student interactions.

Framework

We have framed this study of online class discussions in terms of Kreijns et al.'s (2022) conceptualization of *social interaction*. They began with Short et al.'s (1976) foundational definition of *social presence*, which aimed to explain how different media types shape interpersonal interactions—specifically, the degree to

which an individual is perceived as “real” and salient in mediated communication. Garrison et al. (2000) applied social presence to online learning contexts, emphasizing learners’ ability to project social and emotional aspects of themselves to help other students to perceive them as real people. Rourke et al. (1999) made the concept more concrete, operationalizing social presence through the assessment of interactive content in asynchronous online discussions, categorized as *affective* (e.g., expression of emotions), *interactive* (e.g., continuing a thread), and *cohesive* (e.g., referring to the group using inclusive pronouns) indicators. Finally, Kreijns et al. (2022) argued that with advances in communication technologies since the time of Short et al.’s (1976) original definition, social presence is now better understood in terms of *being present*—that is, in the here and now—rather than in terms of being real.

Recent research presents social interaction as a multidimensional construct that includes factors in addition to social presence. Even three decades ago, Gunawardena (1995) noted that social presence depends not only on the medium itself but also on how participants communicate within it. Kreijns et al. (2022) further clarified this point by reconceptualizing social interaction as social presence in dynamic interplay with social space and sociability. They defined *social space* as the quality of interpersonal and emotional connections between group members and *sociability* as the affordance of the medium or technology to allow for the expression of social presence. In our study, we investigated social interaction in online discussion groups by considering social space through indicators of *interaction structure* (e.g., frequency and mutuality of communication between students) and sociability through indicators of *interaction content* (e.g., the affective tone in the expression of information).

Literature Review

Social Presence in Online Education

Social presence, as a key aspect of social interaction (Kreijns et al., 2022), plays a critical role in relationship-building for student participation and online learning inclusion (Lowenthal et al., 2020). Social presence is influenced not only by technology (e.g., communication media) but also by social context and personal factors (Gunawardena, 1995; Kim et al., 2011). Social presence brings specific benefits to group learning, such as comparing notes and correcting misunderstandings, which helps build trust and learner communities (Lim, 2023; Richardson et al., 2017). Considering the seven main motivations for students to participate in online courses, Van Wart et al. (2020) found that although social presence was a relatively low-ranking factor among the seven motivations, it was closely related to online course satisfaction, course participation, and interaction between students. Similarly, Richardson et al.’s (2017) meta-analysis showed that students who experience higher levels of social presence in online environments experience higher levels of learning satisfaction, engagement, academic self-efficacy, self-regulated learning, and learning performance. Student-student interactions and teacher-student interactions increase social presence, which then reduces learners’ social isolation and loneliness and makes students more engaged in learning activities (Phirangee & Malec, 2017). However, despite the importance of social presence, online learning has been criticized for lacking social presence and having lower sociality compared to face-to-face learning environments (Adnan & Anwar, 2020).

To increase social presence in online learning environments, discussion forums can be established as a space for interaction between students—a viable alternative to traditional face-to-face focus groups (Lowenthal & Dunlap, 2020). However, past research has shown that group size in computer-mediated environments influences students' learning activity participation and learning performance (Luo et al., 2023; Yang et al., 2020). Group size can impact group cohesion, student learning, and social perception; smaller group sizes have been connected to greater perceived social presence, more sense of community, and better learning outcomes (Akcaoglu & Lee, 2016; Michaelsen et al., 2023). Small groups allow members to take more responsibility and be more invested in outcomes (Qiu et al., 2014); students in this format view and comment on their peers' posts more than when in whole-class discussions (Yang et al., 2020). Students in small groups experience less information overload in addition to more social presence; when assigned to small groups, following discussion threads is easier compared to when in whole-class discussions (Qiu & McDougall, 2015; Yang et al., 2020). The reduced reading load, in turn, creates more time for reflection and thoughtful responses.

Measuring Interaction Structure Through Social Network Analysis

Although there are various methods to measure social interaction in online courses, social network analysis (SNA) has been shown to be an effective quantitative approach. Specifically, SNA examines the relationships between social entities and their influence within a network (Dado & Bodemer, 2017; Norz et al., 2023) by considering a set of objects (i.e., nodes) and their connections (i.e., edges). A node is an individual engaged in networks, and an edge represents an interaction between individuals (Dado & Bodemer, 2017; Ye & Pennisi, 2022). Through SNA, researchers can identify an influential node that has more connections, examine the strength of collaboration in the network, and illustrate the structure of connections in the network (Ye & Pennisi, 2022). Because SNA is calculated from observations of a learning environment, it reflects the structure of real interactions rather than perceived ones (Oh et al., 2025).

Past SNA research has used a variety of metrics to examine social networks, including measures of individuals' roles within the network and measures of the overall network structure (e.g., reciprocity, transitivity, diameter, density). In terms of individuals' roles, there are two kinds of centrality: degree and closeness. *Degree centrality* quantifies the connections between nodes; a higher degree-centrality score means that a node has more direct links to other nodes. Furthermore, *in-degree* refers to the number of incoming connections pointing to a node (e.g., how many responses a student receives to a discussion post), and *out-degree* refers to the amount of outgoing connections starting from a node, such as the total number of times a student replies to any of their classmates (Ye & Pennisi, 2022). In addition, *closeness centrality* calculates the proximity of a node to others, which can indicate the key influencers in a network (Ye & Pennisi, 2022).

As a measure of the overall network structure, *reciprocity* measures the relationship between a pair of nodes in terms of the probability of mutual connections; higher reciprocity indicates more back-and-forth conversations, whereas lower reciprocity suggests more one-directional conversations in the network (Staudt Willet & Carpenter, 2020). *Transitivity* is a measure of clustering and group cohesion in the network; higher transitivity indicates that adjacent nodes are more likely to be connected, and lower transitivity means that discussions are more evenly distributed across the whole network (Staudt Willet &

Carpenter, 2020). *Diameter* indicates the size of the network, as understood by the number of nodes required to traverse from one end of the network to the other. Diameter provides insight into the proximity of two network participants by representing the maximum distance between two nodes (Tabassum et al., 2018). Finally, *density* calculates the number of edges present between nodes out of all connections possible in the network (Norz et al., 2023). Density is an indicator of how actively individuals communicate or collaborate with each other; higher density indicates a highly connected network, and lower density implies a loosely connected network.

Researchers have conceptualized online learning communities as social networks and used SNA to investigate the complexity of relationships and connections among network members—that is, students' participation and interactions (Dado & Bodemer, 2017; Lim, 2023). To date, numerous studies have established SNA as an effective tool for mapping social interactions within text-based asynchronous online discussions (Ye & Pennisi, 2022); researchers have also used SNA metrics to analyze social presence, specifically, in online learning environments (e.g., Norz et al., 2023). For instance, Lim (2023) examined SNA metrics such as out-degree, in-degree, betweenness centrality, closeness centrality, and eigenvector centrality to understand the relationship between social presence and students' cognitive or affective learning outcomes. Similarly, Satar and Akcan (2018) found a statistically significant correlation between SNA metrics and social presence indicators.

Measuring Interaction Content Through Linguistic Features

In online discussions, students communicate through written or verbal language. Researchers have claimed that the content of students' interactions is more important than the frequency of posts and comments (Gasell et al., 2022). Previous research has studied students' interactions through qualitative content analysis or thematic analysis (e.g., Rourke et al., 1999). However, Pennebaker et al. (2007) proposed a different approach: to computationally analyze the content of text data by examining psychological characteristics evident through linguistic features. To accomplish this, they developed Linguistic Inquiry and Word Count (LIWC), a software program that categorizes text according to more than 80 linguistic features, such as emotional tone, thinking styles, and linguistic functions (<https://www.liwc.app/>). By examining linguistic features, researchers can gain information on one's attention, emotional status, social relationships, social hierarchy, social coordination and group processes, honesty and deception, close relationships, thinking styles, or individual differences (Tausczik & Pennebaker, 2010).

Although computationally analyzing linguistic features loses the nuanced insights of qualitative coding by hand, it gains the advantage of scalability, automating the assessment of linguistic features across potentially large amounts of text data. This approach can still provide insights into group interactions and dynamics (Scholand et al., 2010), and researchers have used LIWC to study social and cognitive presence in online environments. For instance, Joksimović et al. (2014) used LIWC to analyze 1,747 messages in the online discussion forum of a graduate-level course, finding that linguistic features can serve as indicators of different phases of cognitive presence: triggering, exploration, integration, and resolution. Comparing the linguistic features in user comments on two online platforms (i.e., YouTube and edX) at a larger scale of 6,938 comments from 2,889 contributors, Gao et al. (2024) found that social presence appeared mostly

similar on both platforms. Differences in sentiment were associated with community characteristics and expectations related to the technological affordances of each platform.

Purpose and Research Questions

Although previous research has investigated social interactions and social presence in online learning environments, it remains unclear whether group size and format, such as assigning students to small-group discussions, are associated with the nature of those discussions. The literature on group size in online discussions has largely focused on learning outcomes, participation frequency, or cognitive performance; few studies have examined how discussion formats (e.g., whole class vs. small group) influence the interaction structure and content in asynchronous environments.

To fill this research gap, the purpose of this study was to investigate how the format of asynchronous online class discussions (whole class vs. small group) affects the structure and content of student interactions. This study went beyond self-reported explanations of social interaction and social presence in asynchronous online discussions; we followed a computational text analysis approach to investigate real, observable behavior. To accomplish this purpose, we posed two research questions:

- RQ1. How does the structure of social interactions differ between small-group and whole-class discussions?
- RQ2. How does the content of social interactions differ between small-group and whole-class discussions?

Method

This research followed a quantitative case study design where the case was an online, graduate-level, asynchronous course. Students participated in whole-class discussions in weeks 1–4 and in randomly generated small groups for Weeks 5–8; each small group consisted of four or five class members. Discussion prompts asked students to summarize their learning from the week's materials and reflect on how it related to their practice. They were also required to comment on at least one class or group member's original post.

Data Collection

The data for the current study involved a retrospective analysis of existing student discussion posts that were originally produced as part of regular assignments in the graduate-level course Assessment and Data Analysis in Teaching at a comprehensive university in the southeastern United States. Online class discussions occurred asynchronously in the learning management system for the course, BrightSpace. No data were accessed or analyzed until Institutional Review Board (IRB) approval was obtained following the conclusion of the course. Upon IRB approval, all discussions were fully deidentified prior to analysis. The instructor (a co-author on this paper) exported the discussion forums as text data containing the content of posts and responses, the discussion format, and information regarding senders and receivers. Notably, although the data came from the fall 2014 version of the course, the core structure of this asynchronous online course remains unchanged in that it still covers the same content, same assignments, and discussion structure. In total, the dataset included 460 posts and responses from 19 students.

Data Analysis

We used R (Version 4.5.2) for data cleaning and to conduct all statistical analyses. To answer our first research question regarding the structure of social interactions, we used sender-and-receiver information (i.e., posts and responses) to investigate the social network structures of these discussions as indicated by SNA metrics. The unit of analysis was the number of posts and responses, as well as their connections to each other. We calculated descriptive statistics (e.g., mean, standard deviation, range, median) of the weekly number of posts, response rate (the percentage of posts that received at least one response), reciprocity (the likelihood of mutual connections), transitivity (how clustered interactions were), diameter (how disconnected the furthest two participants were), density (how many actual interactions there were relative to all possible interactions), in-degree (how connected participants were through receiving responses), and closeness centrality (proximity of participants to others, indicating influence). We checked the skewness and kurtosis of each measure; nearly all had a kurtosis score indicating a non-normal distribution. Therefore, we descriptively compared the medians when reporting the findings.

To answer our second research question regarding the content of social interactions, we created a data visualization to compare the likelihood that the most common terms appeared in whole-class versus small-group discussions. We calculated log-odds ratios for each term to show in which discussion format the term was more likely to appear. We also used the 2022 version of LIWC (Pennebaker et al., 2022) to analyze the linguistic features of discussion posts and responses. With LIWC, we calculated the median of numerous linguistic features (Table 1; Boyd et al., 2022) across 460 messages in the online discussion forum. We compared the presence of LIWC features in whole-class versus small-group discussions using a series of Mann-Whitney U tests, specifically Wilcoxon rank sum tests as the special case for two samples (Hollander et al., 2014). We used these non-parametric tests because LIWC features often violated the assumption of normality required for parametric *t*-tests.

Table 1

LIWC Categories and Descriptions

Category	Description
Word count	Total number of words
Words per sentence	Mean number of words per sentence
Big words	Percentage of words seven letters or longer
Dictionary words	Percentage of words captured by LIWC
Analytical thinking	Words indicating logical, formal thinking
Clout	Words indicating leadership, status
Authentic	Words indicating honesty, genuineness
Emotional tone	Degree of positive or negative tone
Linguistic	Words including pronouns (e.g., I, they) and function words (e.g., the, to, and)
Drives	Words indicating power, affiliation, achievement
Cognition	Words indicating thinking processes: insight, causation, differentiation

Category	Description
Affect	Words indicating emotional processes (positive and negative): anxiety, anger, sadness
Social processes	Words indicating social processes: communication, politeness, conflict
Culture	Words indicating facets of human life: politics, ethnicity, technology
Lifestyle	Words indicating leisure, home, work, school, money, religion
Physical	Words indicating health, wellness, illness, mental health, sex, food, death
Perception	Words indicating attention, motion, sensory
Conversational	Words indicating netspeak, assent, fillers (e.g., yeah, oh, okay)

Note. LIWC = Linguistic Inquiry and Word Count software; *netspeak* means informal, online language.

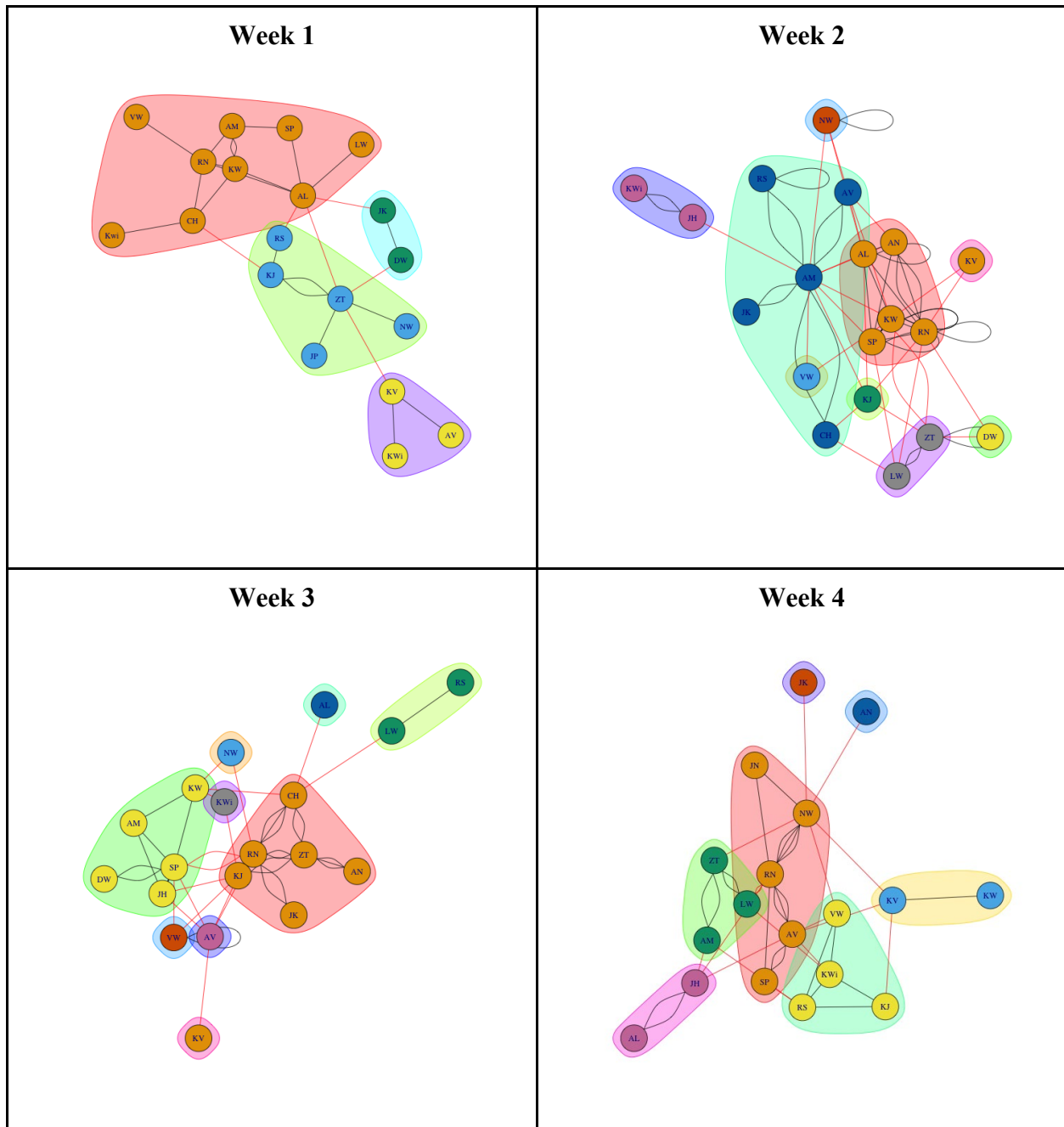
Results

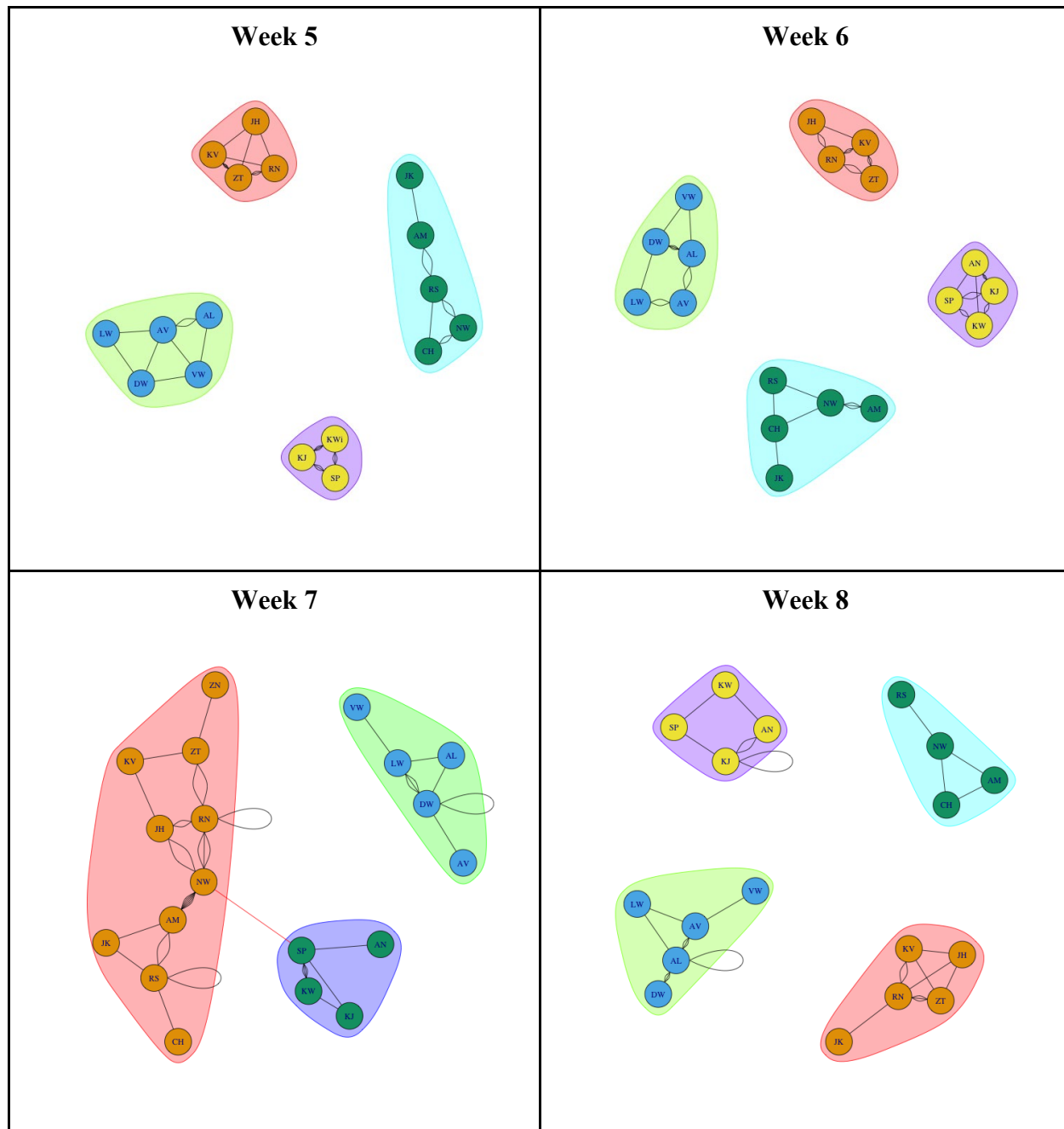
RQ1. Structure of Social Interactions

The analysis of the network structure over 8 weeks showed a shift in interaction dynamics as the class transitioned from a whole-class model (weeks 1–4) to small-group interactions (weeks 5–8), leading to compact and well-defined clusters (Figure 1). In weeks 1–4, most students' interactions were spread out across a wide range of peers, with several isolated students with peripheral engagement. However, their interactions became tighter and more interconnected in clearly bounded small groups over the remaining weeks. One notable exception occurred in week 7, when student NW responded to posts in their assigned small group as well as a post in another group. NW's crossover activity served as a bridge between those two small groups, causing the network diagram to depict those two groups as one.

Figure 1

Network Structure Across Eight Weeks (Weeks 1–4: Whole Class, Weeks 5–8: Small Group)





Note. Letters in the circles represent students in the course. Lines between the circles indicate students responding to other students in the weekly discussion. A line that loops back to where it begins indicates a student who responded to their own discussion post. The colorful clouds around the circles indicate computationally derived clusters in which students are more likely to have responded to other students within the cluster than those outside it.

SNA metrics (Table 2) provide further insight into the differences between whole-class and small-group discussions evident in Figure 1. For instance, the initial, whole-class phase when class discussions exhibited a more loosely connected and distributed network structure was supported by the calculation of the median network diameter (6.0 compared to 3.5), indicating that whole-class discussions spanned many students but were not uniformly connected. In addition, although the response rate was similar between whole-class

and small-group discussions (median of 69.8% compared to 66.4%), the whole-class format exhibited much more variability, with a response rate ranging from 59.6%–75.3% across the 4 weeks, whereas the small-group response rate only ranged from 63.0%–69.5%. This suggests that students were just as likely to receive responses in either format; however, students' interactions were unevenly distributed, with a concentration around a few individuals in the whole-class format.

After the transition to a small-group format in weeks 5–8, we observed notable changes in reciprocity and transitivity. Median reciprocity scores increased from 35.2% to 56.0%, and median transitivity rose from 21.9% to 63.3%; these metrics suggest stronger mutual engagement and more cohesive clustering within the smaller groups. Additionally, the decreased network diameter signifies shorter paths and more focused interactions within groups. Similarly, the median closeness centrality score rose from 0.4 to 0.8, indicating more direct and prompt interactions among discussion participants. However, despite these marks of increased interactivity, the overall volume of contributions decreased slightly from a median of 59.5 posts during whole-class weeks to 58.0 posts during small-group weeks. A larger difference was seen between the maximum weekly posts—the most in a whole-class discussion week was 73 posts, compared to a maximum of 59 posts during small-group discussions. Despite fewer initial posts in small groups, students still received a similar number of responses from classmates, as shown by an identical median in-degree of 2.0.

Table 2

Descriptive Statistics for Social Network Analysis Metrics

Variable	Whole class				Small groups			
	<i>M</i>	<i>SD</i>	Range	<i>Mdn</i>	<i>M</i>	<i>SD</i>	Range	<i>Mdn</i>
Weekly posts	59.8	10.6	47–73	59.5	55.2	6.2	46–59	58.0
Response rate	68.6%	6.7%	59.6–75.3%	69.8%	66.3%	2.9%	63.0–69.5%	66.4%
Reciprocity	34.5%	10.6%	21.4–46.2%	35.2%	52.0%	11.3%	35.7–60.0%	56.0%
Transitivity	22.5%	5.2%	17.5–28.6%	21.9%	58.3%	20.8%	29.0–77.4%	63.3%
Diameter	5.8	1.3	4–7	6.0	3.8	1.0	3–5	3.5
Density	14.6%	2.7%	11.7–18.0%	14.4%	12.4%	2.0%	9.5–13.6%	13.2%
In-degree	2.4	2.3	0–10	2.0	2.1	1.7	0–8	2.0
Closeness	0.5	0.2	0.2–1.0	0.4	0.7	0.2	0.3–1.0	0.8

Overall, the analysis of the structure of social interactions seems to indicate fewer, but more meaningful, exchanges during small-group class discussions. That is, students in small-group discussions received marginally fewer responses to their posts but were much more likely to have mutual exchanges (i.e., higher reciprocity).

RQ2. Content of Social Interactions

To answer the second research question, we analyzed the linguistic differences of discussions in whole-class versus small-group format. First, we investigated the frequency of words used in whole-class compared to small-group discussions (Figure 2). In the 460 discussion posts and responses across 8 weeks, students

discussions showed statistically significantly higher LIWC measures of word count, words per sentence, dictionary words, emotional tone, linguistic, drives, affect, social processes, and conversational. These results align with the observations of word frequencies in Figure 2, particularly that small-group discussions seem to be characterized by more social engagement (as indicated by social processes), positivity (as indicated by emotional tone), and emotional processes (as indicated by affect).

Table 3

Median Scores of LIWC Features and Results of Mann-Whitney U Tests (Wilcoxon Rank Sum)

Feature	Whole-class <i>Mdn</i>	Small-group <i>Mdn</i>	<i>W</i>	<i>p</i>
Word count	64	75	29,359.5	.038
Words per sentence	15.62	16.50	29,793.0	.018
Big words	23.53	21.92	20,916.5	< .001
Dictionary words	88.96	90.99	32,693.0	< .001
Analytical thinking	49.68	47.98	25,090.5	.355
Clout	40.06	42.97	29,036.0	.065
Authentic	68.87	65.87	23,323.5	.030
Emotional tone	46.50	61.50	31,877.5	< .001
Linguistic	70.47	72.28	31,362.5	< .001
Drives	3.90	4.85	32,304.5	< .001
Cognition	17.91	17.62	23,807.0	.068
Affect	3.04	4.00	33,291.0	< .001
Social processes	11.30	12.18	29,900.0	.014
Culture	0.00	0.00	27,664.5	.269
Lifestyle	7.14	6.45	26,403.5	.997
Physical	0.00	0.00	26,158.0	.825
Perception	5.88	6.38	28,104.5	.234
Conversational	0.00	0.00	28,945.0	.044

Discussion

The purpose of this study was to investigate how the format of asynchronous online class discussions (whole-class vs. small-group) affected the structure and content of student interactions. Although previous research has suggested that small-group discussions lead to increased student perceptions of social presence and sociability (Akcaoglu & Lee, 2016; Michaelsen et al., 2023), it was unclear whether different discussion formats would also lead to differences in the structure of social interactions (as measured by SNA) or the content of social interactions (as measured by LIWC assessment of linguistic features).

Findings from this study align with previous research that has shown the benefits of asynchronous online discussions for students, regardless of format (e.g., Yang et al., 2020). Still, our results do show observable

differences between whole-class and small-group formats in the structure of social interactions and statistically significant differences in the content of social interactions. Specifically, we found that whole-class discussions seemed to prioritize more formal and analytical communication, which is consistent with past research showing that whole-class discussions are better suited for tasks that require critical thinking and the synthesis of complex ideas (Cui & Teo, 2023). We also found that whole-class discussions had a higher median and notably higher maximum number of weekly posts, as well as a higher response rate. These differences may suggest an enthusiasm generated by a larger audience, which aligns with past research demonstrating the creative and divergent thinking afforded by large discussion groups (Yang et al., 2020) and participants' intent to appeal to a variety of perspectives, as well as, perhaps, attempts to impress their peers (Cui & Teo, 2023).

The differences in the structure and content of social interactions we observed in small-group discussions suggest a complementary set of benefits to those in whole-class discussions. For instance, higher reciprocity, transitivity, and closeness centrality scores as well as LIWC measures of social processes and positive emotional tone highlight the role that smaller conversation clusters can play in social presence: fostering closer connections, decreasing concerns about group members not contributing, and reducing feelings of isolation in online learning (Luo et al., 2023; Zhan et al., 2022). These benefits may help build trust, social skills, and a sense of community, which are critical for keeping learners engaged in asynchronous environments (Lim, 2023; Richardson et al., 2017). Our findings align with past research suggesting that small-group discussions encourage participants to express their ideas in more detail, promoting deeper engagement and thoughtful discussion (Akcaoglu & Lee, 2016; Qui & McDougall, 2015; Yang et al., 2020; Zhan et al., 2022) while offering a more comfortable social space for shy and reserved students to participate in discussions (Qiu et al., 2014).

Implications for Practice

The findings in this study demonstrate the complementary strengths of whole-class and small-group formats for online discussions, highlighting the importance of instructional design for shaping interaction patterns and optimizing collaborative learning environments (Cheung et al., 2021). Whole-class discussions can generate enthusiasm, fostering creativity (Yang et al., 2020) and critical thinking (Cui & Teo, 2023). Meanwhile, small groups can compensate for some of the perceived weaknesses of online teaching, such as limited social presence and students' feelings of isolation, leading to higher dropout rates and lower satisfaction compared to traditional learning contexts (Phirangee & Malec, 2017; Sung & Mayer, 2012).

Instructional designers and online instructors should consider incorporating both small-group and whole-class discussions into their courses, while writing learning objectives that make clear their expectations for social interactions and align with the strengths of the associated discussion format. For example, whole-class discussions encourage deeper creativity and critical thinking, and small-group discussions foster more reciprocity and support. Dividing students into small groups for discussion could create space for students to focus on mutual understanding and completing shared goals; this, in turn, could disrupt unhelpful discussion behaviors such as completing original posts and comments at the last minute. Meanwhile, whole-class discussions could be designed to stimulate and explore complex ideas.

Educators could apply the computational methods from this study to track and update social interactions in online courses. Building on Kreijns et al.'s (2022) conceptualization of social presence as a key factor for social interactions, we used SNA to investigate the structure of social interactions (i.e., social space) and LIWC to assess the content of social interactions (i.e., sociality). A learning analytics dashboard could automatically and continuously report SNA and LIWC metrics for students' interactions, providing instructors with opportunities to make adjustments during the course as needed and potentially increasing students' engagement as well (cf. Ramaswami et al., 2023). For instance, instructors could make adjustments to the discussion format if they notice trends in the content of social interactions, such as shifts in LIWC scores for social processes, emotional tone, drives, or cognition. In addition, SNA metrics would help instructors gain information on students' interaction patterns and be prepared to intervene, such as identifying isolated students and inviting them into small-group discussions or encouraging other students to respond to their posts.

Limitations and Future Research

Although this study extends previous research (Akcaoglu & Lee, 2016), there remain limitations. First, the dataset is limited to a single course and is dated, as the course discussions were collected in fall 2014. Although the instructor (a co-author on this paper) has maintained the overall design and discussion prompts for this course across the past decade, it is possible that student engagement and interactions may have changed over the years. Future research should follow a similar method for a more current comparison and offer perspectives from a different course at a different type of university, in a different national context for higher education outside the United States, or at the undergraduate level. The methods from the present study, specifically SNA and LIWC metrics, would scale well for analysis of much larger courses or multiple contexts. However, these computational methods are limited by collecting data only through observations; future studies would benefit from validating observational data with self-reported data from students and instructors.

Second, our approach only evaluates social interaction in terms of structure and content. Discussions could have been shaped by how the activity prompts were presented and facilitated. Thus, our findings may be affected by course settings and other contextual factors in addition to the discussion format. It was beyond the scope of the current study to control for confounding variables; however, future studies should investigate the order of offering discussion formats (e.g., small groups before whole-class discussions) as well as other factors, such as course subjects and national context. Similarly, participants' personality features may have had an effect. For instance, shy students might have talked more in small-group discussions, feeling comfortable with fewer people. Future research should consider controlling for personality characteristics and psychological state during statistical analysis.

Conclusion

More students than ever are taking online courses, despite critiques of social presence in online learning. The findings in this study align with past research, providing evidence of the positive benefits of online asynchronous discussions. Here, we found different but complementary strengths of whole-class and small-group discussion formats. Extending Kreijns et al. (2022), we measured social interactions as social presence in conjunction with social space and sociality, using computational methods of SNA to analyze

interaction structure (social space) and LIWC to assess interaction content (sociality). Our study serves as a proof-of-concept for a quantitative approach to computational text analysis that can be updated and scaled up to investigate very large courses or compare multiple contexts. As opportunities for online learning expand and the affordances of educational technologies increase, the structure and content of social interactions will remain key considerations for understanding the various factors that influence success in online learning.

References

- Adnan, M., & Anwar, K. (2020). Online learning amid the COVID-19 pandemic: Students' perspectives. *Journal of Pedagogical Sociology and Psychology*, 2(1), 45–51.
<https://doi.org/10.33902/JPSP.2020261309>
- 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), 1–17. <https://doi.org/10.19173/irrodl.v17i3.2293>
- Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). *The development and psychometric properties of LIWC-22*. University of Texas at Austin.
<https://www.liwc.app/static/documents/LIWC-22%20Manual%20-%20Development%20and%20Psychometrics.pdf>
- Cheung, S. K. S., Kwok, L. F., Phusavat, K., & Yang, H. H. (2021). Shaping the future learning environments with smart elements: Challenges and opportunities. *International Journal of Educational Technology in Higher Education*, 18, Article 16. <https://doi.org/10.1186/s41239-021-00254-1>
- Cui, R., & Teo, P. (2023). Thinking through talk: Using dialogue to develop students' critical thinking. *Teaching and Teacher Education*, 125, Article 104068.
<https://doi.org/10.1016/j.tate.2023.104068>
- Dado, M., & Bodemer, D. (2017). A review of methodological applications of social network analysis in computer-supported collaborative learning. *Educational Research Review*, 22, 159–180.
<https://doi.org/10.1016/j.edurev.2017.08.005>
- Elaish, M. M., Shuib, L., Hwang, G.-J., Ghani, N. A., Yadegaridehkordi, E., & Zainuddin, S. Z. (2023). Mobile English language learning: A systematic review of group size, duration, and assessment methods. *Computer Assisted Language Learning*, 36(3), 430–456.
<https://doi.org/10.1080/09588221.2021.1931341>
- Gao, J., Wei, S., & Dubé, A. K. (2024). Evaluating learners' online learning experience of informal learning environments: A LIWC analysis. *The Journal of Applied Instructional Design*, 13(2), 63–67.
<https://doi.org/10.59668/1269.15691>
- Garrison, R. (2000). Theoretical challenges for distance education in the 21st century: A shift from structural to transactional issues. *The International Review of Research in Open and Distributed Learning*, 1(1), 1–17. <https://doi.org/10.19173/irrodl.v1i1.2>
- Gasell, C., Lowenthal, P. R., Uribe-Flórez, L. J., & Ching, Y.-H. (2022). Interaction in asynchronous discussion boards: A campus-wide analysis to better understand regular and substantive

- interaction. *Education and Information Technologies*, 27(3), 3421–3445.
<https://doi.org/10.1007/s10639-021-10745-3>
- Gunawardena, C. N. (1995). Social presence theory and implications for interaction and collaborative learning in computer conferences. *International Journal of Educational Telecommunications*, 1(2), 147–166. <https://www.learntechlib.org/primary/p/15156/>
- Ho, H. C. Y., Poon, K.-T., Chan, K. K. S., Cheung, S. K., Datu, J. A. D., & Tse, C. Y. A. (2023). Promoting preservice teachers' psychological and pedagogical competencies for online learning and teaching: The TEACH program. *Computers & Education*, 195, Article 104725.
<https://doi.org/10.1016/j.compedu.2023.104725>
- Hollander, M., Wolfe, D. A., & Chicken, E. (2014). *Nonparametric statistical methods* (3rd ed.). Wiley.
- Joksimović, S., Gašević, D., Kovanović, V., Adesope, O., & Hatala, M. (2014). Psychological characteristics in cognitive presence of communities of inquiry: A linguistic analysis of online discussions. *The Internet and Higher Education*, 22, 1–10. <https://doi.org/10.1016/j.iheduc.2014.03.001>
- Kim, J., Kwon, Y., & Cho, D. (2011). Investigating factors that influence social presence and learning outcomes in distance higher education. *Computers & Education*, 57(2), 1512–1520.
<https://doi.org/10.1016/j.compedu.2011.02.005>
- Kreijns, K., Xu, K., & Weidlich, J. (2022). Social presence: Conceptualization and measurement. *Educational Psychology Review*, 34(1), 139–170. <https://doi.org/10.1007/s10648-021-09623-8>
- Lim, J. (2023). Exploring the relationships between interaction measures and learning outcomes through social network analysis: The mediating role of social presence. *International Journal of Educational Technology in Higher Education*, 20, Article 14. <https://doi.org/10.1186/s41239-023-00384-8>
- Lowenthal, P. R., & Dunlap, J. C. (2020). Social presence and online discussions: A mixed method investigation. *Distance Education*, 41(4), 490–514.
<https://doi.org/10.1080/01587919.2020.1821603>
- Lowenthal, P. R., Greear, K., Humphrey, M., Lowenthal, A., Conley, Q., Giacumo, L. A., & Dunlap, J. C. (2020). Creating accessible and inclusive online learning: Moving beyond compliance and broadening the discussion. *Quarterly Review of Distance Education*, 21(2).
<https://www.infoagepub.com/products/Quarterly-Review-of-Distance-Education-21-2>
- Luo, H., Chen, Y., Chen, T., Koszalka, T. A., & Feng, Q. (2023). Impact of role assignment and group size on asynchronous online discussion: An experimental study. *Computers & Education*, 192, Article 104658. <https://doi.org/10.1016/j.compedu.2022.104658>

- Michaelsen, L. K., Knight, A. B., & Fink, L. D. (2023). *Team-based learning: A transformative use of small groups in college teaching*. Taylor & Francis.
- National Center for Education Statistics. (2015, December). *Number and percentage of students enrolled in degree-granting postsecondary institutions, by distance education participation, location of student, level of enrollment, and control and level of institution: Fall 2013 and fall 2014*. Digest of Education Statistics. https://nces.ed.gov/programs/digest/d15/tables/dt15_311.15.asp
- National Center for Education Statistics. (2023, December). *Number and percentage of students enrolled in degree-granting postsecondary institutions, by distance education participation, location of student, level of enrollment, and control and level of institution: Fall 2021 and fall 2022*. Digest of Education Statistics. https://nces.ed.gov/programs/digest/d23/tables/dt23_311.15.asp
- Norz, L.-M., Dornauer, V., Hackl, W. O., & Ammenwerth, E. (2023). Measuring social presence in online-based learning: An exploratory path analysis using log data and social network analysis. *The Internet and Higher Education*, 56, Article 100894. <https://doi.org/10.1016/j.iheduc.2022.100894>
- Oh, H., Liu, Z., & Xing, W. (2025). Do actions speak louder than words? Unveiling linguistic patterns in online learning communities using cross recurrence quantification analysis. In A. Martínez Monés, C. Mills, J. Jovanovic, & X. Ochoa (Chairs), *The 15th International Conference on Learning Analytics and Knowledge* (pp. 992–998). ACM. <https://doi.org/10.1145/3706468.3706569>
- Palau, R., Fuentes, M., Mogas, J., & Cebrián, G. (2021). Analysis of the implementation of teaching and learning processes at Catalan schools during the COVID-19 lockdown. *Technology, Pedagogy and Education*, 30(1), 183–199. <https://doi.org/10.1080/1475939X.2020.1863855>
- Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007). *The development and psychometric properties of LIWC2007*. University of Texas at Austin. <https://liwc.app/static/documents/LIWC2007%20Manual%20-%20Development%20and%20Psychometrics.pdf>
- Pennebaker, J. W., Boyd, R. L., Booth, R. J., Ashakkumar, A. & Francis, M. E. (2022). *Linguistic inquiry and word count: LIWC-22* [Computer software]. Pennebaker Conglomerates. <https://www.liwc.app>
- Phirangee, K., & Malec, A. (2017). Othering in online learning: An examination of social presence, identity, and sense of community. *Distance Education*, 38, 160–172. <https://doi.org/10.1080/01587919.2017.1322457>
- Qiu, M., Hewitt, J., & Brett, C. (2014). Influence of group configuration on online discourse writing. *Computers & Education*, 71, 289–302. <https://doi.org/10.1016/j.compedu.2013.09.010>

- Qiu, M., & McDougall, D. (2015). Influence of group configuration on online discourse reading. *Computers & Education*, 87, 151–165. <https://doi.org/10.1016/j.compedu.2015.04.006>
- Ramaswami, G., Susnjak, T., & Mathrani, A. (2023). Effectiveness of a learning analytics dashboard for increasing student engagement levels. *Journal of Learning Analytics*, 10(3), 115–134. <https://doi.org/10.18608/jla.2023.7935>
- Richardson, J. C., Maeda, Y., Lv, J., & Caskurlu, S. (2017). Social presence in relation to students' satisfaction and learning in the online environment: A meta-analysis. *Computers in Human Behavior*, 71, 402–417. <https://doi.org/10.1016/j.chb.2017.02.001>
- Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (1999). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education*, 14(2), 50–71. <https://www.learntechlib.org/p/92000/>
- Salter, N. P., & Conneely, M. R. (2015). Structured and unstructured discussion forums as tools for student engagement. *Computers in Human Behavior*, 46, 18–25. <https://doi.org/10.1016/j.chb.2014.12.037>
- Satar, H. M., & Akcan, S. (2018). Pre-service EFL teachers' online participation, interaction, and social presence. *Language Learning & Technology*, 22(1), 157–183. <https://doi.org/10.64152/10125/44586>
- Scholand, A. J., Tausczik, Y. R., & Pennebaker, J. W. (2010, February). Social language network analysis. In K. Inkpen & C. Gutwin (Chairs), *The 2010 ACM conference on computer-supported cooperative work* (pp. 23–26). ACM. <https://doi.org/10.1145/1718918.1718925>
- Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. John Wiley & Sons.
- Staudt Willet, K. B., & Carpenter, J. P. (2020). Teachers on Reddit? Exploring contributions and interactions in four teaching-related subreddits. *Journal of Research on Technology in Education*, 52(2), 216–233. <https://doi.org/10.1080/15391523.2020.1722978>
- Steponavičius, M., Gress-Wright, C., & Linzarini, A. (2023). *Social and emotional skills: Latest evidence on teachability and impact on life outcomes* (OECD Education Working Paper No. 304). OECD. <https://doi.org/10.1787/ba34f086-en>
- Sung, E., & Mayer, R. E. (2012). Five facets of social presence in online distance education. *Computers in Human Behavior*, 28(5), 1738–1747. <https://doi.org/10.1016/j.chb.2012.04.014>
- Tabassum, S., Pereira, F. S., Fernandes, S., & Gama, J. (2018). Social network analysis: An overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(5), Article e1256. <https://doi.org/10.1002/widm.1256>

- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24–54. <https://doi.org/10.1177/0261927X09351676>
- Van Wart, M., Ni, A., Medina, P., Canelon, J., Kordrostami, M., Zhang, J., & Liu, Y. (2020). Integrating students' perspectives about online learning: A hierarchy of factors. *International Journal of Educational Technology in Higher Education*, 17, Article 53. <https://doi.org/10.1186/s41239-020-00229-8>
- Wickersham, L. E., & Dooley, K. E. (2006). A content analysis of critical thinking skills as an indicator of quality of online discussion in virtual learning communities. *Quarterly Review of Distance Education*, 7(2), 185–193. <https://doi.org/10.1108/QRDE-06-2006-0007>
- Yang, T., Luo, H., & Sun, D. (2022). Investigating the combined effects of group size and group composition in online discussion. *Active Learning in Higher Education*, 23(2), 115–128. <https://doi.org/10.1177/1469787420938524>
- Ye, D., & Pennisi, S. (2022). Analysing interactions in online discussions through social network analysis. *Journal of Computer Assisted Learning*, 38(3), 784–796. <https://doi.org/10.1111/jcal.12648>
- Zhan, Z., He, G., Li, T., He, L., & Xiang, S. (2022). Effect of groups size on students' learning achievement, motivation, cognitive load, collaborative problem-solving quality, and in-class interaction in an introductory AI course. *Journal of Computer Assisted Learning*, 38(6), 1807–1818. <https://doi.org/10.1111/jcal.12722>



August – 2026

Blended Learning in a Post-Pandemic Context: A Cross-Sectional Study on Turkish EFL Learners' Perspectives

Burak Tomak and Ayşe Yılmaz Virkan
Marmara University, Istanbul, Türkiye

Abstract

During the COVID-19 pandemic, various institutions took measures to prevent the spread of the virus. Educational institutions contributed to these efforts by shifting from face-to-face to online learning. One such approach was blended learning, which served as a transitional model between the two. Blended learning has become an integral part of higher education, particularly in English as a foreign language (EFL) instruction. This study investigated Turkish EFL learners' perceptions of blended learning by comparing two different student cohorts from 2021–2022 and 2024–2025 using a cross-sectional research design. Data were collected through classroom observations and semi-structured interviews with 24 students in total. Although more than 75% of the students acknowledged the flexibility and accessibility of blended learning, persistent concerns remained regarding engagement, socialization, and assessment integrity in online settings. Over time, there has been a slight increase in student acceptance of blended learning, particularly in the areas of autonomous learning and digital literacy. However, many students still expressed a preference for face-to-face instruction, especially for productive skills like speaking and collaborative activities. The study contributed to the growing body of literature by offering a comparative perspective on the challenges and advantages of blended learning in EFL education and providing recommendations for optimizing its effectiveness.

Keywords: blended learning, cross-sectional research, digital maturity, EFL learners, learner autonomy, student engagement

Introduction

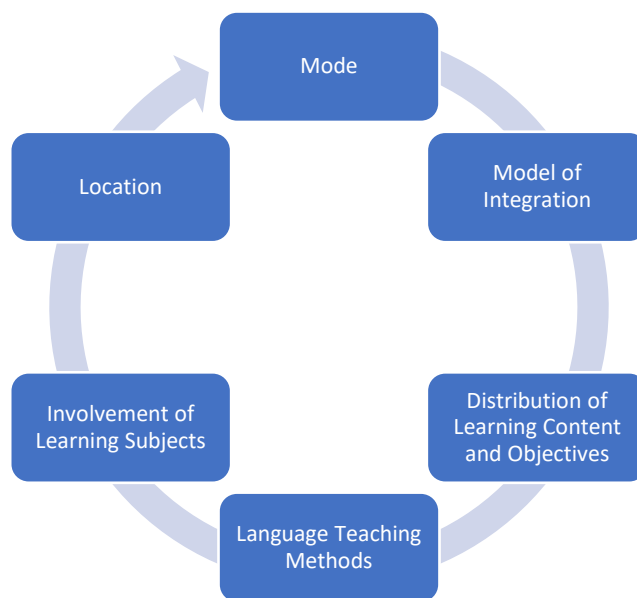
Around the world, the outbreak of the COVID-19 pandemic affected educational institutions who responded with measures much like those in sectors such as health, tourism, and industry. Their first action was to turn to online education and cover all lessons online. In time, some schools started to convert their online education into blended learning, which required both teachers and students to do their lessons online as well as face-to-face. This often meant organizing classes in both real and virtual settings on separate days of the week. What has followed can be considered the transition phase from pandemic (quarantine) life to normal life, with most universities integrating blended learning into their curricula.

Bonk and Graham (2006) described blended learning as the combination of online and face-to-face learning experiences, often leveraging digital tools to facilitate instruction and assessment. Blended instruction has become a trend in education to give students the advantages of each method, as it can also improve learning outcomes by providing personalized learning opportunities and increasing student engagement. Despite extensive research on blended learning, longitudinal cross-cohort analyses remain scarce, leaving open questions about how evolving digital maturity has shaped learner perceptions.

Neumeier (2005) reported that there were six parameters to design an effective blending learning environment: (a) mode (face-to-face or computer-assisted learning); (b) model of integration (optional or compulsory activities); (c) distribution of learning content and objectives; (d) language teaching methods online materials, online tutors, and face-to-face teachers; (e) involvement of learning subjects (types of interaction); and (f) location (physical space). Figure 1 illustrates these parameters.

Figure 1

Parameters of an Effective Design for Blended Learning Environments



This model was comprehensive enough to help us better understand why and how blending learning was adopted in a certain educational institution. This study was conducted in an institution that considered these aspects when blended learning was formally adopted in the 2021–2022 academic year.

Theoretical Framework

Grounded in cognitive load theory (CLT), digital literacy maturity can be considered a factor that helps reduce extraneous cognitive load in blended learning environments. When students become more familiar with learning management systems (LMS) and digital tools, they have fewer resources for germane processing and thus reduce the cognitive load, namely the mental effort that directly promotes learning (Sweller et al., 2019). In such conditions, studies have found that it is critical to cut extraneous load and match intrinsic load with learners' background knowledge for proper design (Mayer, 2002). Within EFL settings, such digital literacy maturity is likely to manifest as more efficient task execution and more positive perceptions of blended learning over time, particularly when instructional materials are structured to lower unnecessary technical complexity. From a constructivist perspective, students learn best through the active construction of meaning, synthesizing earlier knowledge, and working on genuine tasks (Machumu & Zhu, 2017). Awareness of digital media mitigates cognitive load and prompts collaborative and self-regulated learning for deeper insights. CLT imposes design restrictions, with constructivist theory focusing on learner autonomy that develops an understanding of perceptions with growing maturity on digital literacy.

Why Educators Opt for Blended Learning

Many studies have explained why educators applied blended learning in their language classes. One of the most important reasons is that it integrates technology such as educational videos and tools with traditional classroom procedures, which makes learning more attractive for students. Several studies have shown blended learning as more effective than face-to-face or online instruction (Bernard et al., 2014; Brodersen & Melluzzo, 2017). The instructor and technology have a great impact on students' satisfaction with online learning, and students have expressed satisfaction with blended learning when used for their English language competency (Suwantarathip, 2019).

There are many other advantages of blended learning, which can foster the development of students' critical and reflective thinking, along with the flexibility due to the successful combination of both in-class instructions and online lessons (Hirata & Hirata, 2008). Such studies have shown that blended learning contributed positively to the development of foreign language learners' linguistic competence.

The Influence of Blended Learning on Linguistic Skills

Hung (2015) claimed that blended learning enabled students to foster their listening, speaking, and reading comprehension skills, and the degree of learner autonomy improved both their speaking and pronunciation skills. For example, Faridi et al. (2020) claimed that students who took lessons using a hybrid model had better writing scores than a control group who took lessons with a traditional approach because of the effective use of technological platforms and active roles, easy feedback, increased student participation, and peer assessment. However, the influence of blended learning on the linguistic skills of learners in various contexts has differed. Hence, this study aimed to discover more about its effects on the skills development of Turkish learners of English.

Student Perspectives on Blended Learning

Many studies have indicated that students hold positive attitudes toward blended learning, and it has also been linked to increased motivation and learner autonomy. Alias et al. (2012) confirmed that hybrid learning environments encouraged interaction and collaboration, leading to greater motivation, self-confidence, and responsibility. Additionally, Doncheva (2015) underscored the benefits of flexibility and self-directed learning, allowing students to progress at their own pace. Similarly, Hirata and Hirata (2008) mentioned that online materials and interactive feedback increased language acquisition. However, the overwhelming majority of works emphasized the short-term effectiveness of blended learning which left a gap for long-term attitude change, which was addressed by this study.

Drawbacks of Blended Learning

Blended learning offers cognitive advantages but also has gaps that provoke dissatisfaction for students. One of the critical issues is reduced socialization with teachers (Redfern & Naughton, 2002) and with peers (Bernard et al., 2014). Similarly, Ja'ashan (2015) noted that students sometimes viewed it as a waste of time, prone to academic dishonesty, and socially isolating. Gascoigne and Parnell (2013) also noted that students missed the regular face-to-face interactions, especially critical during language learning. Dixon et al. (2021) observed that student performance was not significantly different in traditional and technologically enhanced learning environments. These findings highlighted that even if blended learning provided convenience, it did not necessarily replace the interactivity and the organization of conventional learning environments.

Significance of the Study

The blended learning literature is extensive, but few studies have compared student perceptions longitudinally. Our study was significant as it compared two cohorts and explored the development of perceptions with digital maturity, a rarely documented variable in the context of EFL environments. Different from earlier studies with single groups or short-term studies, this cross-section study explored whether students' attitudes changed as they had more exposure to blended learning over time.

By examining students' evolving perspectives, this study has contributed to the ongoing discussion on best practices for blended learning implementation in language education. Findings from this research have provided insights into how blended learning can be designed more effectively to enhance engagement, skill development, and assessment integrity in EFL contexts.

Research Aim and Questions

This study investigated the evolving perceptions of Turkish EFL learners regarding blended learning by comparing two different student cohorts from 2021–2022 and 2024–2025. The research examined how attitudes toward engagement, socialization, and skill development changed over time. The current study addressed the gap through its longitudinal cross-cohort design, allowing us to examine how digital literacy maturity shaped perceptions over time—a perspective largely absent in current EFL literature.

The following research questions guided this study.

1. Did Turkish EFL learners' perceptions of blended learning differ significantly between the 2021–2022 and 2024–2025 cohorts?
2. What factors contributed to the perceived effectiveness or ineffectiveness of blended learning over time?
3. How did students evaluate their skill development in different modalities of blended learning?
4. What challenges and advantages did students identify regarding engagement and interaction in blended learning environments?

Method

Research Design and Implementation

This study employed a cross-sectional research design to examine Turkish EFL learners' perspectives on blended learning at a state university in Istanbul, Turkey. This study allowed for a comparative analysis of the perceptions of learners over time while the main instruction was carried out in blended learning environments. This system was firstly applied in the 2021–2022 academic year. In Turkish higher education, blended learning was formally integrated into preparatory English programs as part of pandemic-era curricular adaptations regulated by the Council of Higher Education in Turkey. The last implementation was in 2024–2025 because then schools were allowed to return to fully face-to-face education for 2025–2026 academic year. The courses followed a weekly blended schedule combining synchronous online and face-to-face sessions. For example, a typical week included two face-to-face sessions and two asynchronous online sessions. Listening and grammar were taught online with asynchronous materials, and speaking and writing took place face-to-face for immediate interaction and confirmation. Pair-work activities were used for face-to-face classes, and online classes used individual practice with multimedia and automated feedback. Participation required physical attendance at the face-to-face classes, and online participation was encouraged with the use of cameras and structured prompts such as breakout activities.

Participants

The study was conducted in the School of Foreign Languages in one of the state universities located in Istanbul, where a blended learning system was formally implemented between 2021 and 2025. The initial study involved 12 participants, all enrolled in preparatory English courses, while the updated study incorporated an additional 12 students, allowing for a more comprehensive evaluation. Table 1 provides demographic details of both cohorts.

Table 1

Participants' Demographic Information

Cohort	Sex (M/F)	Age range	High school type	Department	Prep year in high school?	Previous blended learning experience
2021– 2022	6M/6F	18–21 years	Anatolian/Science/ Social Science	Engineering, Medicine, Economics	2 had a prep year	none
2024– 2025	7M/5F	18–20 years	Anatolian/Science/ Private	Engineering, Medicine, Business	3 had a prep year	1 student

The participants were selected through purposive sampling, ensuring a diverse representation of students from different academic disciplines. They were also asked whether they had a prep year in which they got intensive language classes in their high school and whether they had any familiarity with blended learning beforehand. This approach aimed to capture a broad range of student experiences while maintaining consistency in data collection. In both phases, all students were informed of the study's purpose and asked for their consent; ethical clearance was obtained from the university's research ethics committee. To ensure confidentiality, student names were anonymized, and all recorded data were securely stored with restricted access.

Data Collection

In both phases, data were collected using semi-structured interviews and classroom observations. This methodological consistency allowed for direct comparisons between the two cohorts. We interviewed each participant in Turkish and made sure that they could express themselves freely. The interviews were conducted online via Zoom, lasted approximately 45 minutes, and were recorded for transcription. The same set of questions was used across both cohorts to maintain consistency and allow for direct comparison.

The researchers conducted classroom observations, noting attendance patterns, engagement levels, and participation in both face-to-face and online lessons over the course of a full semester. These observations provided additional qualitative insights into how students adapted to the blended learning format. The observations focused on four key criteria: (a) frequency of participation in online and face-to-face sessions, (b) use of digital tools during online classes, (c) interaction levels with peers and instructors, and (d) perceived engagement based on instructor feedback. A structured observation checklist was used to ensure consistency across observations.

Data Analysis

Qualitative data were analyzed using thematic analysis through NVivo software. A combination of inductive and deductive coding was employed. Initially, an open coding approach was applied, allowing themes to emerge naturally from the data. This was followed by axial coding, in which themes were refined and

categorized based on their relevance to engagement, assessment, and skill development in blended learning environments. A chi-square test was also conducted to identify any significant difference in students' perceived satisfaction. To distinguish between cohorts, students from 2024–2025 were numbered from S13 onward.

Results

Comparison of Student Satisfaction Across Cohorts

To answer the first research question, a chi-square test was conducted to examine whether there was a significant difference in student perceptions of blended learning satisfaction between the 2021–2022 and 2024–2025 cohorts. Perceived satisfaction categories were derived from interview responses and observation notes, then coded dichotomously (satisfied / not satisfied) for chi-square analysis. As seen in Table 2, the results showed no statistically significant difference in terms of satisfaction between the two groups, $\chi^2 (df = 1) = 1.54, p = 0.214$.

Table 2

Results of Chi-square Test

Variable	2021–2022 Cohort (<i>n</i> = 12)	2024–2025 Cohort (<i>n</i> = 12)	Total (<i>N</i> = 24)	χ^2	<i>p</i> value
Satisfied (Yes)	5 (41.6%)	9 (75.0%)	14 (58.3%)	1.54	0.214
Not Satisfied (No)	7 (58.3%)	3 (25.0%)	10 (41.6%)		

Although the 2024–2025 cohort reported a higher satisfaction rate (75.0%) compared to the 2021–2022 cohort (41.6%), the difference was not significant, likely due to the small sample size and variability within groups.

The Efficiency of Blended Learning

As for the second research question, we investigated the perceived effectiveness of blended learning over time. Table 3 presents a comparison of students' responses regarding the efficiency of blended learning.

Table 3

Comparison of Students' Perceptions of Blended Learning Efficiency

Cohort	Efficient (%)	Partially efficient (%)	Not efficient (%)
2021–2022	2 (16.7%)	2 (16.7%)	8 (66.6%)
2024–2025	3 (25%)	6 (50%)	3 (25%)

While more students in 2024–2025 recognized some benefits of blended learning, the majority still expressed concerns about the lack of a real learning environment, limited social interaction, and concentration challenges in online sessions. One student stated that “it was no good for us not to have an attendance obligation in our face-to-face lessons as I could not socialize with my friends” (S5, 2021–2022). Another participant from the new cohort expressed a similar concern, saying “online lessons feel detached; I can’t engage with my classmates or the instructor in the same way” (S18, 2024–2025).

To better understand the results, we also examined student responses in detail (see Table 4) based on whether they found blended learning efficient, along with reasons for their evaluations, as well as the positive and negative aspects.

Table 4

The Efficiency of Blended Learning Across Two Cohorts

Participant	Efficient?	Reason	Positive aspect	Negative aspect
S1	No	Lack of discipline	Easy access	Concentration problems
S2	No	Lack of stability	No transportation waste	Lack of socializing
S3	No	No real learning environment	Convenience	Lack of discipline
S4	No	Lack of discipline	Easy access	No real learning environment
S5	Partially Yes	Active grammar teaching	No transportation waste	No cameras on, lack of socializing
S6	No	Lack of stability	No transportation waste	Technical problems
S7	No	Lack of discipline	More personal time	Isolation, concentration issues
S8	Yes	Easy tracking	Comfort, access	Lack of real learning environment
S9	Yes	Socializing face-to-face	No transportation waste	Concentration problems

Participant	Efficient?	Reason	Positive aspect	Negative aspect
S10	No	Lack of discipline, focus issues	Easy access	Lack of attendance enforcement
S11	No	Lack of engagement, real learning environment	No transportation waste	No active participation
S12	Partially Yes	Intensive lessons	No time/money waste	Lack of discipline, participation
S13	Partially Yes	No waste of time for transportation	No waste of time for transportation	Concentration problems
S14	Partially Yes	1 day off advantage	No waste of time for transportation	Less participation, lack of socializing
S15	Yes	More personal time	No waste of time for transportation	Concentration problems, lack of discipline
S16	Partially Yes	Wise use of technology	Easy access	No real learning environment, lack of discipline
S17	No	Lack of real learning environment	No waste of time for transportation	Less participation, lack of socializing
S18	Yes	1 day off advantage	No waste of time for transportation	Technical problems
S19	Partially Yes	Some flexibility	More personal time	Lack of real learning environment, concentration problems
S20	Partially Yes	Variety in methods	Comfort, access	Lack of real learning environment
S21	Partially Yes	Real-life preparation	No waste of time for transportation	Concentration problems
S22	No	Lack of discipline, focus issues	Easy access	Lack of attendance enforcement
S23	No	Lack of engagement, real learning environment	No waste of time for transportation	No active participation
S24	Yes	More time for self-study, easy teacher control	No waste of time and money	Lack of discipline, less participation

Among the positive aspects, the most frequently mentioned benefit was no transportation waste, cited by 10 students (41.7%) across both cohorts. This was followed by easy access, reported by six students (25%), and more personal time, which was valued by four students (16.7%). Some students also appreciated the ability to study independently, with more time for self-study and easy teacher control each noted by two students (8.3%). As one student from the latest cohort stated, “I didn’t like online learning at first, but I now realize that it saves me a lot of time” (S13, 2024–2025), reflecting an increased appreciation for time management benefits.

However, negative aspects remained dominant, with concentration problems being the most frequently reported issue, mentioned by seven students (29.2%). This was followed by lack of discipline (six students, 25%) and lack of a real learning environment (six students, 25%). Social engagement challenges were also apparent, with lack of socializing noted by four students (16.7%), and technical problems cited by three students (12.5%). One student stated that “even though I like the flexibility, I still find it hard to focus during online lessons” (S19, 2024-2025).

Students expressed mixed opinions about the efficiency of blended learning (see Table 5). While some appreciated the flexibility and accessibility it provided, others found it difficult to maintain engagement and follow the lesson structure.

Table 5

Variables That Affect the Efficiency of Blended Learning

Participant	Flow of lessons	Teacher motivation	Contact with teachers	Feedback efficiency	Material preference	Digital resource satisfaction
S1	Split	Low	More in face-to-face	Low in online	Face-to-face	✓
S2	Split	Low	More in face-to-face	Low in online	Face-to-face (better for writing)	✓
S3	Split	Low	Smooth	Efficient for attentive students	Mixed preference	✓
S4	Split	Improved with tech use	More in face-to-face	Easy	Hardcopy preferred	✓
S5	Good	Low	More in face-to-face	Low in online	Sufficient	✓
S6	Good	Normal	More in face-to-face	Low in online	Face-to-face	✓
S7	Split	Low	More in face-to-face	Low in online	Sufficient	✓
S8	Split	Low	Easy	Efficient for attentive students	Hardcopy preferred	(Technical issues)

Participant	Flow of lessons	Teacher motivation	Contact with teachers	Feedback efficiency	Material preference	Digital resource satisfaction
S9	Normal-fast online	Low	More in face-to-face	Low in online	Face-to-face	✓
S10	Split	Low	Limited	Low in online	Sufficient	✓
S11	Good	Low	More in face-to-face	Low in online	Sufficient	✓
S12	Split	Low	Easy	More efficient in face-to-face	Sufficient	✓
S13	Good	Positive	More in face-to-face	Low in online	Face-to-face	✓
S14	Split	Positive	More in face-to-face	Low in online	Electronic preferred	✓
S15	Good	Positive	Smooth	More efficient in face-to-face	Hardcopy preferred	✓
S16	Split	Low	More in face-to-face	Low in online	Hardcopy preferred	✓
S17	Split	Low	More in face-to-face	Low in online	Sufficient	✓
S18	Good	Normal	More in face-to-face	Low in online	Softcopy preferred	✓
S19	Good	Positive	More in face-to-face	More efficient	Mixed preference	✓
S20	Split	Low	Easy	Satisfactory	Hardcopy preferred	(Technical issues)
S21	Split	Low	More in face-to-face	Low in online	Sufficient	✓
S22	Split	Low	More in face-to-face	Low in online	Softcopy preferred	✓
S23	Split	Low	Easy	More efficient in face-to-face	Sufficient	✓
S24	Split	Low	More in face-to-face	More efficient	Electronic preferred	✓

Several participants noted that the division between face-to-face and online lessons disrupted the flow, making it hard to integrate what was learned in each setting. According to Table 5, 15 students (62.5%) found the lesson flow to be fragmented (split), while nine students (37.5%) described it as smooth (good). One student described feeling lost when transitioning between the two, stating that “when I missed the face-

to-face lesson, I felt lost in the online session because they were meant to be a continuation of each other” (S22, 2024–2025). Another pointed out the difference in pacing, explaining that “lessons were covered slowly in face-to-face sessions, but the pace significantly increased in online lessons, making it difficult to adjust” (S4, 2021–2022). A few students, however, found the structure manageable and believed that consistent attendance was key to making the system work effectively. According to S6 (2021–2022):

The flow was good. As long as they attend both the face-to-face lessons and online lessons regularly, they can keep up. However, if they miss the face-to-face lessons at the beginning of the week, they will not benefit from the online sessions.

Most participants noted that teachers found it difficult to keep themselves motivated, owing to inattentive students, especially when it came to online classes in which cameras were not on, thus fostering an unengaged atmosphere. One student emphasized that

The ones who can increase the motivation of the teachers are students themselves. Unfortunately, my classmates did not attend regularly. They just turned off their cameras in online sessions, so there was no real classroom environment, which lowered the efficiency of the lessons and negatively affected teachers' motivation. (S21, 2024–2025)

Despite this, some students acknowledged teachers' efforts to adapt to the new system. S4 (2021–2022) commented that “the low attendance and lack of participation in lessons discouraged teachers, but this also pushed them to improve their methods. They had to adapt to new technologies and find ways to make lessons more engaging.”

Communication with instructors emerged as another challenge. A total of 17 students (70.8%) reported that they had better contact with teachers in face-to-face settings, while four students (16.7%) found online communication manageable (easy), and two students (8.3%) described their interaction with instructors as smooth. However, others found communication to be inadequate, especially during online sessions where instant feedback was not always available. One student expressed frustration over the lack of accessibility, stating that “in traditional schooling, it was easy to approach teachers whenever we had a question. Yet, in blended learning, communication was more challenging. In online lessons, asking a question was difficult, and responses lacked the immediacy of face-to-face feedback” (S22, 2024–2025).

The efficiency of feedback in blended learning was another concern. A total of 15 students (62.5%) stated that feedback was ineffective in online lessons, while five students (20.8%) found it more efficient in face-to-face settings. A smaller percentage (three students, 12.5%) noted that feedback was efficient for attentive students, and just one student (4.2%) rated it as satisfactory. As S3 (2021–2022) explained, “there were no contact problems for students who attended both face-to-face and online lessons regularly. However, those who were inconsistent in attendance had trouble keeping in touch with the instructor.”

The selection of materials used in blended learning was generally well received. Most students agreed that both printed and digital materials were effective, though preferences varied. Six students (25%) favored face-to-face materials, five students (20.8%) preferred hardcopy materials, while three students (12.5%) opted for softcopy versions. Another seven students (29.2%) considered the available materials sufficient,

and one (4.2%) found electronic materials preferable. Digital assignments were considered beneficial for reinforcing concepts covered in lessons. However, a few students pointed out occasional technical difficulties that interfered with accessing digital resources. Even so, despite some challenges, digital materials were generally perceived as beneficial for learning.

Apart from these variables affecting the efficiency of blended learning, there were other variables on which participants elaborated. Table 6 provides insight into the students' perspectives on the efficiency of blended learning overall.

Table 6

Additional Variables That Affect the Efficiency of Blended Learning

Category	Findings	Percentage (n)
Learning environment	No real learning environment	75% (18 students)
Evaluation criteria	Unfair/Susceptible to cheating	58.3% (14 students)
	Prefer face-to-face evaluation	41.7% (10 students)
Self-evaluation	Some improvement (speaking, writing skills)	83.3% (20 students)
	No significant improvement	16.7% (4 students)
Self-studying	More free time and materials for study	66.7% (16 students)
	Did not engage in additional studying	33.3% (8 students)
Learner autonomy	More responsibility for self-learning	70.8% (17 students)
	Found it stressful/difficult	29.2% (7 students)

As can be seen in Table 6, many participants raised concerns about the fairness of evaluation methods, with a majority believing that assessments were unreliable in an online setting. As one student explained, “in online exams, it is hard to know if everyone follows the rules, which makes it feel unfair” (S17, 2024–2025). Another added that “I found online evaluations unreliable, and I believe face-to-face exams should be mandatory” (S18, 2024–2025).

Regarding self-studying habits, most students reported having more time to study independently, while others admitted they were not motivated to engage in additional study. A student highlighted that “I had more time to study on my own, which helped reinforce what I learned in class” (S14, 2024–2025), whereas another remarked, “I did not feel motivated to study outside of class because I was not used to the format” (S5, 2021–2022). Finally, when discussing learner autonomy, most students reported taking on more responsibility for their learning, particularly in managing online sessions. Some students found this beneficial, while others described it as stressful.

The Efficiency of Skills Improvement Through Blended Learning

To answer our third research question, we asked which specific skills students felt they developed with the blended learning applied in their school. The results are shown in Table 7.

Table 7

Skills Improved with Blended Learning

Skill category	2021–2022 percentage (n)	2024–2025 percentage (n)
Speaking	75% (9 students)	66.7% (8 students)
Writing	50% (6 students)	50% (6 students)
Listening	50% (6 students)	41.7% (5 students)
Grammar	25% (3 students)	25% (3 students)
Reading	0% (0 students)	8.3% (1 student)

The comparative analysis of the two cohorts indicated a slight decline in the percentage of students reporting improvement in speaking skills, from 75% in 2021–2022 to 66.7% in 2024–2025. Writing skills, on the other hand, remained consistent, with 50% of students in both cohorts reporting improvement. Listening skills saw a decrease from 50% to 41.7%, while grammar skills remained stable at 25%. A slight increase in reading skills was noted, with 8.3% of students in 2024–2025 mentioning improvement; no students from the earlier cohort highlighted reading as an area of progress.

One student reflected on their development, stating that “I noticed an improvement in my speaking and writing, but I struggled more with grammar compared to traditional learning” (S15, 2024–2025). Another student pointed out that “listening was harder to improve because online sessions made it difficult to stay focused” (S18, 2024–2025). This suggested that while blended learning supports speaking and writing development, challenges remain in maintaining engagement for skills like listening and grammar.

Table 8

Students' Preferences for Developing Various Skills

Skill	Face-to-Face (2021–2022)	Face-to-Face (2024–2025)	Online (2021–2022)	Online (2024–2025)
Grammar	50% (6 students)	75% (9 students)	50% (6 students)	25% (3 students)
Vocabulary	42% (5 students)	58% (7 students)	58% (7 students)	42% (5 students)
Reading	58% (7 students)	50% (6 students)	42% (5 students)	50% (6 students)
Writing	75% (9 students)	83% (10 students)	25% (3 students)	17% (2 students)
Listening	42% (5 students)	33% (4 students)	58% (7 students)	67% (8 students)
Speaking	83% (10 students)	92% (11 students)	17% (2 students)	8% (1 student)

The results from Table 8 suggested that students increasingly favored face-to-face instruction for grammar (75%), writing (83%), and speaking (92%) in 2024–2025 compared to previous years. Grammar instruction, which was equally divided between online and face-to-face preferences in the earlier cohort, was seen in 2024–2025 as more effectively taught in a physical classroom. One student explained that “we can only ask grammar questions immediately in face-to-face lessons, but in online lessons, we cannot interrupt the teacher” (S7, 2021–2022).

Vocabulary instruction preferences increased, with 58% preferring face-to-face instruction in 2024–2025, an increase from 42% in the earlier cohort. As one student noted, “to learn vocabulary is within the responsibility of the students, so it does not matter whether it is online or face-to-face. We can have it online with applications that make it even easier” (S10, 2021–2022).

Reading instruction preferences remained relatively stable, with a small decline in face-to-face preference. One participant admitted that “I did not read the texts assigned in online lessons, while I felt responsible in face-to-face lessons” (S9, 2021–2022). However, others argued that online reading allowed them to concentrate better at home without distractions.

Listening instruction saw a shift toward online learning, rising from 58% to 67%. A student explained that “listening is better at home with headphones and no distractions” (S13, 2024–2025), whereas another claimed, “I concentrate better in class; listening depends on focus” (S6, 2021–2022).

Writing and speaking remained the most favored skills for face-to-face instruction in both cohorts, with 83% and 92% in the first and second cohorts respectively. One student highlighted the benefit of instant feedback in writing, stating that “it is much more beneficial when we get immediate feedback from the teacher” (S17, 2024–2025). For speaking, nearly all students agreed that real-time interaction was essential. “Speaking is more beneficial in face-to-face lessons because of the instant feedback from the teacher. There is a huge difference between talking to a screen and a real person” (S1, 2021–2022). Another emphasized the importance of body language, adding that “gestures and mimics are essential in speaking and can only be noticed in face-to-face conversations” (S14, 2024–2025).

Challenges and Advantages of Engagement and Interaction

To address the fourth research question, we conducted systematic classroom observations with checklists to examine systematic engagement and interaction in face-to-face and online classrooms throughout a complete semester. These observations focused on participation frequency, digital tool usage, peer and instructor interaction, and perceived engagement, revealing both challenges and advantages of blended learning.

Inconsistent class attendance during live sessions was among the most important challenges we observed. In online settings, even where students participated in sessions, engagement was still limited because they kept cameras turned off and made very minimal contributions in terms of conversation. Interaction rates among students and instructors were lower in online courses compared to classroom learning and produced a weaker classroom community.

On the other hand, certain advantages of blended learning were observed. The flexibility of online components allowed students to access course materials at their convenience, reducing logistical constraints. Some students also felt more comfortable participating in discussions through digital platforms rather than speaking in traditional classroom settings. Additionally, the availability of asynchronous resources enabled students to revisit content, potentially supporting comprehension and retention.

Discussion and Conclusion

This cross-sectional study aimed to evaluate the efficiency of blended learning by analyzing two different student cohorts from 2021–2022 and 2024–2025. While several studies report rather positive effects of blended learning on engagement and achievement, our findings indicated cohort- and skill-dependent variations: as students' digital literacy increased, engagement and acceptance shifted in patterned ways. This may have stemmed from institutional policies and infrastructural factors affecting online delivery, yet, it highlighted the role of program design and contextual adaptation in explaining the mixed results observed.

Our findings indicated persistent dissatisfaction with blended learning, although there was evidence of a gradual increase in acceptance over time. It was observed that the more linguistically capable students in the 2024–2025 cohort were better able to adjust to the system and benefit from its flexibility and potential for self-paced learning. However, some students struggled to keep up, particularly in online settings where they had fewer direct interactions with teaching staff. These findings aligned with other research suggesting that blended learning is more suited to higher-level learners (Doncheva, 2015; Gascoigne & Parnell, 2013).

The most common advantage of blended learning was the time gained and cost reduced through online learning. The downside, though, was the lack of responsibility and organization in online classes, which made students more easily distracted. This issue was also present in the 2024–2025 group, though students indicated they had stronger self-regulation than in the previous cohort (Hubackova & Semradova, 2016).

The second issue raised by participants was the split course progression with both face-to-face and online classes. Participants who did not attend face-to-face classes were unable to integrate knowledge in both modes. It seems that the 2024–2025 cohort managed better using an appropriate coping strategy, but learning was not consistent. According to previous studies, student participation and program design significantly influenced the effectiveness of blended learning (Hrastinski, 2019).

Teacher engagement and enthusiasm influenced students' attitudes towards blended learning. Students in both cohorts frequently reported that the teachers were less engaged during online classes due to decreased learner participation. The 2024–2025 cohort observed more use of digital technology and interactive approaches by teachers, which can be explained by learning and adaptation at the institutional level. Nevertheless, student interaction remained low in online sessions, leading to an ongoing issue of decreased instructor enthusiasm. These findings emphasized the importance of establishing interactive and participatory strategies in online learning environments (Akkoyunlu & Soylu, 2006). From a constructivist perspective, little interaction in online sessions undermines the collaborative processes essential for language development. To address engagement issues, students recommended several strategies, surprisingly including requiring cameras to be on (while they themselves preferred to leave them turned

off) during online lessons, integrating interactive activities, and using structured participation incentives. These insights aligned with previous research that advocated for more structured and engaging blended learning implementations (Akkoyunlu & Soylu, 2006).

The validity of assessment was the primary concern for both cohorts. Students complained that online testing was uneven and susceptible to cheating. The 2024–2025 cohort, being more familiar with LMS, was more aware of these issues and called for stricter regulations to ensure fairness. Mistrust of online testing was in alignment with previous results warning against vulnerabilities of cheating in remote testing contexts (Ja'ashan, 2015). This highlighted the importance of aligning assessment design with integrity-focused open and distributed learning (ODL) frameworks to ensure validity.

Despite all, the 2024–2025 cohort expressed a greater appreciation for the autonomy and flexibility of blended learning. They indicated that they felt more responsible for their own learning, particularly in terms of time management and self-study habits. This was supported by research suggesting that blended learning can foster learner autonomy if appropriately scaffolded (Lalima & Dangwal, 2017), and lead to better individualized learning opportunities (Castro, 2019). However, some students were still intimidated by the extra burden and indicated that additional guidance and support mechanisms were needed. Thus, it was clear that more guidance and instructions be provided by teachers for learners to better keep track of what they should do in their own studies. In other words, they should be given the chance to create their own products under the guidance of the teacher, considering constructivist theory. Li (2016) confirmed that a blended learning environment guided by the theory of constructivism had a positive effect on learners' growth. Similarly, Machumu and Zhu (2017) asserted that there was a significant positive correlation between students' conceptions of constructivist learning and their engagement in blended learning environments.

In terms of skill development, the patterns revealed in this study aligned with CLT and constructivist principles, emphasizing real-time interaction for speaking and writing. In that sense, the findings supported constructivist perspectives on active meaning-making while refining CLT applications by showing how digital literacy maturity mitigates extraneous load over time.

Finally, the study focused on the cohort-maturity effect to explain temporal shifts in blended learning perceptions, contributing a novel lens to EFL blended learning research. The quality of blended learning depends on structured interaction, intensified interaction, and fair evaluation; as students get used to it, institutions must change their methods of implementation to marry flexibility and quality.

Our results highlighted the need for more complex blended learning programs that take into account the various requirements of various language proficiency levels. They specifically fit with common trends seen in blended EFL and ODL contexts and resonated with global trends in ODL, where similar challenges of engagement, modality integration, and assessment integrity have been reported. Building on this alignment, we offer three generalizable design principles to direct future blended programs: (a) the use of modality with skill so that each mode's strengths (online or face-to-face) match up with the skill in question; (b) the development of structured interactional routines that encourage activity and discourage passive online activity; and (c) the creation of integrity-based assessments that respect academic integrity but also

support variable learning spaces, which will contribute to refining CLT and constructivist applications in blended EFL contexts.

Theoretical Implications

Beyond its practical implications, the study also advanced theoretical understanding of blended learning. By integrating CLT and constructivist perspectives to explain how digital literacy maturity influenced students' acceptance of blended learning over time, the study showed that as learners' digital competence increased, cognitive load was reduced while opportunities for active meaning-making expanded. By taking this double-theoretical approach, deeper understandings of relationships and interaction patterns between skill and modality can be obtained, thus extending previous research on blended EFL contexts.

Limitations

This study was conducted in a single institution with a small sample size, which limited the generalizability of our findings. Although including two independent cohorts provided a comparative viewpoint, the study was limited to the preparatory English classes at a Turkish state institution. Future studies could build on these findings by incorporating a more varied participant pool from different institutions and educational environments. Furthermore, this study focused solely on student perceptions, with no input from teachers. Given that teacher engagement and instructional strategies are critical to blended learning effectiveness, future research ought to examine the experiences of teachers to gain a more comprehensive picture of blended learning dynamics.

References

- Akkoyunlu, B., & Soylu-Yılmaz, M. (2006). A study on students' views on blended learning environment. *Turkish Online Journal of Distance Education*, 7(3), 43–56.
- Alias, A. A., Manan, N. A. A., Yusof, J., & Pandian, A. (2012). The use of Facebook as language learning strategy (LLS) training tool on college students' LLS use and academic writing performance. *Procedia - Social and Behavioral Sciences*, 67, 36–48.
<https://doi.org/10.1016/j.sbspro.2012.11.305>
- Bernard, R. M., Borokhovski, E., Schmid, R. F., Tamim, R. M., & Abrami, P. C. (2014). A meta-analysis of blended learning and technology use in higher education: From the general to the applied. *Journal of Computing in Higher Education*, 26(1), 87–122. <https://doi.org/10.1007/s12528-013-9077-3>
- Bonk, C. J., & Graham, C. R. (2006). *The handbook of blended learning*. Pfeiffer.
- Brodersen, R. M., & Melluzzo, D. (2017). Summary of research on online and blended learning programs that offer differentiated learning options. *Regional Educational Laboratory Central*, 1–22. Retrieved from <https://files.eric.ed.gov/fulltext/ED572935.pdf>
- Castro, R. (2019). Blended learning in higher education: Trends and capabilities. *Education and Information Technologies*, 24, 2523–2546. <https://doi.org/10.1007/s10639-019-09886-3>
- Dixon, T., Christison, M., Dixon, D. H., & Palmer, A. S. (2021). A meta-analysis of hybrid language instruction and call for future research. *The Modern Language Journal*, 105(4), 792–809. <https://doi.org/10.1111/modl.12732>
- Doncheva, L. (2015). An implementation of a hybrid form of education in foreign language acquisition. *Activities in Physical Education & Sport*, 5(2), 244–247. Retrieved from <https://fsprm.mk/wp-content/uploads/2022/03/Pages-from-APES-2-2015-28.pdf>
- Faridi, A., Saleh, M., & Fitriati, S. W. (2020). The effect of hybrid task-based language teaching and critical thinking on writing performance in Indonesia. *The New Educational Review*, 61(3), 109–118. <https://doi.org/10.15804/tner.2020.61.3.09>
- Gascoigne, C., & Parnell, J. (2013). Hybrid language instruction: A case study in beginning French. *The French Review*, 87(2), 141–154. <https://doi.org/10.1353/tfr.2013.0269>
- Hirata, Y., & Hirata, Y. (2008). Japanese students' attitudes towards hybrid learning. In J. Fong, R. Kwan & F. L. Wang (Eds.), *Hybrid learning and education: Lecture notes in computer science* (Vol. 5169). International Conference on Hybrid Learning, Springer.
- Hrastinski, S. (2019). What do we mean by blended learning? *TechTrends*, 63, 564–659. <https://doi.org/10.1007/s11528-019-00375-5>

- Hubackova, S., & Semradova, I. (2016). Evaluation of blended learning. *Procedia - Social and Behavioral Sciences*, 217, 551–557. <https://doi.org/10.1016/j.sbspro.2016.02.044>
- Hung, H. T. (2015). Flipping the classroom for English language learners to foster active learning, *Computer Assisted Language Learning*, 28(1), 81–96
<https://doi.org/10.1080/09588221.2014.967701>
- Ja'ashan, M. M. N. H. (2015). Perceptions and attitudes towards blended learning for English courses: A case study of students at University of Bisha. *English Language Teaching*, 8(9), 40–50.
<http://dx.doi.org/10.5539/elt.v8n9p40>
- Lalima, & Dangwal, K. L. (2017). Blended learning: an innovative approach. *Universal Journal of Educational Research*, 5(1), 129–136. <https://doi.org/10.13189/ujer.2017.050116>
- Li, X. (2016). A blended learning model of English teaching methodology course guided by constructivism. *International Journal of Continuing Engineering Education and Life-Long Learning*, 27(1–2), 101–110.
- Machumu, H. & Zhu, C. (2017). The relationship between student conceptions of constructivist learning and their engagement in constructivist based blended learning environments. *International Journal of Learning Technology*, 12(3), 253–272. <https://doi.org/10.1504/IJLT.2017.088408>
- Mayer, R. E. (2002). Multimedia learning. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 41, pp. 85–139). Academic Press. [https://doi.org/10.1016/S0079-7421\(02\)80005-6](https://doi.org/10.1016/S0079-7421(02)80005-6)
- Neumeier, P. (2005). A closer look at blended learning: Parameters for designing a blended learning environment for language teaching and learning. *ReCALL*, 17(2), 163–178.
<https://doi.org/10.1017/S0958344005000224>
- Redfern, S. & Naughton, N. (2002). Collaborative virtual environments to support communication and community in internet-based distance education. *Journal of Information Technology Education*, 1(3), 210–220. <https://doi.org/10.28945/356>
- Suwantharathip, O. (2019). Predictors of students' satisfaction with a hybrid English course. *Turkish Online Journal of Distance Education*, 20(1), 115–130. <https://doi.org/10.17718/tojde.522427>
- Sweller, J., Ayres, P., & Kalyuga, S. (2019). *Cognitive load theory* (2nd ed.). Springer.



August – 2026

From Consumers to Creators: How Student Participation in OER Development Transforms Understanding of Open Licensing and Improves Digital Literacy

Natasha Ramroop Singh, Ayisha Najeeha C O K, Olivia Mendez Romero, and Kira Reiersen
Thompson Rivers University, Kamloops, BC, Canada

Abstract

Open educational resources (OER) have emerged as key tools in democratizing education, and the levels of student awareness of OER, as well as engagement and use in many Canadian institutions, are increasing. This study explores whether involving undergraduates in the creation of OER improves their understanding of open licensing, their perceptions of OER utility, and their ability to independently access such resources. A pre-/post-activity study was conducted with 57 undergraduate students enrolled in an advanced biochemistry course. As part of the course's activities, students collaboratively contributed to the creation of a new open resource aiming to increase understanding of enzyme mechanisms, to be used by themselves as well as their peers. Data were collected using structured surveys and analyzed via the Wilcoxon signed-rank test across three domains: Creative Commons knowledge, OER perception, and OER access. Statistically significant improvements were observed in students' understanding of open licensing, particularly their knowledge of Creative Commons licenses (e.g., understanding license types; $p < .001$), perceptions of OER credibility (e.g., perceived accuracy and source citation; $p = .0049$), and behaviors related to accessing diverse types of OER materials ($p = .0397$). Thematic analysis supported these findings, highlighting learning gains and increased engagement in areas previously unfamiliar to students. Participation in the creation of OER fostered a deeper conceptual understanding and more favorable attitudes toward open education. However, changes in access behavior were modest, indicating the need for institutional scaffolding to fully support learner autonomy in open environments.

Keywords: open educational resources, OER, student co-creation, Creative Commons, digital literacy

Introduction

Open educational resources (OER) have emerged as a powerful force reshaping how knowledge is produced, shared, and accessed across educational institutions. As higher education grapples with rising textbook costs and persistent inequities in learning access (Mullens & Hoffman, 2023), OER offers a sustainable and inclusive solution. Defined by UNESCO (n.d., para. 1) as “learning, teaching and research materials in any format and medium that reside in the public domain or are under copyright that have been released under an open license, that permit no-cost access, re-use, re-purpose, adaptation and redistribution by others,” OER remove legal and financial barriers while supporting more collaborative and flexible forms of learning. At the heart of the OER movement lies a broader commitment to open education, which emphasizes access, equity, participation, and learner-centered pedagogy. Using open licenses, most commonly Creative Commons licenses, educators and learners can not only freely access materials but also adapt and redistribute them, making education more responsive to diverse learning needs (Adil et al., 2024; Mishra, 2025; Santos-Hermosa, 2024). Recent international guidance, such as the Dubai Declaration on OER, foregrounds the role of open-licensed content and emerging technologies in promoting equitable and inclusive access to knowledge, underscoring the global relevance of efforts to build and share collectively authored OER (UNESCO, 2024).

Student participation in the creation of OER appears to positively influence student engagement and perceptions of course content. When students actively create materials, such as essays, videos, or remixed texts, they transition from passive recipients to active contributors in their learning ecosystem (Trust et al., 2023). This process, aligning with the principles of OER-enabled pedagogy, transforms disposable assignments into renewable artifacts that provide ongoing value for the creator and for future learners (Tietjen & Asino, 2021; Werth & Williams, 2021a; Wiley & Hilton, 2018).

Initially, the adoption of OER in higher education was driven largely by a desire to lower costs for students (Wiley & Hilton, 2018). However, recent scholarship has revealed their broader pedagogical potential. Building on this, educators are moving beyond simply using open materials to embracing open pedagogy, a practice that actively involves students in the creation, adaptation, and dissemination of learning resources (Cronin, 2017; DeRosa & Robison, 2017). This participatory approach encourages students to shift from passive recipients of knowledge to active contributors in the educational process (Clinton-Lisell et al., 2023; Short et al., 2024).

Engaging students in the creation of OER has shown significant educational benefits. Research suggests that this involvement promotes the student’s understanding of the subject, improves learning outcomes, and fosters a stronger sense of ownership over their learning (Cho & Permzadian, 2024; Lubicz-Nawrocka & Bovill, 2023; Spencer et al., 2025). A study by Grimaldi et al. (2019) found that high school students who codeveloped interactive OER demonstrated gains in academic performance. Similarly, in higher education, students who adapted and integrated OER into their coursework reported a better grasp of open practices and digital literacy skills (Farias-Gaytán et al., 2023; López-Núñez et al., 2024; Nobre, 2021; Tinmaz et al., 2022).

These hands-on experiences promote deeper cognitive engagement, as students are required not only to comprehend the content but also to critically assess, reframe, and communicate it for peer and public use.

This process also encourages reflection on issues such as licensing, audience diversity, and accessibility—key aspects of openness often overlooked in traditional educational contexts (Schmied et al., 2024; Werth & Williams, 2021c).

Both students and faculty generally express positive perceptions of OER, citing benefits such as cost savings, increased accessibility, and flexibility in learning (Hilton, 2020). When faculty engage in creating or adapting OER, their enhanced understanding of open educational practices often translates into more meaningful student learning experiences (Meger et al., 2021). Students involved in such practices report greater engagement and motivation and a sense of empowerment in their learning journeys (LeMire et al., 2024; Werth & Williams, 2021c).

Moreover, OER contributes directly to educational equity by making high-quality resources freely available to all learners, regardless of economic background. This is especially critical for students who may otherwise be unable to afford textbooks or other required materials (Hilton, 2020). When students participate in the creation of these resources, they not only benefit from free access but also gain an appreciation for inclusive educational design (Griffiths et al., 2022; Jhangiani et al., 2025; Meger et al., 2021).

Despite the promise of OER and open pedagogy, many students remain unaware of what OER are or how they function. Key concepts such as reuse rights, attribution, and open licensing remain unfamiliar to most learners. Additionally, students are rarely invited to participate in conversations about the creation, purpose, or future use of educational materials. This lack of exposure limits their ability to critically engage with the resources they use and undermines the transformative potential of open education (Adil et al., 2024; Cheung et al., 2022).

While research has extensively explored the cost-effectiveness of OER and faculty perspectives (Clinton, 2018; Clinton & Khan, 2019; Hilton, 2020), fewer studies have examined how students' direct involvement in OER creation affects their understanding, perceptions, and ability to navigate open resources independently. As Nobre (2021) and Meger et al. (2021) have pointed out, meaningful student engagement in OER creation requires supportive institutional policies, faculty commitment, and training. However, these elements are still developing in many contexts (Arispe & Hoyer, 2023; Maultsaid & Harrison, 2023; Mercer-Mapstone et al., 2017; Omland et al., 2025).

This study seeks to address this gap by exploring how undergraduate students' participation in creating OER influences their understanding of open licensing, their perception of OER utility, and their ability to access and reuse such materials independently. The activity was implemented as part of a project in a third-year undergraduate biochemistry course, where students collaborated to develop openly licensed content. The full resource developed by students can be found at <https://enzymemechanisms.pressbooks.tru.ca/>. Pre- and post-activity surveys were used to measure changes in students' knowledge, attitudes, and behaviors related to OER (Lan et al., 2025; Tili et al., 2023; Tili et al., 2025).

The central research question guiding this study is: Does participation in the creation of open educational resources improve students' understanding, perception, and access to OER-related concepts?

By examining the effects of open pedagogical engagement using both statistical and thematic analyses, this

study contributes new insights into the role of students as active agents in open education ecosystems. The findings aim to inform future efforts to integrate student-led OER initiatives into higher education curricula.

Research Methodology

Research Design

This study used a quantitative pre- and post-activity within-subjects design to evaluate changes in students' understanding, perceptions, and access to OER due to participating in creating a course-based OER. The activity was embedded within an undergraduate biochemistry course and structured as a collaborative, hands-on module in which students developed openly licensed educational content. Students' videos were evaluated using a rubric that emphasized scientific accuracy, clarity of communication, proper citation and attribution practices, and correct application of Creative Commons licensing, ensuring that only work meeting these quality standards was published. To capture changes in key constructs, data were collected at two time points: prior to the activity (pre-survey) and immediately after its completion (post-survey).

Institutional Context and Participants

As part of the university's commitment to advancing open education to reduce barriers to learning and enhance student success, the institution supports zero textbook cost approaches that replace commercial textbooks with OER and/or library-licensed materials, which lowers costs while keeping high-quality learning materials in students' hands. Instructors are encouraged to support these efforts, and as such, research on the efficacy and success of these OER-enabled and open pedagogy activities are required for building a sustainable model for faculty and students alike.

The 57 participants in this study were undergraduate students enrolled in an advanced biochemistry course (BIOL 3130—Introduction to Biochemistry) at a Canadian postsecondary institution. The course incorporated an open pedagogy model, emphasizing learner participation in the co-creation of educational resources. All students were invited to voluntarily participate in the study, and informed consent was obtained in accordance with the institution's ethical research guidelines. Survey participation was anonymous, and individual responses were matched across the two survey phases using non-identifiable coded keys to enable paired statistical analysis.

Survey Instrument

Data were collected using a structured survey instrument specifically designed to assess student knowledge, perceptions, and experiences with OER. The instrument consisted of 17 primary questions, many of which included multiple subitems, resulting in a total of over 30 analyzable variables. The survey elements were aligned with key thematic domains:

- Access to and familiarity with OER platforms and materials
- Perceptions of OER quality, usefulness, and educational value
- Experience with and attitudes toward contributing to OER
- Knowledge of Creative Commons licensing and its applications
- Confidence in the accuracy and credibility of OER

Most items were measured using 5-point Likert-type scales (e.g., 1 = Strongly Disagree to 5 = Strongly Agree or 1 = Never to 5 = Always), while a few items employed matrix-style formats or matching exercises to assess categorical knowledge. Both survey instruments (pre- and post-) were structurally identical to ensure a valid direct comparison.

Data Collection and Processing

The surveys were administered in class as paper-based questionnaires. Participating students responded to them at two points: before and after the OER creation activity. Participants were encouraged to complete both surveys to enable matched-pair analysis. For the final analysis, only cases in which participants completed both surveys were considered. Survey responses were exported to a statistical software environment for cleaning and preprocessing. The variables were anonymized and matched based on unique identifiers. “No Answer” and “Not Applicable” responses were retained for descriptive completeness but excluded from inferential testing. A detailed mapping of survey items to thematic codes is provided in Table 1.

The de-identified dataset, codebook, and analysis scripts that support the findings of this study are openly available in GitHub at <https://github.com/nramroopsingh/IRRODL2025> and may be reused under license CC BY 4.0.

Table 1

Summary of Survey Questions and Themes

Question no.	Survey question/item	Theme
1a	Have you used OER in your academic studies?	Access to OER
1b(i)	Frequency of OER use vs. traditional resources	Access to OER
1b(ii)	Types of OER materials accessed (textbooks, videos, etc.)	OER types
1b(iii)	Frequency of use of: OER Commons, OpenStax, LibreTexts, etc.	Platform usage
2	Overall perception of OER	Perception
3	Importance of factors: quality, access, variety, relevance, peer review, editability	Perception drivers
4	Engagement with course content when using OER	Engagement
5	Importance of features: interactivity, multimedia, collaboration, contribution	Engagement
6	Past experiences contributing to or modifying OER	Participation
7	Belief that contribution affects OER perception	Perception

Question no.	Survey question/item	Theme
		change
8	Belief that contribution increases engagement	Engagement
		change
9	Concern about licensing when contributing	Licensing
		awareness
10	Understanding of CC licenses	CC knowledge
11	Frequency of using CC-licensed content	CC usage
12	Matching CC licenses with correct descriptions	CC license
		differentiation
13	Appropriateness of CC licenses in various contexts (educational, personal, commercial)	CC application
14	Confidence in the accuracy of OER	Trust in OER
15	Importance of author credibility, peer review, citation, updates	Trust factors
16	Frequency of verifying OER accuracy with external sources	Verification behavior
17a	Plan to take BIOL 3230 (Biochemistry)	Course planning
17b	Reason for taking BIOL 3230	Course motivation

Note. OER = open educational resources; CC = Creative Commons.

Ethics Approval and Funding

Ethics approval for this research project was obtained from the university's research ethics board. Participation was entirely voluntary, and students were informed that their decision to participate—or not—would have no impact on their course grades or academic standing. No personal identification was used in the analysis or reporting, and all data were stored securely.

Funding for this work was provided by the Centre for Excellence in Learning and Teaching's (CELT) Scholarship of Teaching and Learning (SoTL) Scholars Development Grant.

Handling Missing Data

We excluded questions with missing responses over 20% to ensure data quality and reliability. The missingness threshold was applied independently to both the pre-survey and post-survey datasets. Questions exceeding this

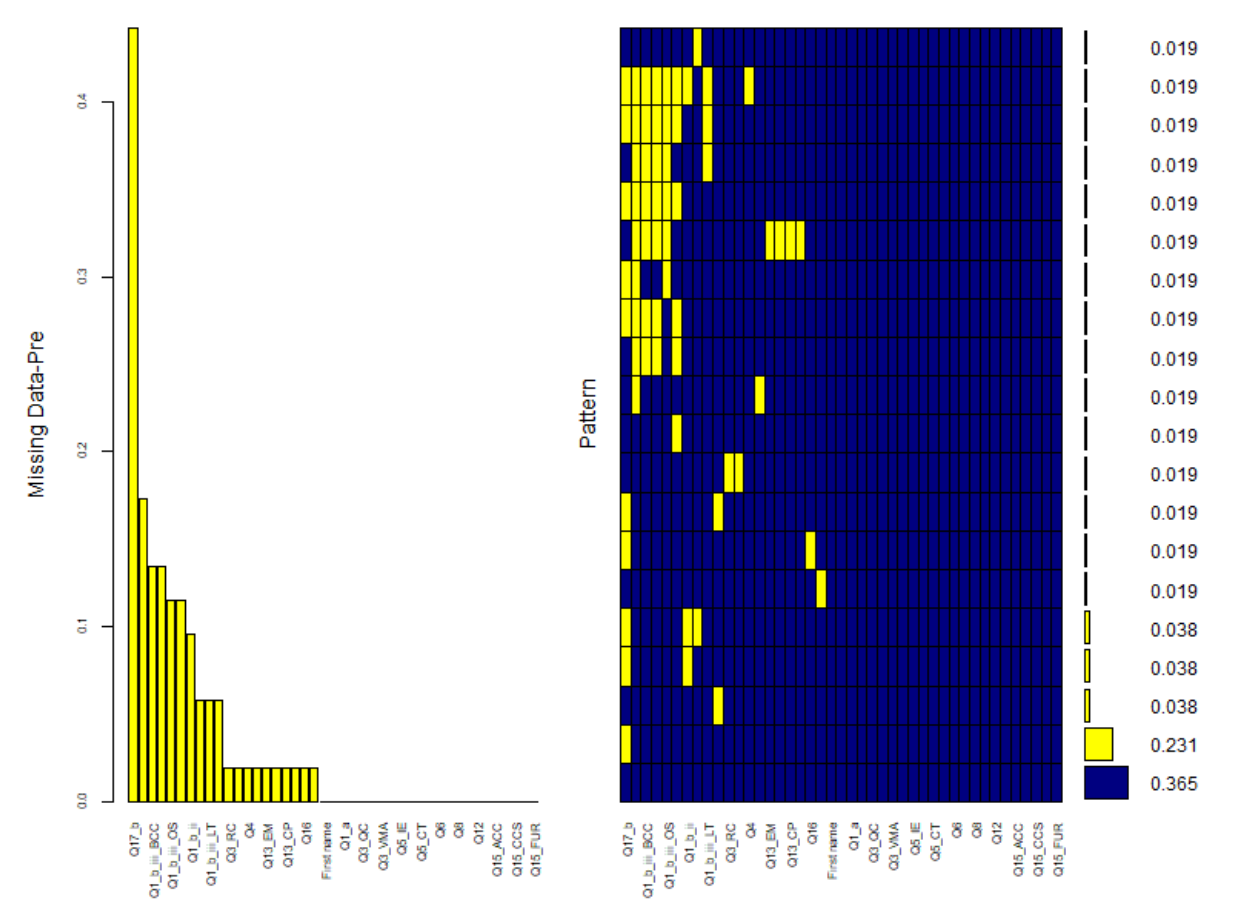
threshold in either dataset were removed from further analysis. A 20% threshold is widely accepted in survey-based research to minimize bias and reduce the risk of unreliable findings (Mirzaei et al., 2022). High levels of missing data can arise from unclear wording, lack of relevance, or survey fatigue, and retaining such questions may compromise the accuracy and validity of the results. Excluding these questions ensured the remaining data were sufficiently robust for meaningful statistical interpretation, as well as ensuring reliable conclusions.

Figures 1 and 2 show the distribution of missing data for pre-survey and post-survey items, respectively. Items with high missingness were mostly clustered in optional or multipart questions, suggesting possible survey fatigue or lack of relevance to certain respondents.

The final cleaned dataset allowed for direct, matched comparisons between the pre- and post-activity responses for each student. This paired structure was ideal for the application of the Wilcoxon signed-rank test, which does not require normality and is appropriate for ordinal Likert data.

Figure 1

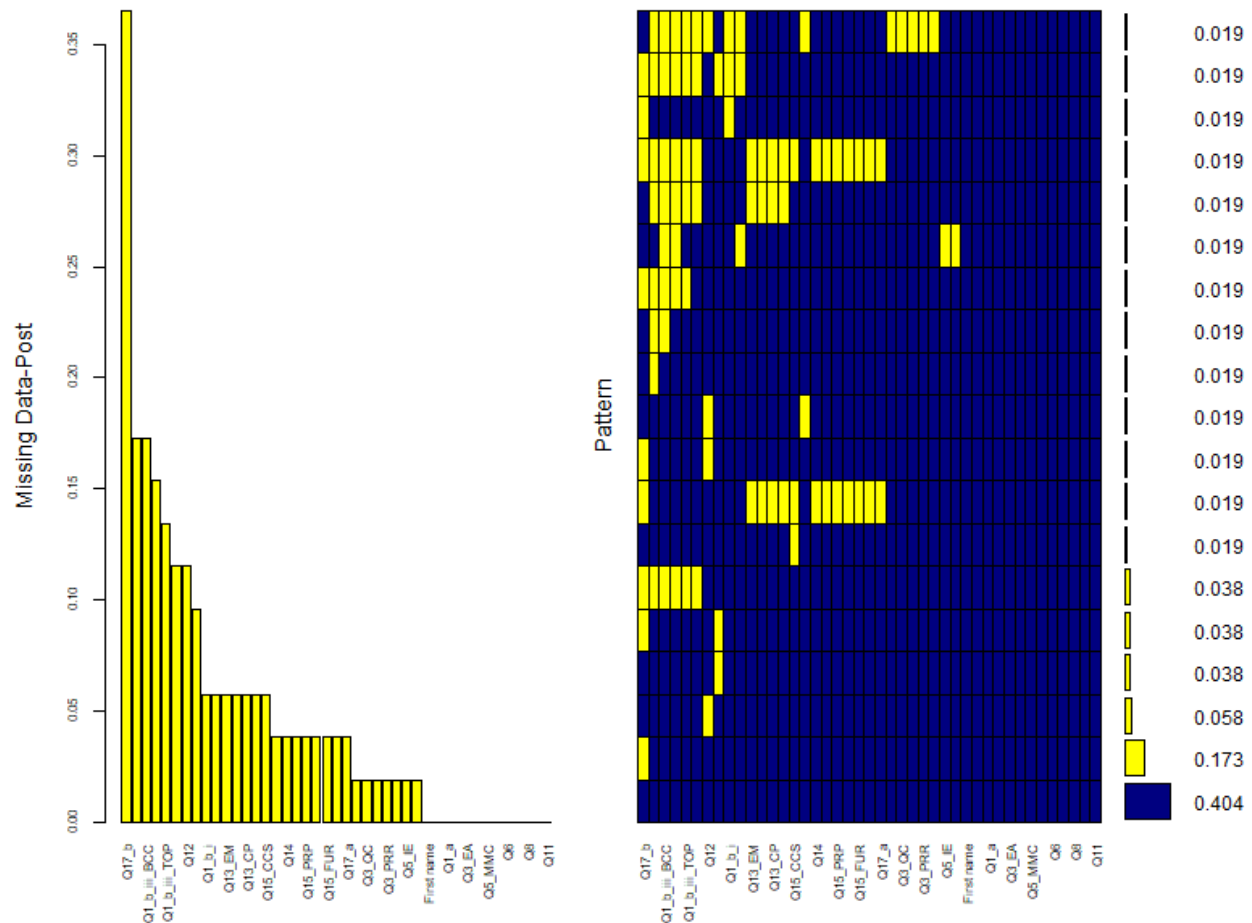
Proportion and Patterns of Missing Data in the Pre-Survey



Note. Variables with the highest levels of missing responses are highlighted.

Figure 2

Proportion and Patterns of Missing Data in the Post-Survey



Note. Like in the pre-survey, select variables showed elevated levels of nonresponse. Variables with the highest levels of missing responses are highlighted.

Statistical Analysis

Due to the ordinal nature of Likert scale data and potential violations of normality assumptions, a nonparametric approach was used. The Wilcoxon signed-rank test was applied to compare paired responses for each question before and after the activity.

Wilcoxon Signed-Rank Test

To assess changes in students' responses before and after the activity, the Wilcoxon signed-rank test was employed. This nonparametric statistical method is suitable for evaluating paired ordinal data, such as Likert scale responses, and does not assume a normal distribution of the underlying population.

The Wilcoxon signed-rank test is used to determine whether the median of differences between two related samples is significantly different from 0. In this study, it tested the null hypothesis that the educational activity

would have no effect on students' understanding, perceptions, or behaviors related to OER, as measured by matched pre-survey and post-survey responses.

The Wilcoxon signed-rank test is appropriate under the following conditions:

- Observations are paired and drawn from the same participants.
- The measurement scale is at least ordinal.
- The distribution of the differences is symmetric.

For each survey item, paired differences were calculated between post-survey and pre-survey responses:

$$D_i = y_i - x_i$$

where x_i and y_i represent the pre- and post-survey scores, respectively, for participant i . Pairs for which $D_i = 0$ were excluded. The absolute differences $|D_i|$ were ranked in ascending order, with tied ranks assigned their average.

Each rank was then given the sign of the original difference D_i . The test statistics W^+ and W^- were calculated as the sum of the positive and negative signed ranks, respectively:

$$W^+ = \sum_{\text{positive } D_i} \text{rank}(|D_i|), \quad W^- = \sum_{\text{negative } D_i} \text{rank}(|D_i|)$$

The final test statistic is given by

$$V = \min(W^+, W^-)$$

For sample sizes greater than 20, a normal approximation was applied. The standardized test statistic Z is calculated as:

$$Z = \frac{W^+ - \mu}{\sigma}$$

where the mean and standard deviation under the null hypothesis are

$$\mu = \frac{n(n+1)}{4}, \quad \sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

Here, n denotes the number of nonzero difference pairs.

The resulting Z score or rank-sum statistic was used to compute p values for each item. A result was considered statistically significant if $p < .05$, indicating that the activity produced a meaningful change in the students' responses. Items with $p \geq .05$ were interpreted as showing no statistically significant difference.

This test allowed for rigorous evaluation of learning and attitudinal shifts resulting from participation in the

OER creation project while accommodating the ordinal nature of the data and the absence of normality assumptions.

Thematic Analysis of Survey Responses

To better understand the effects of the OER activity, the survey responses were grouped into three key thematic domains: (a) knowledge of Creative Commons licenses, (b) perception of OER, and (c) access to OER. These themes reflect the study's core objective of evaluating shifts in students' understanding, perceptions, and behaviors related to OER. Each theme was assessed using a set of targeted survey items and visualized using box plots comparing pre-survey and post-survey responses. Statistical comparisons were conducted using the Wilcoxon signed-rank test to detect significant changes in response distributions across time points.

Analysis of Results

Descriptive Trends in Survey Responses

To explore the initial patterns in the changes in responses between pre- and post-activity surveys, we visualized the distribution of responses on the Likert scale before and after the activity (Figure 3). Several key themes emerged.

Licensing Knowledge

The questions assessing students' self-reported understanding of Creative Commons licenses showed a clear upward shift. This indicates that students gained familiarity with open licensing terminology and its applications after participating in the OER creation process.

Participation in OER Creation

The question about students' past experiences with creating or improving OER also demonstrated a noticeable increase in post-survey responses. This suggests that the activity succeeded in providing students with practical exposure to open resource development.

Platform Familiarity

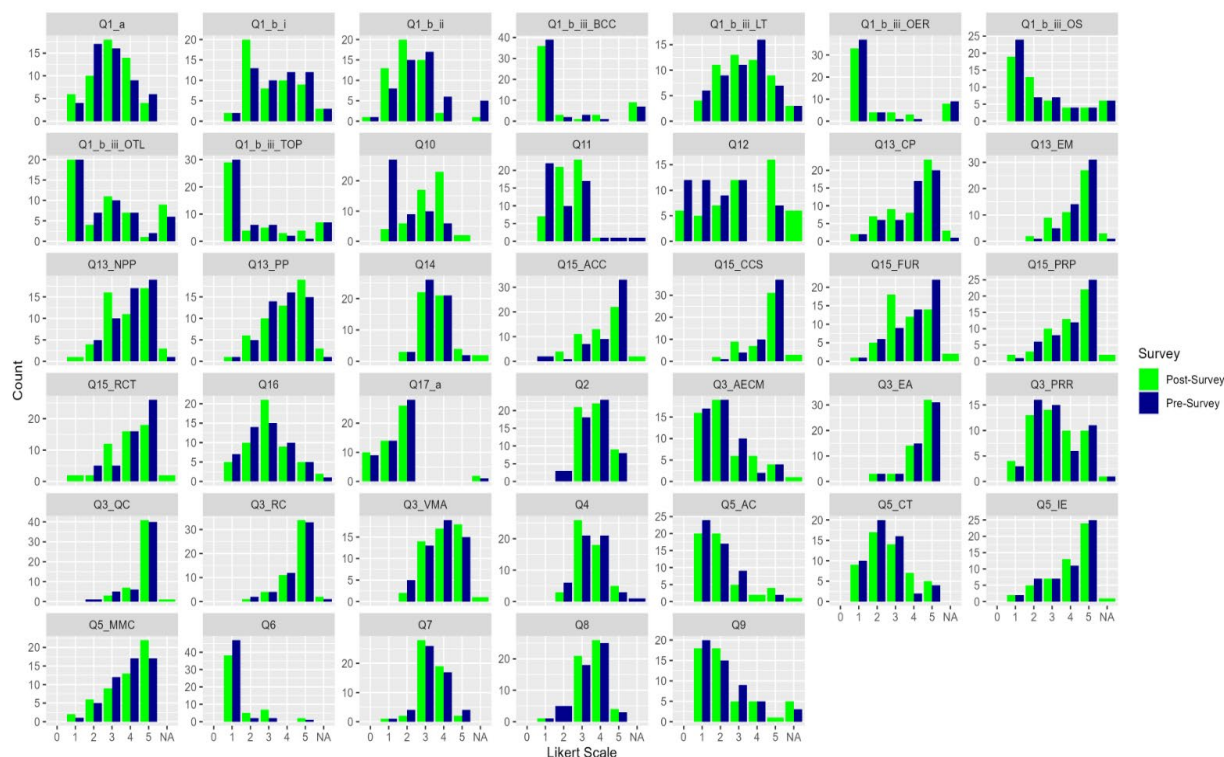
In contrast, little to no change was revealed for items addressing the frequency of use of specific OER platforms (such as BCCampus and OER Commons). This may reflect students' preexisting familiarity and comfort with these platforms, limiting the potential for growth in this domain.

Perception and Uncertainty

A high concentration of neutral responses was shown for several items in the pre-survey, particularly the perceived importance of content quality in shaping OER perception. This suggests that many students initially lacked strong opinions or familiarity with these quality indicators. However, the post-survey distribution showed movement away from the neutral midpoint, implying increased confidence or more informed perspectives.

Figure 3

Comparison of Pre-Survey and Post-Survey Responses Across Likert Scale Items



Note. Pre-survey responses are shown in blue, while post-survey green responses are shown in green.

Analysis Between the Pre-Survey and Post-Survey

The results of the Wilcoxon signed-rank test for the comparison of pre-survey and post-survey responses are presented in Table 2. The test was used to determine whether the students' knowledge of, perceptions of, and engagement with OER changed significantly after participating in the OER creation activity. Several survey items showed statistically significant improvements. For instance, students reported accessing a wider variety of OER materials after the activity (Q1b[ii]), $p = .0397$, and more students indicated experience contributing to or improving OER (Q6, $p = .0050$). The most significant change was observed in the students' understanding of Creative Commons licenses, which improved markedly after the activity (Q10, $p = 4.54 \times 10^{-7}$). Additional gains were seen in the use of Creative Commons–licensed content (Q11, $p = .0238$), correct identification of license types (Q12, $p = .0227$), and perceived importance of citing credible sources (Q15 CCS, $p = .0284$) and maintaining the OER content updated (Q15 FUR, $p = .0487$).

Although these items showed significant changes, most did not. Items such as previous OER usage (Q1a, $p = .6857$), ease of access (Q3 EA, $p = .7819$), and engagement with interactive features (Q5 IE, $p = .7422$) remained relatively stable. This suggests that some perceptions and habits were already well formed before the activity or were not strongly affected by it. For certain items, such as the belief that contributing to OER would change one's perception (Q7, $p = 1.000$), all responses remained the same, indicating no shift in opinion.

Table 2

Wilcoxon Signed-Rank Test Results for Pre-Survey and Post-Survey Responses

Question	<i>p</i>	Test statistic (V)
Q1a	.6857	274.5
Q1b(i)	.1025	291.5
Q1b(ii)	.0397	368
Q1b(iii) OER	.1371	7
Q1b(iii) OS	.5211	97
Q1b(iii) LT	.7380	265
Q1b(iii) OTL	.5538	118.5
Q1b(iii) BCC	.4568	6.5
Q1b(iii) TOP	.2478	52
Q2	.5298	118.5
Q3 QC	.1658	22
Q3 EA	.7819	79
Q3 YMA	.2436	166
Q3 RC	.4239	60
Q3 PRR	.6717	198
Q3 AEEM	.4774	115.5
Q4	.4636	105.5
Q5 IE	.7422	175.5
Q5 MMC	.6028	222
Q5 CT	.1144	135
Q5 AC	.2507	92
Q6	.0050	0
Q7	1.0000	95.5
Q8	.1902	72
Q9	.5692	130
Q10	4.54e-07	54
Q11	.0238	117
Q12	.0227	189
Q13 EM	.2496	203.5
Q13 PP	.5513	263.5
Q13 CP	.6213	167.5
Q13 NPP	.3378	197
Q14	.2632	125

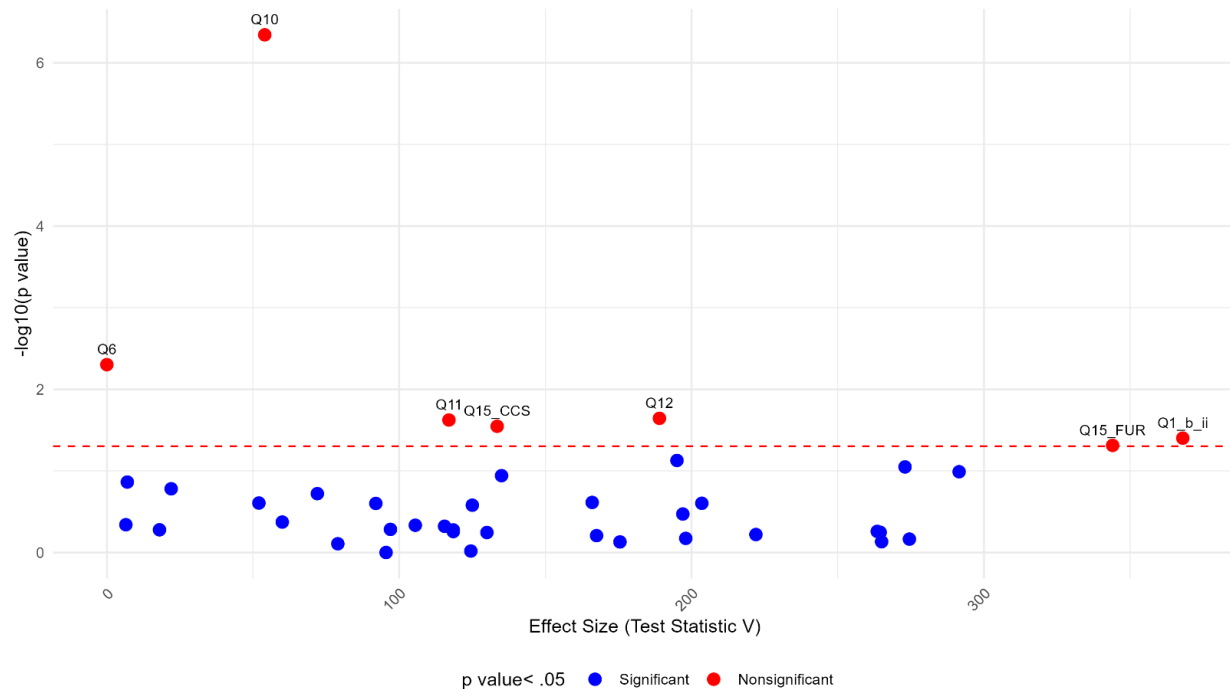
Question	<i>p</i>	Test statistic (V)
Q15 ACC	.0746	195
Q15 PRP	.9598	124.5
Q15 CCS	.0284	133.5
Q15 RCT	.0895	273
Q15 FUR	.0487	344
Q16	.5661	264.5
Q17 a	.5271	18

Note. OER = OER Commons; OS = OpenStax; LT = LibreTexts; OTL = Open Textbook Library; BCC = BCcampus; TOP = TRU Open Press; QC = Content Quality; EA = Ease of Access; VMA = Variety of Materials Available; RC = Relevance to Coursework; PRR = Peer Reviews or Recommendations; AECM = Ability to Edit or Change Material; IE = Interactive Elements; MMC = Multimedia Content; CT = Collaborative Tools; AC = Ability to Contribute or Modify Content; EM = Educational Materials; PP = Personal Projects; CP = Commercial Projects; NPP = Non-profit Projects; ACC = Author Credibility; PRP = Peer Reviews/Editing Process; CCS = Citing Credible Sources; RCT = Relevance to Coursework/Topic; FUR = Frequency of Updates/Refresh.; Statistically significant *p* values are **bolded**.

To help visually interpret the results, a volcano plot was created (Figure 4). This plot displays both the test statistic and the significance level for each item. Red markers highlight survey items where significant changes occurred, while blue markers indicate nonsignificant results. Overall, the findings suggest that the activity had the strongest impact on students' knowledge of licensing and content contribution practices, while perceptions of access, usage, and platform familiarity showed little change.

Figure 4

Volcano Plot of Wilcoxon Signed-Rank Test Results for Pre-Survey and Post-Survey Responses



Note. The x-axis represents the effect size (Wilcoxon test statistic V), while the y-axis represents the significance level as $-\log_{10}(p \text{ value})$. Red points indicate statistically significant questions ($p < .05$); blue points represent nonsignificant questions. Labels are added for significant questions to facilitate interpretation.

Thematic Results

The thematic summaries presented below reflect descriptive, domain-based interpretations aligned with the survey structure and are intended to contextualize statistically significant findings rather than represent an inductive or qualitative thematic analysis.

Students Demonstrated Significant Gains in Creative Commons Licensing Knowledge

Participation in the OER creation activity produced the largest and most consistent gains in students' understanding of Creative Commons (CC) licensing. This theme was assessed using responses to questions Q10, Q11, and Q12 (Figure 5). A clear improvement in the understanding of CC licensing by the students' post-activity was observed. Q10 showed the most dramatic change, with a highly significant increase in self-reported understanding of CC licenses ($p = 4.54 \times 10^{-7}$). Q11 showed a more modest but still statistically significant improvement in the frequency of using CC-licensed content ($p = .0238$). Q12 responses indicated an improved ability to correctly match the license types with descriptions ($p = .0227$). Box plots revealed not only upward shifts in median values but also reduced variance, particularly for Q10. The substantial pre-post shift on self-reported CC knowledge (Q10), coupled with improved accuracy in matching license types with their descriptions (Q12), suggests movement beyond familiarity toward functional understanding. Notably, the reduction in response variance for Q10 indicates convergence around a higher shared baseline of

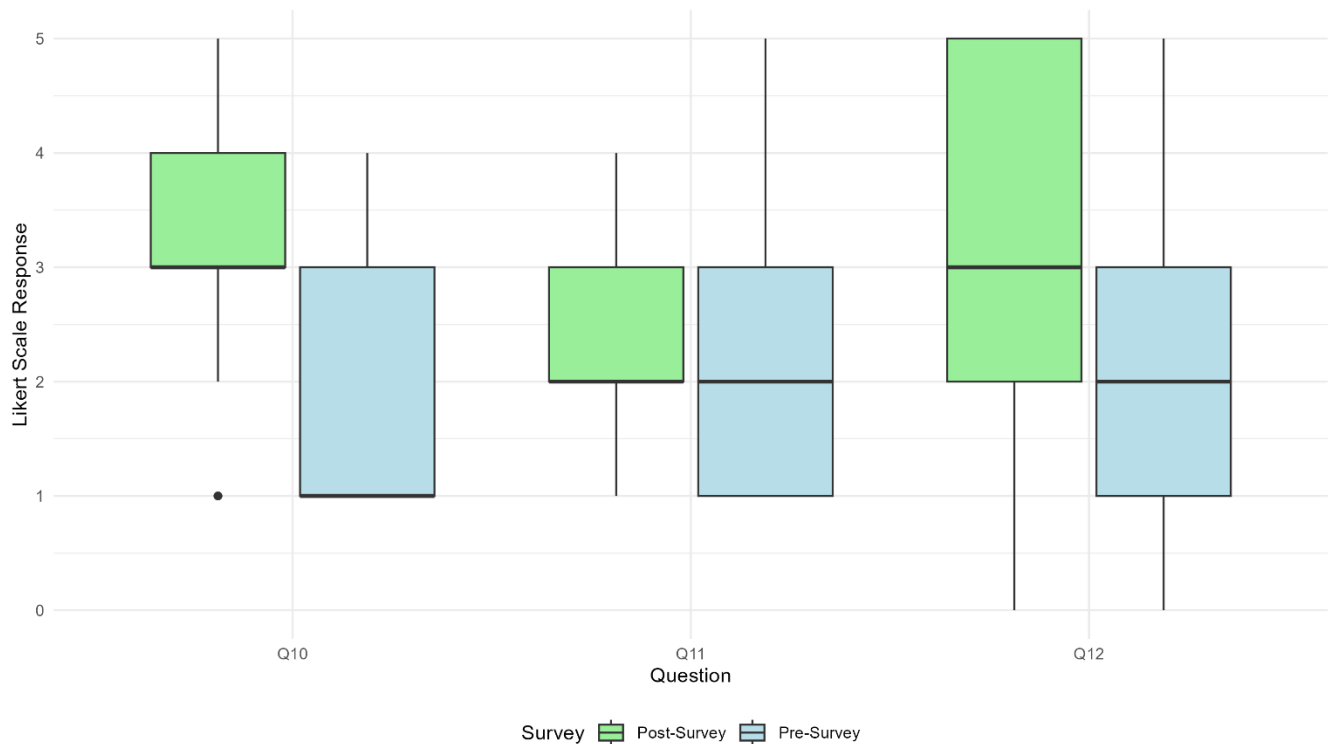
licensing literacy. This matters in group writing, as one person's misunderstanding can compromise the licensing compliance of the whole group.

Two mechanisms likely underpin these changes. The first is learning by doing: Making openly licensed artifacts requires students to select a license, justify that choice, and apply attribution standards to third-party assets. These activities are known to transform abstract policy into situated practice (e.g., OER-enabled pedagogy's focus on authentic, public-facing tasks; Wiley & Hilton, 2018). Second, social accountability, in the form of public dissemination, raises the stakes of correct licensing and citation, prompting attention to reuse conditions and versioning. The modest but significant increase in the reported frequency of using CC-licensed materials (Q11) is consistent with a near-term shift in information-seeking habits; however, it likely represents early adoption rather than wholesale behavior change. This pattern aligns with prior research showing that co-creation primarily catalyzes conceptual and procedural knowledge first, with behavioral routinization emerging later with continued opportunities and support (Tili et al., 2025; Trust et al., 2023).

Taken together, these results indicate that integrating explicit licensing decisions and attribution tasks into coursework is a high leverage feature of open pedagogy: It directly addresses a common barrier to reuse/remix and equips students with transferable digital civic competencies (Hilton, 2020; Werth & Williams, 2021a).

Figure 5

Comparison of Pre-Survey and Post-Survey Responses for the Theme “Knowledge of Creative Commons Licenses”



Note. Post-survey responses (green) show significant upward shifts, particularly in Q10 and Q12.

Students’ Perceptions of OER Improved in Targeted Areas Related to Credibility and Contribution

Students’ perceptions of OER also shifted positively, most clearly in their reported experience in contributing to or improving OER (Q6) and in the importance they placed on credible citation and currency (Q15 CCS; Q15 FUR; Figure 6). Students’ experience contributing to OER (Q6) showed a significant improvement ($p = .0050$), moving from a median response of 1 in the pre-survey to 2 in the post-survey. For Q15 CCS (citing credible sources) and Q15 FUR (importance of frequent updates), the responses were already high at baseline but showed small yet statistically significant gains ($p = .0284$ and $p = .0487$, respectively).

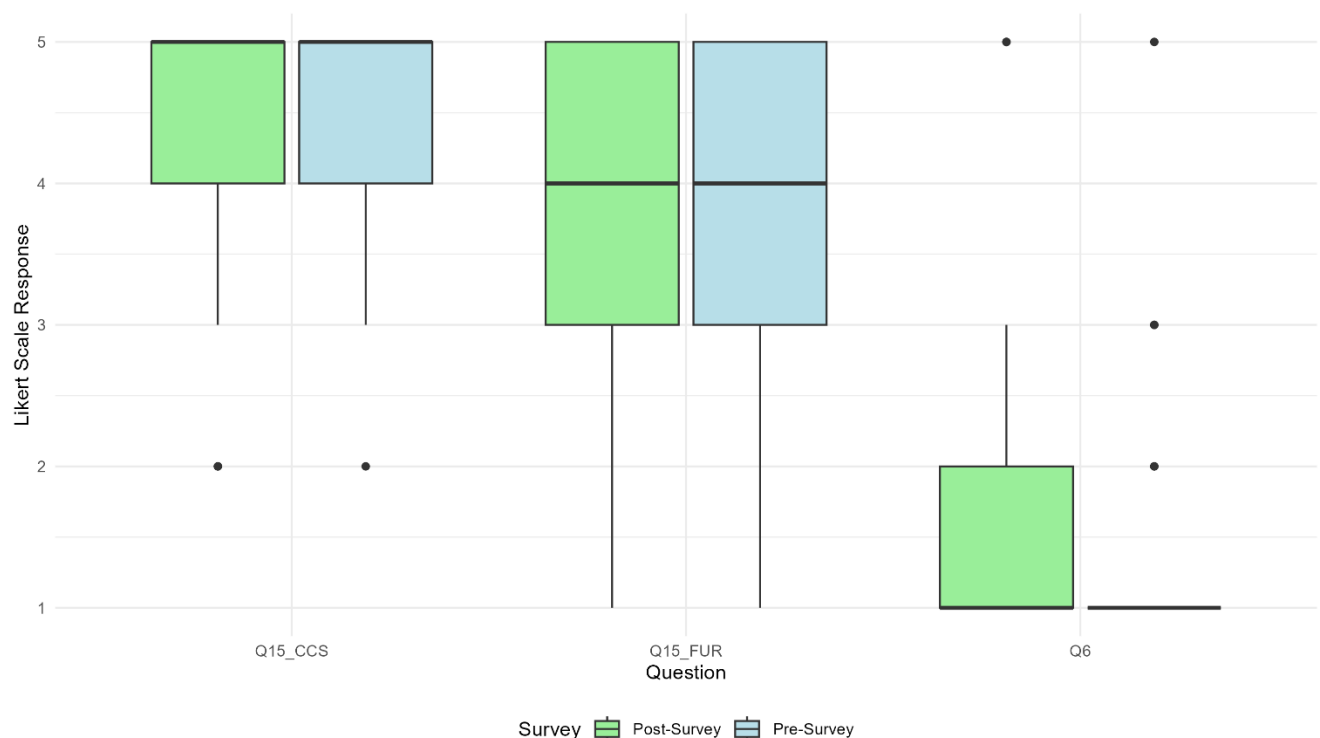
These findings suggest that the activity positively impacted on students’ trust of and perception of OER. The median increase for contribution experience was from “never” to “rarely.” Although this may appear small, it is substantively meaningful in a population that typically has had no prior authorship history. Moreover, the trust-related items were already high at baseline, suggesting ceiling effects; the observed gains, therefore, reflect reinforcement of an already favorable stance. When students have experience contributing and start paying closer attention to source credibility and how often materials are updated, they take on a producer mindset. Once their work is public, they prioritize where information comes from, how strong the

evidence is, and keeping it up to date—core practices of OER stewardship (Cronin, 2017; DeRosa & Robison, 2017).

An additional qualitative signal is the movement away from neutral responses noted elsewhere in the results: As students encounter concrete editorial decisions (e.g., selecting sources to cite, deciding when an OER needs revision), uncertainty gives way to more informed judgments. This clearer, more critical way of judging sources is a key step for students toward participating on their own in open communities. It also aligns with research showing that cocreating students develop greater agency and a stronger sense of identity (Lubicz-Nawrocka & Bovill, 2023; Werth & Williams, 2021a).

Figure 6

Comparison of Pre-Survey and Post-Survey Responses for the Theme “Perception of OER”



Note. Q6 showed the most notable shift; the other questions displayed consistently positive perceptions.

Modest but Meaningful Change in Access Behaviors

Access to OER and use behaviors were evaluated using Q1b(ii), which asked about the types of OER materials accessed (Figure 7). Changes in access were statistically significant but comparatively modest. Although the median response remained around 2, a statistically significant shift was observed after the activity ($p = .0397$). This indicates a larger, if modest, increase in exposure to different types of OER content, likely facilitated by the hands-on nature of the course module.

Figure 7

Comparison of Pre-Survey and Post-Survey Responses for the Theme “Access to OER” (Q1b[ii])



Note. Responses show a slight increase in access behaviors post-activity.

Students reported engaging with a broader variety of OER formats (e.g., videos, textbooks, interactives) after the activity, yet platform-specific usage patterns (e.g., BCcampus, OER Commons) were largely unchanged. This broadened but shallow adoption suggests that the course successfully exposed students to multiple open modalities without fully converting that exposure into habitual discovery and reuse.

Several factors can explain this pattern. First, the ability to find good OER is a learned skill: Because resources are scattered, students need to know how to search, use the right terms, and navigate each platform. Such competencies are typically underemphasized in disciplinary curricula (Farias-Gaytán et al., 2023; Tinmaz et al., 2022). Second, tight timelines can push students to finish an assignment rather than build repeatable search-and-review habits. Finally, in advanced courses, students often rely on instructor-curated links; without intentional transfer tasks (e.g., find and justify), those habits may slip after the activity or course ends.

In summary, thematic findings collectively highlight that the activity was effective in improving students' understanding of Creative Commons licensing and their perceptions of OER quality and credibility. Although gains in access behaviors were present, they were more modest and suggest that further institutional support may be needed to reinforce regular OER use.

Summarizing Discussion

This study examined the impact of engaging undergraduate students in the creation of OER within a course-based activity. Using a pre–post-survey design and Wilcoxon signed-rank test analysis, we identified statistically significant improvements in students’ understanding of Creative Commons licenses, their perception of OER, and, albeit more modestly, their access to OER materials.

The most substantial gains were observed in the domain of Creative Commons knowledge, particularly in items assessing recognition of license types (Q10) and the ability to match licenses with their appropriate use (Q12). These findings suggest that active participation in OER creation can help demystify complex licensing frameworks and foster applied digital literacy competencies, a pattern similarly reported in studies of OER-enabled pedagogy and student authorship (Trust et al., 2023; Werth & Williams, 2021a; Wiley & Hilton, 2018). Because uncertainty about licensing frequently inhibits reuse and remixing, these gains represent both pedagogical and practical value within open education ecosystems.

The perception of OER theme also showed improvement, particularly in Q6. Prior to the activity, responses clustered at the lower end of the scale, suggesting unfamiliarity or initial skepticism. Post-activity results indicate that even a single structured engagement in OER creation can shift student attitudes, making learners more receptive to open practices, which is a finding that aligns with prior work on co-creation, student agency, and openness (LeMire et al., 2024; Lubicz-Nawrocka & Bovill, 2023; Maultsaid & Harrison, 2023). Trust-related items (Q15 CCS and Q15 FUR) remained high across both time points, reflecting preexisting confidence in OER credibility.

Improvements in the access to OER theme, although statistically significant, were comparatively modest. This suggests that while the learning activity increased awareness and strengthened perceptions, these did not fully translate into autonomous access behaviors. Prior research highlights that students often lack the search skills, repository familiarity, and procedural digital literacy needed to independently locate and evaluate open resources (Cheung et al., 2022; Farias-Gaytán et al., 2023; Tinmaz et al., 2022). This underscores a key challenge in open pedagogy: supporting learners not only in understanding openness but in developing the habits and competencies necessary to act on that understanding.

The absence of change across several items (e.g., prior OER use, ease of access, interactivity features) is also informative. Three considerations help contextualize these patterns. First, ceiling and floor effects likely constrained detectable movement where baseline ratings were already high or conceptions were already formed (Clinton & Khan, 2019). Second, research shows that changes in sourcing habits often require extended practice or repeated exposure across multiple courses rather than a single-semester intervention (Grimaldi et al., 2019; Tlili et al., 2023). Thirdly, the use of single-item Likert measures may have limited sensitivity to small but meaningful changes; multi-item scales could strengthen future assessments (Clinton & Khan, 2019).

These findings support the growing evidence that open pedagogy can be a powerful mechanism for enhancing students’ engagement with and understanding of OER. At the same time, the results suggest that familiarity with licensing and access infrastructures may require continued opportunities, structured guidance, and institutional support to produce deeper and more durable changes in student behavior.

This study had several strengths, including the use of a within-subjects pre–post design that provided a direct measure of change; the alignment of survey constructs with clearly defined domains of licensing knowledge, perception, and access; and the focus on Creative Commons literacy, which is an area that remains underexamined in OER scholarship. At the same time, several limitations should be acknowledged. The modest improvement observed in access behaviors suggests that short-term interventions may be insufficient to shift habitual sourcing practices, which often require repeated opportunities and broader institutional scaffolding. Additionally, because the activity was implemented within a single biochemistry course at one institution, the findings may not be generalizable across disciplines or contexts where students have different levels of digital literacy or prior exposure to OER. Future research could build on this work by incorporating qualitative approaches to capture students’ experiences more deeply, examining longitudinal changes in access behaviors and exploring how similar OER creation activities function in diverse disciplinary or institutional environments.

Conclusions

The integration of OER creation into an undergraduate biochemistry course significantly enhanced students’ knowledge of Creative Commons licensing, improved their perception of open resources, and modestly increased their reported access to OER. Students are not merely passive consumers of educational content but can be meaningfully positioned as co-creators in the learning process.

Engagement with open pedagogy led to more informed, confident, and positive attitudes toward open resources, particularly in areas where prior knowledge was limited. Students demonstrated measurable gains in both knowledge and perception, which are critical for the sustainability and success of open education initiatives. The distinctive contribution of this work lies in documenting measurable gains in licensing literacy, which is an essential yet often overlooked competency in ODL environments.

While access behavior showed some improvement, it remained an area requiring additional support. This underscores the importance of providing clear pathways for students to discover, evaluate, and use OER materials independently.

Although our focus was on quantitative changes, future studies could use qualitative methods to capture students’ voices and experiences as they move through the process of creating OER. These insights would complement the present results by highlighting how students make sense of licensing, authorship, and digital literacy in authentic, public-facing work.

When students create real, public work, they learn practical licensing skills (competency), feel more connected by coauthoring with peers and instructors (relatedness), and make editorial choices for real audiences (autonomy). These drivers help explain why knowledge and critical judgment improved most, while everyday access habits changed less (Werth & Williams, 2021c). Put simply, students started to think like open practitioners before they consistently acted like them.

Incorporating open pedagogy activities within courses should involve conceptual grounding, coached practice, and ideally, independent application across multiple courses. Institutional infrastructure should facilitate repositories, guidance, and sufficient recognition to sustain momentum.

In closing, we recommend that instructors build OER work into the heart of courses rather than adding it on the side. When students create and use open materials as a core activity, they practice digital authorship, learn about information ethics, and contribute to public knowledge. Students should be taught how to find, judge, adapt, and properly attribute OER. Clear instruction can close the gap between students' positive views of OER and their ability to use it independently. Although our work was with biochemistry students, the methods can be adapted to students in other disciplines.

As higher education continues to evolve in response to affordability, equity, and access challenges, student-centered open pedagogy represents a promising avenue for transformative, inclusive education. This study adds to the growing body of evidence that involving students in OER creation not only supports content understanding but also fosters empowerment and engagement in the learning process.

Transparency of Authorship and Use of AI Tools

No generative artificial intelligence tools were used in the design of the study, the collection or analysis of data, or the writing of this manuscript. All content, interpretations, and conclusions are the sole work and responsibility of the authors.

References

- Adil, H. M., Ali, S., Sultan, M., Ashiq, M., & Rafiq, M. (2024). Open education resources' benefits and challenges: A systematic review. *Global Knowledge, Memory and Communication*, 73(3), 274–291. <https://doi.org/10.1108/GKMC-02-2022-0049>
- Arispe, K., & Hoyer, A. (2023). Partnering higher education and K–12 institutions in OER: Foundations in supporting teacher OER-enabled pedagogy. *The International Review of Research in Open and Distributed Learning*, 24(2), 196–212. <https://doi.org/10.19173/irrodl.v24i2.6856>
- Cheung, S. K. S., Wong, B. T. M., & Li, K. C. (2022). Perceived usefulness of open educational resources: Impact of switching to online learning for face-to-face and distance learners. *Frontiers in Psychology*, 13, Article 1004459. <https://doi.org/10.3389/fpsyg.2022.1004459>
- Cho, K. W., & Permzadian, V. (2024). The impact of open educational resources on student achievement: A meta-analysis. *International Journal of Educational Research*, 126, Article 102365. <https://doi.org/10.1016/j.ijer.2024.102365>
- Clinton, V. (2018). Savings without sacrifice: A case report on open-source textbook adoption. *Open Learning: The Journal of Open, Distance and e-Learning*, 33(3), 177–189. <https://doi.org/10.1080/02680513.2018.1486184>
- Clinton, V., & Khan, S. (2019). Efficacy of open textbook adoption on learning performance and course withdrawal rates: A meta-analysis. *AERA Open*, 5(3). <https://doi.org/10.1177/2332858419872212>
- Clinton-Lisell, V. E., Roberts-Crews, J., & Gwozdz, L. (2023). SCOPE of open education: A new framework for research. *The International Review of Research in Open and Distributed Learning*, 24(4), 135–153. <https://doi.org/10.19173/irrodl.v24i4.7356>
- Cronin, C. (2017). Openness and praxis: Exploring the use of open educational practices in higher education. *The International Review of Research in Open and Distributed Learning*, 18(5), 15–34. <https://doi.org/10.19173/irrodl.v18i5.3096>
- DeRosa, R., & Robison, S. (2017). From OER to open pedagogy: Harnessing the power of open. In R. S. Jhangiani & R. Biswas-Diener (Eds.), *Open: The philosophy and practices that are revolutionizing education and science* (pp. 115–124). Ubiquity Press. <https://doi.org/10.5334/bbc.i>
- Farias-Gaytán, S., Aguaded, I., & Ramírez-Montoya, M. S. (2023). Digital transformation and digital literacy in the context of complexity within higher education institutions: A systematic literature review. *Humanities and Social Sciences Communications*, 10, Article 386. <https://doi.org/10.1057/s41599-023-01875-9>
- Griffiths, R., Mislevy, J., & Wang, S. (2022). Encouraging impacts of an open education resource degree initiative on college students' progress to degree. *Higher Education*, 84(5), 1089–1106.

<https://doi.org/10.1007/s10734-022-00817-9>

- Grimaldi, P. J., Basu Mallick, D., Waters, A. E., & Baraniuk, R. G. (2019). Do open educational resources improve student learning? Implications of the access hypothesis. *PLOS ONE*, 14(3), Article e0212508. <https://doi.org/10.1371/journal.pone.0212508>
- Hilton, J., III. (2020). Open educational resources, student efficacy, and user perceptions: A synthesis of research published between 2015 and 2018. *Educational Technology Research and Development*, 68(3), 853–876. <https://doi.org/10.1007/s11423-019-09700-4>
- Jhangiani, R., Pakkal, O., & Xia, X. (2025). The multi-year impact of Canada's first zero textbook cost initiative. *Open Praxis*, 17(2), 326–348. <https://doi.org/10.55982/openpraxis.17.2.832>
- Lan, M., Liu, H., & Pan, Q. (2025). Unpacking the digital literacy–self-regulated learning nexus: A systematic review and a three-level meta-analysis. *Educational Research Review*, 48, Article 100713. <https://doi.org/10.1016/j.edurev.2025.100713>
- LeMire, S., Anders, K., & Pantuso, T. (2024). Exploring first-generation student experiences with OER textbooks. *College & Research Libraries*, 85(7), Article 952. <https://doi.org/10.5860/crl.85.7.952>
- López-Núñez, J.-A., Alonso-García, S., Berral-Ortiz, B., & Victoria-Maldonado, J.-J. (2024). A systematic review of digital competence evaluation in higher education. *Education Sciences*, 14(11), Article 1181. <https://doi.org/10.3390/educsci14111181>
- Lubicz-Nawrocka, T., & Bovill, C. (2023). Do open students experience transformation through co-creating curriculum in higher education? *Teaching in Higher Education*, 28(7), 1744–1760. <https://doi.org/10.1080/13562517.2021.1928060>
- Maultsaid, D., & Harrison, M. (2023). Can open pedagogy encourage care? Student perspectives. *The International Review of Research in Open and Distributed Learning*, 24(3), 77–98. <https://doi.org/10.19173/irrodl.v24i3.6901>
- Meger, E., Schwartz, M., & Freeman, W. (2021). Participation in OER creation: A trajectory of values. *The Open/Technology in Education, Society, and Scholarship Association Journal*, 1(2), 1–10. <https://doi.org/10.18357/otessaj.2021.1.2.13>
- Mercer-Mapstone, L., Dvorakova, L. S., Matthews, K. E., Abbot, S., Cheng, B., Felten, P., Knorr, K., Marquis, E., Shammass, R., & Swaim, K. (2017). A systematic literature review of students as partners in higher education. *International Journal for Students as Partners*, 1(1), 15–37. <https://doi.org/10.15173/ijsap.vii1.3119>
- Mirzaei, A., Carter, S. R., Patanwala, A. E., & Schneider, C. R. (2022). Missing data in surveys: Key concepts, approaches, and applications. *Research in Social and Administrative Pharmacy*, 18(2), 2308–2316. <https://doi.org/10.1016/j.sapharm.2021.03.009>

- Mishra, S. (2025). A review of reviews on open educational resources. *Open Praxis*, 17(2), 305–325. <https://doi.org/10.55982/openpraxis.17.2.851>
- Mullens, A. M., & Hoffman, B. (2023). The affordability solution: A systematic review of open educational resources. *Educational Psychology Review*, 35, Article 72. <https://doi.org/10.1007/s10648-023-09793-7>
- Nobre, A. (2021). Open educational practices and resources in the higher education learning environment. In D. Chatham (Ed.), *Advancing online course design and pedagogy for the 21st century learning environment* (pp. 93–111). IGI Global Scientific Publishing. <https://doi.org/10.4018/978-1-7998-5598-9.ch006>
- Omland, M., Hontvedt, M., Siddiq, F., Amundrud, A., Hermansen, H., Mathisen, M. A. S., Rudningen, G., & Reiersen, F. (2025). Co-creation in higher education: A conceptual systematic review. *Higher Education*, 90, 1017–1047. <https://doi.org/10.1007/s10734-024-01364-1>
- Santos-Hermosa, G. (2024). Impact and implementation of UNESCO’s recommendation on open educational resources in academic libraries: SPARC Europe case study. *Research in Learning Technology*, 32, Article 3183. <https://doi.org/10.25304/rlt.v32.3183>
- Schmied, A., Ntonia, I., Ng, M. K. J., Zhu, Y., Gibbs, F., & Zou, H. G. (2024). Co-creating with students to promote science of learning in higher education: An international pioneer collaborative effort for asynchronous teaching. *Trends in Neuroscience and Education*, 35, Article 100229. <https://doi.org/10.1016/j.tine.2024.100229>
- Short, C. R., Hilton, B., Hilton, J., III, Wiley, D., Chaffee, R., Guilmett, J., & Darrow, J. (2024). Higher education instructors’ perceptions of open pedagogy: An exploratory study of open pedagogy definitions in practice. *Open Learning: The Journal of Open, Distance and e-Learning*. Advance online publication. <https://doi.org/10.1080/02680513.2024.2334237>
- Spencer, C., Angra, A., Dósa, K., & Jones, A. (2025). Undergraduate learning gains and learning efficiency in a focused open education resource. *International Review of Research in Open and Distributed Learning*, 26(2), 184–204. <https://doi.org/10.19173/irrodl.v26i2.8117>
- Tietjen, P., & Asino, T. I. (2021). What is open pedagogy? Identifying commonalities. *The International Review of Research in Open and Distributed Learning*, 22(2), 185–204. <https://doi.org/10.19173/irrodl.v22i2.5161>
- Tinmaz, H., Lee, Y.-T., & Baber, H. (2022). A systematic review on digital literacy. *Smart Learning Environments*, 9, Article 21. <https://doi.org/10.1186/s40561-022-00204-y>
- Tlili, A., Garzón, J., Salha, S., Huang, R., Xu, L., Burgos, D., Denden, M., Farrell, O., Farrow, R., Bozkurt, A., Amiel, T., McGreal, R., López-Serrano, A., & Wiley, D. (2023). Are open educational resources (OER) and practices (OEP) effective in improving learning achievement? A meta-analysis and research synthesis. *International Journal of Educational Technology in Higher Education*, 20(1),

Article 54. <https://doi.org/10.1186/s41239-023-00424-3>

- Tlili, A., Zhang, X., Lampropoulos, G., Salha, S., Garzón, J., Bozkurt, A., Huang, R., & Burgos, D. (2025). Uncovering the black box effect of open educational resources (OER) and practices (OEP): A meta-analysis and meta-synthesis from the perspective of activity theory. *Humanities and Social Sciences Communications*, 12, Article 504. <https://doi.org/10.1057/s41599-025-04644-y>
- Trust, T., Maloy, R. W., & Edwards, S. (2023). College student engagement in OER design projects: Impacts on attitudes, motivation, and learning. *Active Learning in Higher Education*, 24(3), 353–371. <https://doi.org/10.1177/14697874221081454>
- UNESCO. (n.d.). *Open educational resources (OER)*. <https://en.unesco.org/themes/building-knowledge-societies/oer>
- UNESCO. (2024, November 20). *Dubai Declaration on Open Educational Resources (OER): Digital public goods and emerging technologies for equitable and inclusive access to knowledge*. <https://unesdoc.unesco.org/ark:/48223/pf0000392271.locale=en>
- Werth, E., & Williams, K. (2021a). Exploring student perceptions as co-authors of course material. *Open Praxis*, 13(1), 53–67. <https://doi.org/10.5944/openpraxis.13.1.1187>
- Werth, E., & Williams, K. (2021b). Learning to be open: Instructor growth through open pedagogy. *Open Learning: The Journal of Open, Distance and e-Learning*, 38(4), 301–314. <https://doi.org/10.1080/02680513.2021.1970520>
- Werth, E., & Williams, K. (2021c). What motivates students about open pedagogy? Motivational regulation through the lens of self-determination theory. *The International Review of Research in Open and Distributed Learning*, 22(3), 34–54. <https://doi.org/10.19173/irrodl.v22i3.5373>
- Wiley, D., & Hilton, J. L., III. (2018). Defining OER-enabled pedagogy. *The International Review of Research in Open and Distributed Learning*, 19(4), 133–147. <https://doi.org/10.19173/irrodl.v19i4.3601>



August – 2026

The 20-Year Evolution of Distance Education: A Bibliometric and LDA-Based Topic Model Analysis

Mehmet Yavuz^{1*}, Şener Balat¹, Bünyami Kayalı², and Emirhan Gülen¹

¹Bingöl University, Türkiye; ²Bayburt University, Türkiye; *Corresponding Author

Abstract

Although research on distance education has increased substantially in recent years, comprehensive analyses capturing the overall thematic development of the field remain limited. Modeling studies are necessary to address this gap and to uncover long-term research trends. This study aimed to reveal the trends of research conducted in the field of distance education in the last two decades by using machine learning methodology and bibliometric analysis. The unique aspect of the study was its use of the Latent Dirichlet Allocation (LDA) algorithm, a machine learning method for analysing large-scale data. Within the scope of the study, 54,444 articles in the Web of Science database were analysed. Bibliometric analysis revealed that distance education had gained significant momentum, especially in the post-pandemic period, with a wide range of applications in different disciplines such as education, management, and health. As a result of LDA analysis, 19 thematic topics were identified. Among these, digitalization, artificial intelligence, Web-based learning, and social interaction and collaboration stood out. Time series analysis showed an increasing trend for topics such as artificial intelligence and emergency distance education over the years, with less interest shown in topics such as Web-based learning and program design. The study emphasized that distance education is in a technology-driven transformation process and pointed to areas that will offer new opportunities for researchers. Furthermore, recommendations for decision-makers, such as investments in digital infrastructure, integration of artificial intelligence-based systems, and the establishment of quality standards, also emerged.

Keywords: distance education, machine learning, topic modeling, LDA, bibliometric analysis

Introduction

Distance education has driven an important paradigm shift in the field of education, thanks to rapidly developing technological infrastructures and increasing accessibility opportunities (Kaldoudi et al., 2010). With the widespread use of Internet technologies, distance education applications, which offer learning opportunities independent of time and space, have both diversified and enriched individuals' learning processes (Park et al., 2012). Distance education has not only replaced traditional educational environments but also paved the way for the development of innovative pedagogical approaches (Watermeyer et al., 2021).

The last two decades have been a remarkable period for distance education research as technologies, methods, and pedagogical approaches used in distance education have rapidly diversified and developed significantly (Bates, 2014). In parallel with these developments, there has been a significant increase in research on the effects of distance education on learning outcomes (Bozkurt et al., 2015). Analysing these studies systematically and revealing general trends in the field is of great importance in terms of both understanding the existing body of knowledge and shaping future research directions (Özmen & Kan, 2021). However, with the increase in technology-oriented research and publishing opportunities, it has become almost impossible for researchers to determine the trends and development of the field by analysing the literature in a holistic manner in the face of the wide-ranging scientific content (Debortoli et al., 2016).

Technological developments in the field of machine learning have brought about a significant transformation in educational research (Peña-Ayala, 2014). Thanks to their superior ability to analyse large data sets and uncover complex relationships, machine learning algorithms have enabled faster, easier, and more detailed examination of trends in the field of distance education (Aggarwal & Zhai, 2012). Integrating these technologies into distance education research can enable a more comprehensive and objective evaluation of the studies conducted in the past (Blei, 2012). Thus, clearer and more reliable answers can be found for questions such as which topics are prominent in distance education research, which methodologies are preferred, and which gaps are found.

Theoretical Framework

The concept of distance education emerged from flexible and decentralized learning needs and developed later than general education research (Zawacki-Richter & Naidu, 2016). Early studies in the 1960s were limited (Peters, 2014), descriptive, and criticized for weak design (Saba, 2000). In the 1980s, distance education began to professionalize with journals such as *Distance Education* (Delling, 1971). During the 1980s, research first emphasized defining the field and founding universities, then shifted to instructional technologies, design, and interaction strategies (Bozkurt & Zawacki-Richter, 2021).

In the 1990s, topics that addressed improving quality, developing student support services, and early online learning became central (Berge & Mrozowski, 2001). Technological advances reduced costs and expanded access (Howell et al., 2003). Instructional design remained important (Bozkurt & Zawacki-Richter, 2021), alongside calls for new methods and paradigms (Bozkurt et al., 2022). In the 2000s, research focused on collaborative learning (Oviatt et al., 2012), interaction (Aydın et al., 2020) and blended learning (Gunawardena & McIsaac, 2013). Studies also explored student and instructor experiences, as well as cultural, social, and ethical dimensions (Aydın et al., 2020).

From 2010 to 2020, debates on methodological rigor and comparisons with face-to-face education dominated (National Forum, 2021). MOOCs, technology-supported collaboration, and the shift from instructional to learning design stood out (Xiao, 2023). Open education and student support services gained emphasis (Bozkurt & Zawacki-Richter, 2021). After 2020, the COVID-19 pandemic accelerated online learning, leading to increased publications on personalized learning (Pai et al., 2021), learning analytics (Hamal et al., 2022), data mining, and machine learning (Yörük, 2024).

In this study, the interpretation of distance education research and the construction of the proposed taxonomy were supported by the synthesis of five established theoretical perspectives in the field. Moore's transactional distance theory (Moore & Kearsley, 2012) provided a lens for explaining the relationship between instructional structure and dialogue, while the community of inquiry (CoI) framework (Garrison et al., 2000) contributed to the understanding of social presence, collaboration, and affective dimensions. Self-regulated learning (SRL) theory (Zimmerman, 2002) grounded issues of learner autonomy and motivation. Connectivism (Siemens, 2005) conceptualized the rise of digitalization and network-based learning, and the technology acceptance model (TAM; Davis, 1989) helped explain technological adoption and perceived usefulness. Taken together, these theoretical perspectives guided the formulation of research questions, the labelling of topic clusters, and the interpretation of the taxonomy. This enabled a theory-driven reading of the thematic evolution of distance education research rather than a purely descriptive account.

Topic Modeling

Topic modeling is a text mining technique that uses probabilistic models to uncover hidden semantic topics in large text collections (Blei, 2012). LDA, one of the most widely used algorithms, enables semantic analysis by identifying latent topics in document collections, especially in text mining and natural language processing (Griffiths & Steyvers, 2004).

LDA is a three-level hierarchical Bayesian model that represents each document as a combination of topics and each topic as a probabilistic distribution of words (Wallach, 2006). It treats words in documents as variables observed in an iterative probabilistic process based on a Dirichlet distribution (Karl et al., 2015). Rather than relying on a predefined taxonomy or fixed dictionary, LDA identifies topics based on the co-occurrence and contextual neighborhoods of words. This probabilistic framework allows the algorithm to group conceptually related terms into the same topic even if their lexical forms vary over time; for instance, if evolving terms share a similar set of surrounding words, they are mapped to the same underlying theme. By automating the extraction of these semantic structures, LDA has surpassed traditional bibliometric methods and manual reviews in both efficiency and depth (Aziz et al., 2022).

With its robust analytical capabilities, LDA has become an essential tool in academic and industrial research (Griffiths & Steyvers, 2004). It has been widely applied in the social sciences: (a) Gürcan et al. (2021) analysed 41,720 articles on human-computer interaction; (b) Aziz et al. (2022) examined 5,942 publications in finance; (c) Fu et al. (2024) assessed 27,316 student comments on MOOCs; and (d) Zankadi et al. (2023) analysed 120,000 tweets from 12,187 students. In the present study, the resulting topic structures were further interpreted and validated through independent expert labelling and consensus discussions, transforming the statistically derived topic clusters into a theory-informed and reliable thematic taxonomy.

This study examined trends in distance education research over the past two decades using a combined bibliometric and machine learning–based approach. Analysing a large corpus of 54,444 Web of Science (WoS)-indexed publications, it provided a data-driven taxonomy that mapped thematic evolution and identified priority areas for theory, policy, and practice. Unlike prior distance education and open and distance learning (ODL) syntheses, which largely relied on manual content analysis and shorter time windows, (e.g., 2000–2008 or 2014–2019), this study applied LDA to capture long-term thematic shifts and emerging trends at scale. This approach enabled a more precise identification of key developments—such as digitalisation, artificial intelligence, and social interaction—thereby offering a clearer and more systematic roadmap for future distance education research. The study also contributed to the redefinition of distance education theories in the context of digital transformation, with potential to offer a theoretical framework to broaden the impact of technology in education. From a practical perspective, the study yielded concrete recommendations for policy development regarding the integration of artificial intelligence (AI)-based systems and investments in digital infrastructure. Machine learning-based analysis has been touted as a useful methodology not only for distance education but also expected to bring a new perspective to educational research. In this study, published research on distance education was examined, and trends were analysed using machine learning algorithms. The following research questions guided this study.

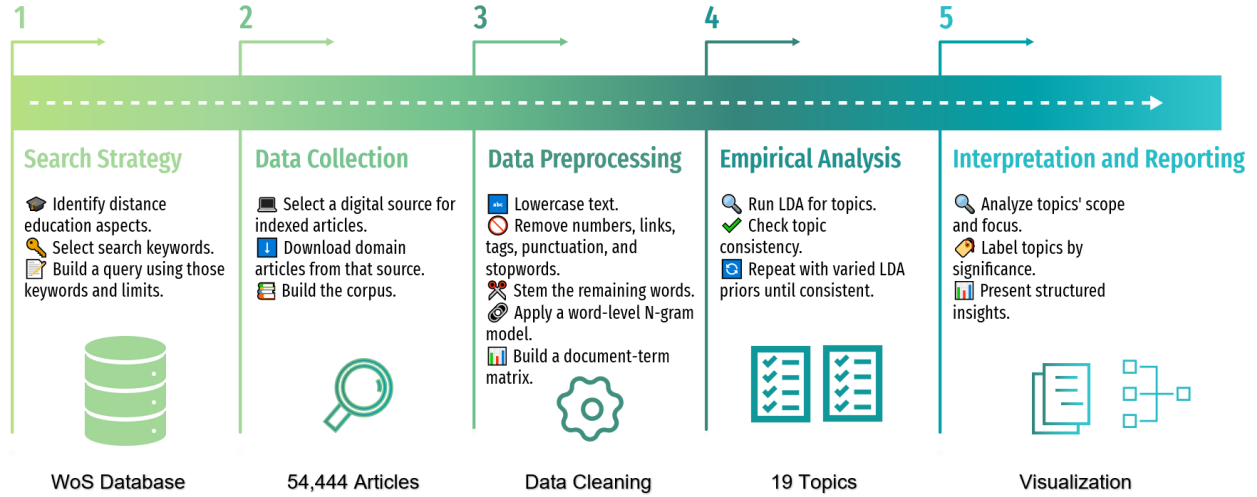
1. What have been the bibliometric characteristics of distance education research over the last 20 years?
2. What have been the main topics of distance education research over the last 20 years?
3. What have been the temporal trends in distance education research topics over the last 20 years?

Methodology

This study was conducted within the framework of machine learning methodology and using LDA, a text mining technique, to reveal the trends of distance education research over the last two decades. The methodological approach of the study consisted of five main stages: (a) search strategy, (b) data collection, (c) data preprocessing, (d) empirical analysis, and (e) interpretation and reporting. The scope was limited to English-language, WoS-indexed publications (2005–2024), which may have limited generalizability to non-English and non-WoS research. Works published in local journals or in different languages may offer different perspectives on thematic evolution and research trends. During model selection, topic coherence values for nearby topic numbers yielded comparable thematic patterns, supporting the robustness of the selected model. Records without abstracts were excluded due to insufficient textual information; however, as these constituted a small proportion of the corpus and were not concentrated in specific years or subject areas, their exclusion is unlikely to have systematically biased the identified topics. Figure 1 illustrates the stages and activities in the research process.

Figure 1

Research Methodology and Process Flow



The first stage involved developing a search strategy on distance education by analysing relevant concepts in the literature to obtain consistent data. Then, a dataset was created and academic articles fitting the criteria were compiled. This textual dataset was preprocessed with various techniques to make it suitable for topic modeling analysis.

Subsequently, the dataset was analysed using the LDA method. In the final stage, the topics obtained were evaluated, their contents analysed, and temporal trends revealed. Based on these analyses, a systematic taxonomy for distance education research was proposed.

In the first stage, a general search was conducted in the WoS database using the term distance education to determine frequently used keywords in the field. The distribution frequencies of these keywords were analysed and then included in the search query. Using them, a comprehensive search was carried out in the WoS database, which was preferred as it covers journals indexed in Science Citation Index Expanded, Social Sciences Citation Index, Emerging Sources Citation Index, and Arts & Humanities Citation Index (Charan, 2014). The exact search query and filtering criteria used in Web of Science are provided below.

TS=(“distance education” OR “e-learning” OR “online learning” OR “emergency remote teaching” OR “distance learning” OR “online education” OR “online teaching” OR “e-teaching” OR “distance teaching” OR “web-based learning” OR “web-based teaching”) AND DT=(Article) AND LA=(English) AND PY=(2005-2024)

On January 3, 2025, a total of 54,812 articles on distance education were accessed. In preprocessing, the title, abstract, and keywords of the articles were included, texts reformatted, and meaningless components removed. For 499 articles without abstracts, an additional search found 131, while 368 were excluded. Thus, 54,444 articles were included. Bibliometric data such as subject area, journal, author, and country were also obtained, forming a secondary dataset for descriptive analyses.

The data cleaning and preprocessing steps were carefully designed to ensure both data quality and conceptual integrity. Texts were tokenized, and noise was removed by eliminating HTML tags, punctuation, numbers, very short words, and common stop words (e.g., and, is, the) while preserving meaningful content. All words were lowercased and lemmatized to standardize word forms and make the texts suitable for modeling. Each article was then represented as word vectors using the bag-of-words approach and transformed into a document term matrix for topic modeling. All preprocessing and data cleaning procedures were implemented programmatically using Python-based natural language processing libraries, enabling a systematic, reproducible, and transparent workflow.

The LDA model was applied to the corpus to identify latent thematic structures. LDA is a probabilistic topic modeling approach that uncovers topics based not only on word frequency but also on patterns of word co-occurrence. Within this framework, each document is represented as a mixture of topics, and each topic as a probability distribution over words. Through an iterative estimation process, the model simultaneously learns which topics characterize which documents and which words represent each topic, thereby enabling the identification of conceptually coherent latent topics distributed across the corpus rather than simple clusters of frequently used keywords. Prior parameters α , β , and K were estimated for model optimization (Blei, 2012). The α parameter indicated topic distribution per document, β referred to word distribution per topic, and K represented the number of topics. Following recommendations for short texts, $\alpha = 0.1$ and $\beta = 0.01$ were used (Li et al., 2016). The number of topics (K) was tested from 10 to 50, and the coherence metric CV was calculated to select the optimal K (Konrad, 2017), with the maximum score achieved at 19 topics. Model coherence was then used to evaluate topic meaningfulness. The final stage involved interpreting topics, analysing temporal trends, and presenting findings systematically. Independent labelling by multiple coders was followed by consensus discussions, ensuring agreement on topic interpretation. Topic momentum was calculated using LDA-derived yearly topic proportions. These proportions were normalized to control for annual publication volume differences and smoothed using a moving average. Momentum was then defined as the year-to-year change in the smoothed topic prevalence, indicating increases or decreases in research attention over time. Each topic included up to 20 descriptive keywords, and in addition to topic modeling, the bibliometric profile of the corpus was analysed in detail.

Based on these procedures, the identified topics reflected empirically derived latent topics interpreted at a conceptual level, rather than surface-level keyword clusters. Accordingly, the proposed taxonomy captured conceptually grounded shifts in research focus rather than surface-level changes in educational vocabulary. The mappings among latent topics, theoretical lenses, and taxonomic categories are summarized in Figure 2.

Figure 2

Functional Mapping of ODL Theories to Research Design and Taxonomy

Theoretical Lens	Primary Focus	Research Question Rationale	Topic Labelling	Topic Labelling
Transactional Distance	System & Structure: Focuses on institutional and technical organization.	RQ1 & RQ2: Investigates if ODL has evolved from individual trials into a professional structure.	Labels technical clusters (design, management) as structure based on Moore's framework.	Interprets ODL as an evolving institutional system rather than just a delivery method.
Community of Inquiry	Interactivity & Presence: Focuses on social and cognitive bonds in digital spaces.	RQ2: Questions if technological shifts have sacrificed human interaction (dialogue) for technical tools.	Categorizes discussion and collaboration clusters under the theoretical concept of social presence.	Explains social interaction shifts as an evolution of format rather than a loss of presence.
Self-Regulated Learning	Individual Agency: Focuses on learner's psychological and strategic processes.	RQ1 & RQ3: Questions the shift in responsibility from the teacher to the individual learner.	Groups motivation and success clusters under the theoretical umbrella of learner agency.	Validates the high frequency of SRL as empirical proof of a pedagogical shift toward the learner.
Connectivism	Networked Ecosystems: Focuses on knowledge distributed across digital nodes.	RQ3: Questions if knowledge remains human-centric or is distributed across networks and algorithms.	Labels AI and machine learning clusters as digital nodes of knowledge.	Provides a theoretical rationale for the shift from platforms to AI-driven intelligent ecosystems.
Technology Acceptance	User Adoption & Intent: Focuses on the perceived value and ease of new tools.	RQ1 & RQ2: Investigates if infrastructure growth is driven by technical existence or human acceptance.	Labels perceived usefulness and ease of use directly as technology acceptance.	Interprets infrastructure relevance through the lens of user behavioral intent and adoption cycles.

Findings

The findings of this study have been organized by the three main research questions. First, bibliometric analysis results show the distribution of articles by years, subject areas, productive journals, and high-frequency terms. Second, the topics modeling discovered through the LDA method are presented and interpreted in context. Third, the temporal trends of these topics modeling are analysed and visualized with appropriate graphics.

Bibliometric Analysis

To answer the first research question, this section presents the bibliometric characteristics of studies in the field of distance education. First, the distribution of these studies by year is shown in Figure 3 with frequency values.

Figure 3

Distribution of Articles by Year (between 2005 and 2024)

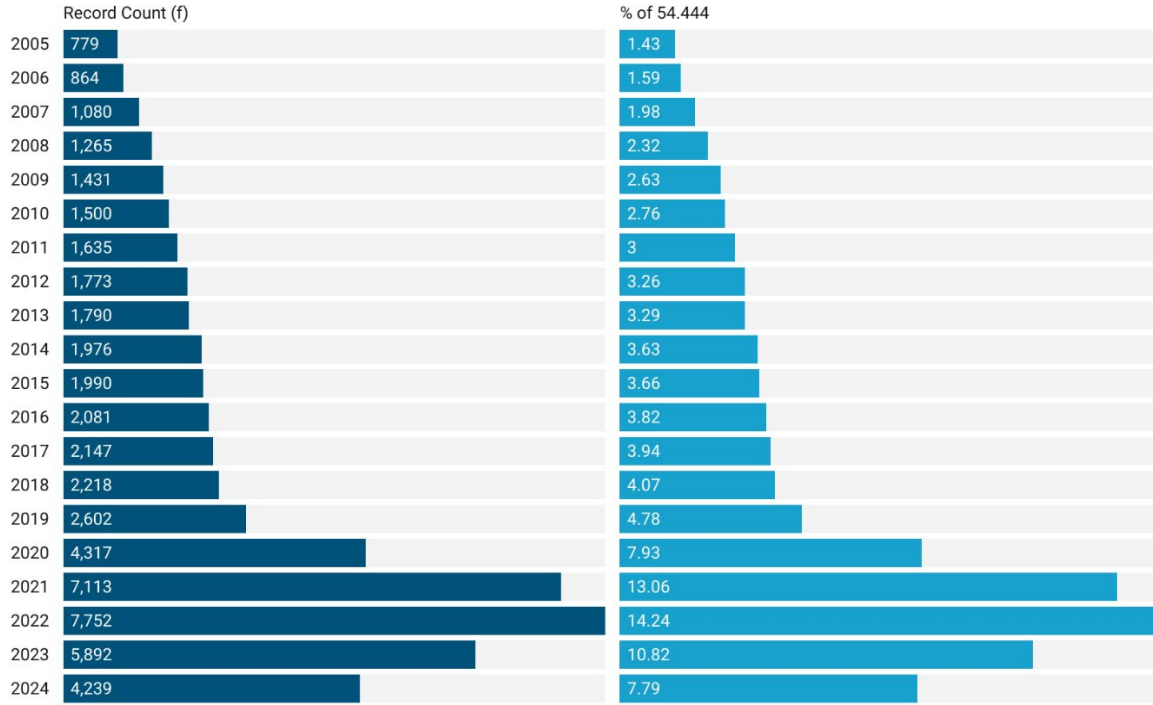
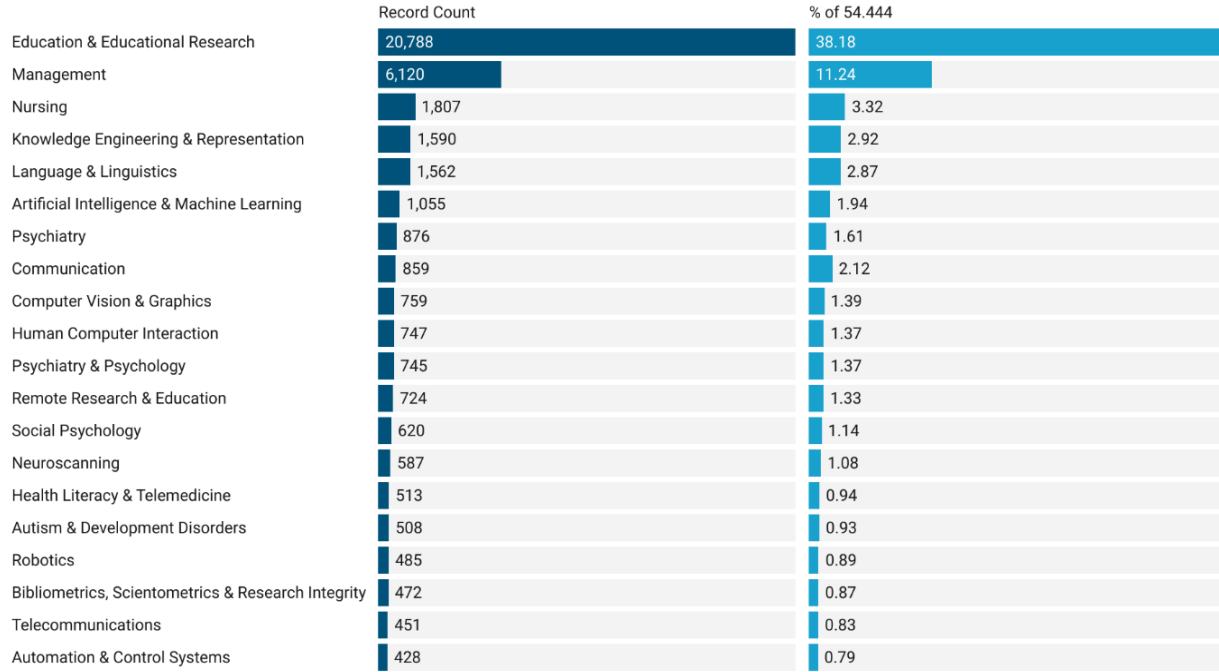


Figure 3 shows that academic interest in distance education has risen markedly since 2005. While 779 articles were published in 2005, the number grew steadily, reaching 1,500 in 2010 and peaking at 7,752 in 2022. In the last four years, 4,317 articles appeared in 2020, increasing to 7,113 in 2021 and 7,752 in 2022, but dropping to 5,892 in 2023 and 4,239 in 2024. These figures highlighted distance education as a growing research focus, with a post-2020 surge driven by the COVID-19 pandemic.

Another bibliometric result was concerned with subject areas. Figure 4 presents the percentage and frequency values of the top 20 fields represented in distance education studies, revealing which disciplines have dominated and which have been explored less.

Figure 4

Topics with the Highest Number of Articles



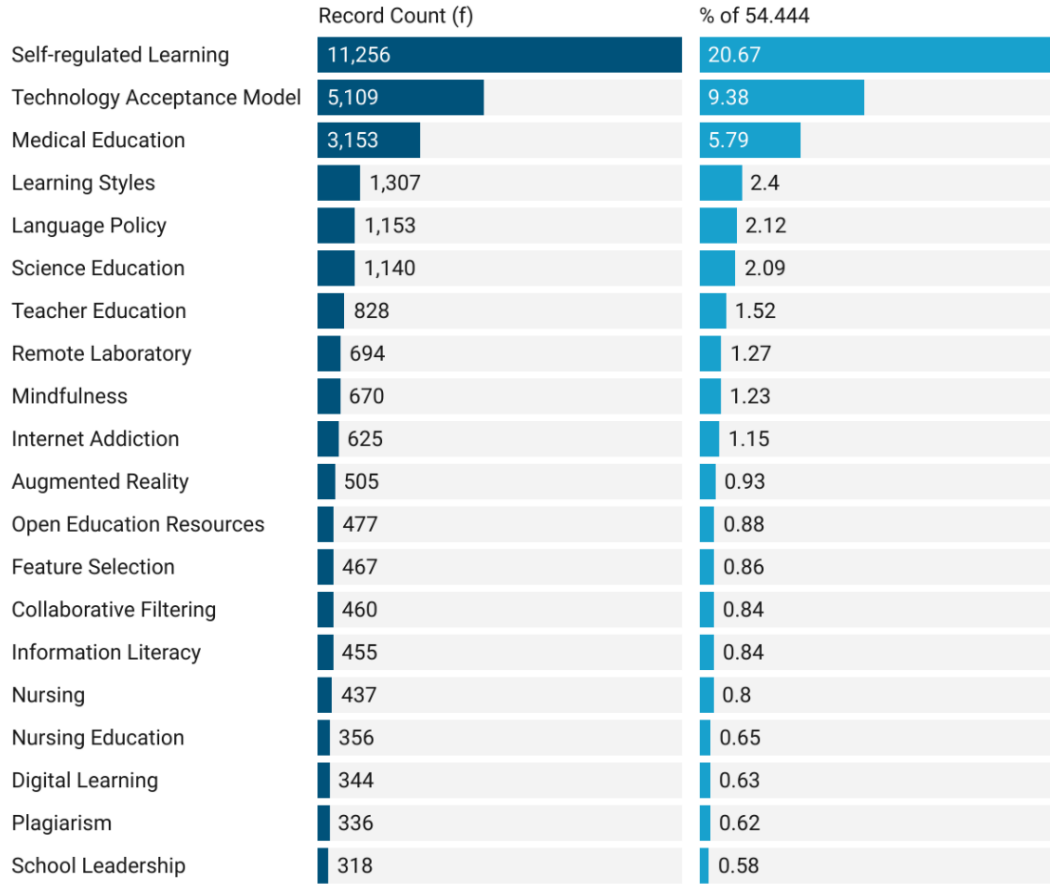
Created with Datawrapper

Figure 4 illustrates that most studies in distance education fell under education and educational research, with 20,788 studies (38.18%); this highlights the decisive role of pedagogical approaches as the basis of the field. The second-largest area was management with 6,120 studies (11.24%), which reflected interest in the managerial dimension. This was followed by health (1,807; 3.32%), then knowledge engineering and representation (1,590; 2.92%). Other notable areas included language and linguistics (2.87%), artificial intelligence and machine learning (1.94%), and psychiatry (1.61%). Although smaller in share, disciplines such as robotics (0.89%), bibliometrics and scientometrics (0.87%), telecommunications (0.83%), and automation and control systems (0.79%) have enriched the literature by addressing technological and methodological aspects of distance education.

The sub-topics variable was also analysed. The first 20 sub-subjects, showing how distance education research has focused on more specific topics, are presented in Figure 5 with their frequency and percentage values.

Figure 5

Sub-Topics with the Highest Number of Articles



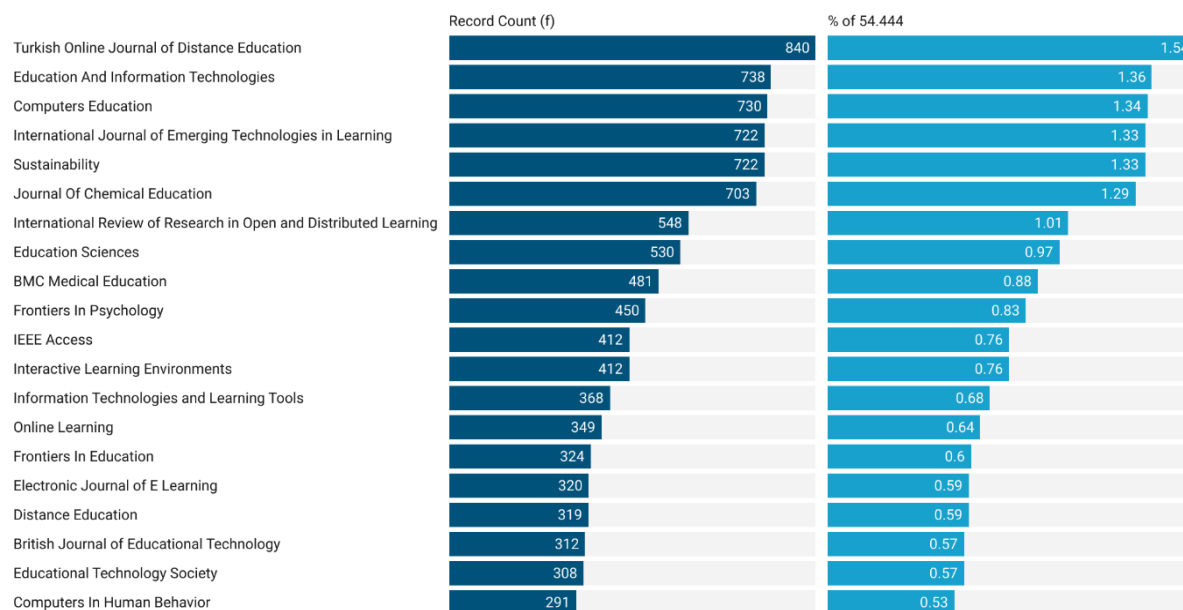
Created with Datawrapper

When Figure 5 is analysed, self-regulated learning emerged as the leading sub-subject with 11,256 articles (20.67%), indicating the centrality of individual learning skills in distance education. The technology acceptance model followed with 5,109 articles (9.38%), stressing technology adoption, while medical education ranked third with 3,153 articles (5.79%), reflecting the significance of distance education in health fields. Other sub-subjects included teacher education (1.52%), Internet addiction (1.15%), and open educational resources (0.88%), with smaller contributions from digital learning (0.63%), plagiarism (0.62%), and school leadership (0.58%).

Journals with the highest publication counts were also analysed as a variable. Results revealed that distance education studies were distributed across diverse journals, with leading outlets shown in Figure 6.

Figure 6

The Most Productive Journals



Created with Datawrapper

Figure 6 presents the distribution of the top 20 journals publishing the highest number of articles on distance education. Although Turkish Online Journal of Distance Education ranked first (1.54%), followed closely by Education and Information Technologies (1.36%) and Computers & Education (1.34%), the relatively low percentage values indicate that publications are not concentrated within a single dominant outlet. Instead, the findings suggest a dispersed publication pattern across multiple journals.

Notably, many of the leading journals focus on educational technology and digital learning environments, highlighting the strong connection between distance education research and technology-oriented scholarship. Furthermore, the presence of interdisciplinary journals such as Sustainability suggests that distance education has expanded beyond traditional educational research boundaries. Overall, the distribution reflects the field's multidisciplinary and widely disseminated nature.

Topic Modeling Analysis in Distance Education

An LDA-based topic modeling analysis of distance education research was conducted to determine the topics modeling and percentage distribution of studies. Figure 7 shows the extent to which different topics have been examined in the literature, along with the prominent keywords. The analysis included 19 topics ranging from digital education technologies to learner-centred designs, and from the effects of the pandemic to mobile computing security.

Figure 7

Ratings of Topics and Top Keywords Discovered by LDA

Rate	Topic name	LDA keywords
11.06	Digitalization in Distance Education	education, distance, educational, technology, digital, university, teacher, process, information, high, use, development, research, teach, article, new, communication, training, train, method
10.96	Web-Based Learning	Web, base, user, learning, content, resource, information, learner, tool, use, platform, provide, material, game, system, video, technology, paper, design, educational
10.61	Social Interaction and Collaboration	social, environment, learning, interaction, experience, community, discussion, collaborative, communication, support, activity, learner, practice, virtual, work, group, peer, presence, role, medium
10.34	Artificial Intelligence	data, model, propose, method, algorithm, base, network, performance, time, feature, accuracy, approach, improve, prediction, paper, real, neural, machine, problem, classification
8.24	Technology-Based Education	research, design, education, approach, paper, technology, development, framework, practice, literature, review, educational, provide, finding, challenge, article, base, focus, high, future
8.23	Course and Program	education, university, program, distance, academic, faculty, institution, high, support, need, service, programme, open, project, provide, library, offer, work, course, country
6.06	Problem-Based Learning	model, process, problem, base, knowledge, evaluation, present, concept, approach, develop, work, design, task, propose, paper, solve, project, different, environment, level
5.73	Motivation and Success	self, learner, engagement, learning, academic, level, motivation, factor, satisfaction, high, efficacy, environment, cognitive, analysis, performance, relationship, significant, interaction, effect, scale
5.57	Health Education	health, care, patient, program, train, training, intervention, knowledge, practice, nurse, professional, participant, nursing, clinical, skill, education, base, improve, module, method
4.01	Medical Education Research	medical, method, clinical, year, conclusion, survey, score, question, teach, questionnaire, assess, objective, module, curriculum, resident, evaluate, dental, medicine, mean, examination
3.63	Blended Learning	course, face, teach, assessment, blend, instructor, class, feedback, classroom, learning, design, traditional, activity, instruction, semester, assignment, year, evaluation, undergraduate, grade
3.37	Technology Acceptance	model, factor, use, influence, technology, perceive, intention, quality, satisfaction, acceptance, structural, research, effect, affect, usefulness, adoption, equation, impact, significant, relationship
3.20	Emergency Remote Teaching	pandemic, COVID, teach, face, education, challenge, experience, remote, university, emergency, change, transition, class, impact, shift, time, learning, institution, crisis, academic
2.15	Experiential Learning	group, test, video, control, score, difference, effect, significant, experimental, compare, post, base, pre, knowledge, performance, participant, high, significantly, show, lecture
1.75	Research and Method	data, research, analysis, interview, qualitative, method, collect, questionnaire, perception, quantitative, participant, finding, conduct, survey, faculty, university, aim, sample, structure
1.56	System and Infrastructure	mobile, laboratory, device, virtual, application, service, remote, experiment, cloud, engineering, LAB, environment, energy, real, technology, security, network, smart, robot, Internet
1.28	Affective Skills	health, stress, anxiety, mental, emotion, level, high, emotional, psychological, social, physical, risk, increase, negative, time, associate, effect, COVID, age, female
1.08	MOOC	teacher, teach, language, English, teaching, MOOC, MOOCs, foreign, classroom, write, skill, practice, professional, Chinese, read, EFL, development, learner, massive, service
0.79	Learning Environment	school, child, parent, family, adult, age, rural, young, home, secondary, primary, people, support, Internet, old, education, access, disability, country, low

Prominent topics in Figure 7 included digitalisation in distance education (11.06%), Web-based learning (10.96%), social interaction and collaboration (10.61%), and artificial intelligence (10.34%). In contrast learning environment (0.79%) and system and infrastructure (1.56%) showed the lowest representation.

Other notable topics were technology-based education (8.24%), course and programme (8.23%), and problem-based learning (6.06%). Emergency remote teaching (3.20%) highlighted the effects of the COVID-19 pandemic, while less frequent topics such as affective skills (1.28%) pointed to growing attention on the emotional dimension of education. The distribution of distance education topic modeling is visualised in Figure 8.

Figure 8

Percentage Distribution of Distance Education Topic Modeling

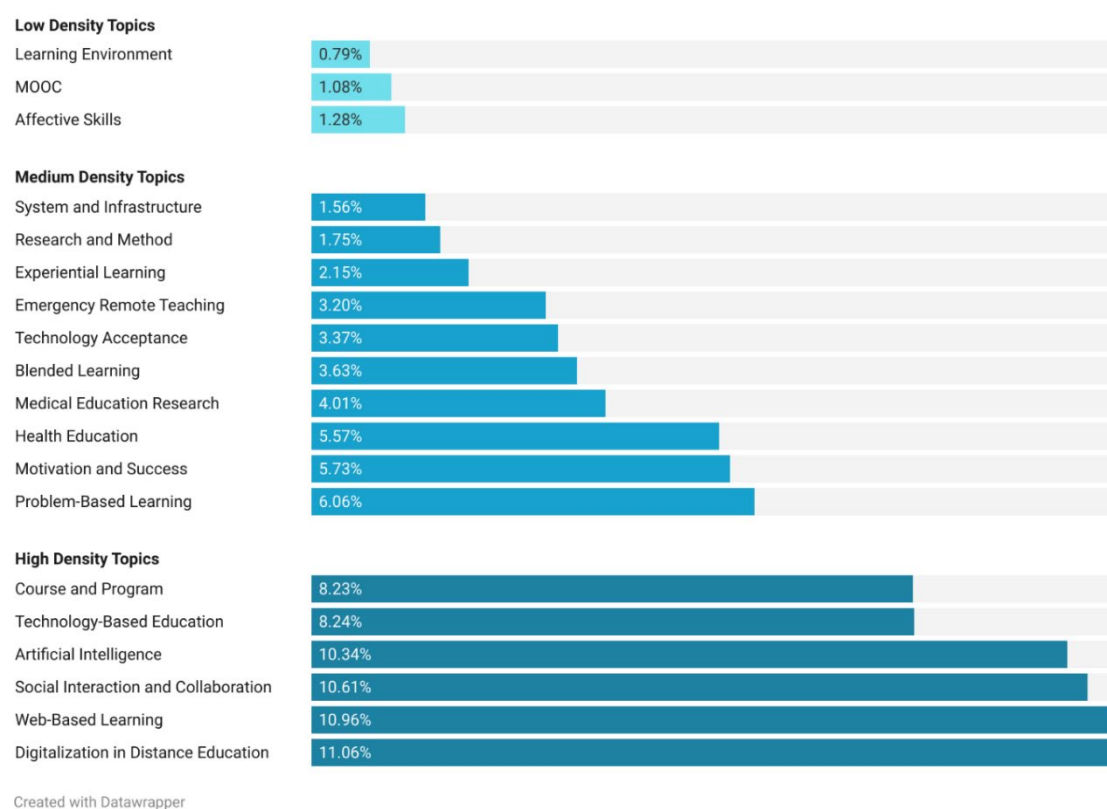


Figure 8 shows the percentage distribution of topics from the topic modeling analysis, grouped as high, medium, and low intensity. High-intensity topics (8.23%–11.06%) centred on technology-based learning. Digitalisation in distance education (11.06%) was studied most, followed by Web-based learning (10.96%), social interaction and collaboration (10.61%), and artificial intelligence (10.34%). Technology-based education (8.24%) and course and programme (8.23%) also drew strong interest.

Medium-intensity topics (1.56%–6.06%) focused on technological applications and pedagogy. While system and infrastructure (1.56%) and research and method (1.75%) were studied less, problem-based

learning (6.06%), motivation and success (5.73%), and health education (5.57%) received moderate attention, showing the importance of methodological and assessment approaches.

Low-intensity topics (0.79%–1.28%) included learning environment (0.79%), MOOC (1.08%), and affective skills (1.28%). Although limited, topics like affective skills became more visible post-pandemic. Overall, findings revealed the dominance of technology integration, while areas such as psychological health, access, and traditional engineering remained less explored. After COVID-19, distance education and technology-based learning intensified significantly.

Temporal Analysis of Topics Explored

This section reports annual topic volume percentages and momentum patterns in distance education research. Figure 9 summarises temporal changes in topic prominence using a colour-coded heat map, where green indicates above-average and red below-average values.

Figure 9

Percentage of Volume and Acceleration Trends of Subjects by Year

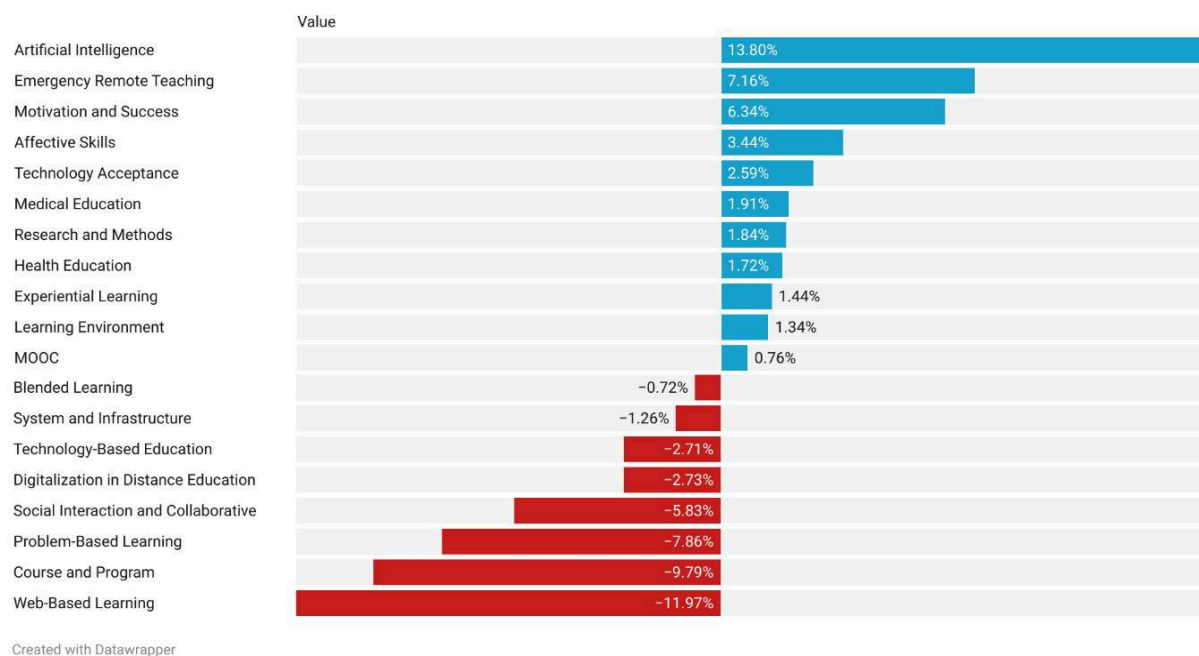
Topic Name / Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	Trend
Artificial Intelligence	3.58	6.07	5.32	5.71	6.54	6.76	6.51	8.82	8.86	9.82	11.42	11.04	13.68	16.43	18.77	13.86	10.40	12.19	13.58	17.38	13.80
Emergency Remote Teaching	0.53	0.23	0.28	0.31	0.35	0.14	0.37	0.23	0.11	0.15	0.35	0.10	0.23	0.13	0.24	11.61	15.46	14.33	11.23	7.69	7.16
Motivation and Success	3.58	3.27	3.54	3.60	4.71	4.80	5.47	5.50	5.44	6.53	5.61	7.12	5.28	6.24	5.97	5.48	6.19	7.65	8.62	9.92	6.34
Affective Skills	0.27	0.35	0.56	0.55	0.21	0.41	0.18	0.57	0.50	0.71	0.35	0.57	0.32	0.45	0.61	1.96	4.14	4.85	4.39	3.71	3.44
Technology Acceptance	2.39	1.17	2.33	2.66	2.11	3.04	3.81	3.78	3.31	2.43	3.31	2.72	3.67	3.59	4.39	3.49	4.02	5.01	5.12	4.98	2.59
Medical Education Research	2.39	4.09	2.89	2.82	3.59	3.45	3.75	3.15	3.70	4.15	4.66	4.16	4.13	4.18	4.10	4.46	6.13	5.36	4.64	4.30	1.91
Research and Methods	1.19	0.58	0.75	0.86	1.48	1.15	1.66	1.37	1.74	1.62	1.20	1.53	1.79	1.80	1.99	2.04	2.98	2.91	3.25	3.03	1.84
Health Education	5.04	3.15	4.20	4.62	4.15	4.33	5.28	6.24	4.66	6.33	6.91	6.50	7.16	7.41	7.11	6.24	5.00	4.90	5.44	6.76	1.72
Experiential Learning	1.46	1.87	1.21	1.96	1.41	1.56	1.90	2.00	2.02	1.77	2.91	3.11	2.75	2.60	2.97	2.06	1.85	2.21	2.63	2.90	1.44
Learning Environment	0.53	0.35	0.56	0.86	0.28	0.41	0.68	0.46	0.22	0.81	0.60	0.57	0.28	0.54	0.69	0.82	1.91	1.59	1.75	1.87	1.34
MOOC	0.53	0.47	0.75	1.02	0.49	1.49	0.80	1.09	0.73	0.76	1.05	1.15	1.42	1.62	0.97	1.04	1.30	1.70	1.90	1.29	0.76
Blended Learning	3.45	2.45	3.26	4.54	3.38	3.25	3.87	3.38	4.88	3.54	4.21	4.54	4.41	4.31	3.82	3.95	3.06	2.87	2.68	2.73	-0.72
System and Infrastructure	2.12	1.87	1.59	2.66	2.39	1.76	2.27	1.45	2.24	1.27	1.20	2.10	1.38	1.12	1.10	0.94	1.01	1.03	0.94	0.86	-1.26
Technology-Based Education	12.20	9.00	8.86	9.78	9.01	8.25	7.37	9.28	7.85	8.76	8.37	7.50	8.67	7.14	8.16	6.37	6.07	5.77	6.80	9.49	-2.71
Digitalization in Distance Education	9.02	9.35	10.91	11.74	11.05	13.12	12.84	10.48	10.66	12.76	10.22	10.70	13.17	12.08	10.60	12.61	12.53	11.22	9.75	6.29	-2.73
Social Interaction and Collaboration	13.13	12.85	13.15	10.88	13.23	12.78	12.78	12.83	13.24	11.44	11.47	11.90	9.78	9.30	8.08	8.46	6.41	6.15	7.08	7.30	-5.83
Problem-Based Learning	9.81	10.28	8.68	8.29	8.73	8.45	7.86	6.64	6.56	6.03	6.16	6.16	4.77	4.45	4.96	3.29	2.97	2.72	2.44	1.95	-7.86
Course and Program	12.73	13.90	13.15	10.56	10.84	9.60	9.09	9.28	9.98	9.37	8.42	8.22	6.84	7.23	7.11	5.15	3.67	3.08	3.38	2.94	-9.79
Web-Based Learning	16.05	18.11	17.82	16.35	15.62	14.74	13.33	13.00	12.90	10.94	11.02	9.94	9.73	8.89	7.92	5.88	4.70	4.19	3.89	4.08	-11.97

According to Figure 9, some topics have shown a strong upward trend in recent years. Artificial intelligence (13.80%) stood out with the highest growth, reaching 17.38% in 2024, highlighting its expanding role in distance education. Emergency remote teaching (7.16%) also rose rapidly during the COVID-19 pandemic and has remained influential. Interest in motivation and success (6.34%), affective skills (3.44%), and technology acceptance (2.59%) has increased, while health sector applications expanded significantly post-pandemic. More moderate growth has been seen in medical education research (1.91%), research and methods (1.84%), experiential learning (1.44%), and learning environment (1.34%).

Conversely, several topics have declined. Web-based learning (-11.97%) showed the sharpest drop, steadily falling until 2024. Once a leading area, it has become one of the lowest growth fields. Social interaction and collaboration (-5.83%), problem-based learning (-7.86%), and course and programme (-9.79%) have also decreased. Analyses indicated that while these topics dominated two decades ago, they have largely lost prominence since then. The topics with the highest and lowest momentum are shown in Figure 10.

Figure 10

Acceleration of Issues Between 2005 and 2024



To summarise, Figure 10 illustrates how research trends in distance education have shifted over the years. Topics such as machine learning, data analytics, emergency remote teaching, and success and motivation have become increasingly prominent, while the decline of topics like Web-based learning, once highly important, revealed changing research priorities.

Discussion

This study aimed to reveal research trends in distance education over the last two decades using a combined bibliometric and machine learning-based approach. In this context, 54,444 articles were identified in the

Web of Science database through the specified keywords. The findings from both the bibliometric analysis and the LDA-based topic modeling of this dataset are discussed in detail in this section in line with the related literature and research questions.

Bibliometric Analysis

In the research, it was determined that the number of academic studies on distance education has shown a continuous increase, reaching its peak in 2022. During the COVID-19 pandemic, this growth accelerated. The main reason has been the integration of high-potential technologies such as the Internet into educational environments (Whalley et al., 2021), which paralleled the increase in distance education activities. During the pandemic, the fact that most educational activities were carried out remotely under social distancing measures (Bozkurt & Sharma, 2020) further increased research interest (Mishra et al., 2021). Similar trends have been reported in Bozkurt et al. (2015) and Yılmaz Özden et al. (2023).

Another bibliometric finding concerned the subject areas where studies were concentrated. Distance education has been researched in education, management, nursing, engineering, and language. The adoption of digital technologies such as simulation, augmented reality, and virtual reality in health, engineering, and environmental sciences, and especially in education and computer sciences, has increased research in these disciplines (Blyznyuk et al., 2021). Moreover, opportunities such as flexibility, interaction, communication, and collaboration offered by distance education (Henderson et al., 2017) have increased the desirability of DE (Kanishcheva, 2022). Similar trends were reported by Amoozegar et al. (2018) and Gürcan et al. (2021).

Self-regulated learning, technology acceptance model, medical education, learning styles, and language policy were the sub-topics most often researched. In distance education, self-regulated learning is crucial as students must manage their learning processes independently (Bozkurt & Zawacki-Richter, 2021). The adoption of systems also depends on how students and teachers perceive technology; hence, the technology acceptance model has been studied frequently (Zawacki-Richter et al., 2009). Distance education also offers a different structure for medical students, which has led to increased research on its effectiveness (Djeki et al., 2022). Studies focused on self-regulated learning, technology acceptance, and medical education stood out (Teo, 2019).

Finally, bibliometric data showed that academic publications on distance education were concentrated in technology- and education-oriented journals. The *Turkish Online Journal of Distance Education*, *Education and Information Technologies*, *Computers & Education*, *International Journal of Emerging Technologies in Learning*, and *Sustainability* were prominent publication platforms. In particular, *Computers & Education*, *Educational Technology & Society*, and *Distance Education* were prominent as pioneering journals (Bozkurt, 2020). Similar emphasis was also noted in Amoozegar et al. (2018), Djeki et al. (2022), and Gürcan et al. (2021).

Topic Modeling Analysis in Distance Education

Within the scope of the second research question, LDA-based topic analysis was performed on studies in distance education. As a result, 19 topics were identified and classified into low, medium, and high intensity. Among them, digitalisation in distance education, social interaction and collaboration, Web-based learning,

and artificial intelligence had the highest intensity. In recent years, distance education expanded rapidly with technological developments, and during the COVID-19 pandemic, both digitalisation processes and the transition to online learning accelerated globally (Bozkurt & Zawacki-Richter, 2021; Zawacki-Richter et al., 2020). Today, distance education has become mainstream, enabling students to access learning opportunities without time and space constraints (Hodges et al., 2020).

Digitalisation has reduced costs and allowed students to customise learning according to personal needs (Zawacki-Richter et al., 2020). Online platforms have integrated group work, peer assessments, and forum discussions to enhance interaction (Anderson, 2008). Innovative technologies such as AI, virtual reality, and augmented reality have served to simulate interactive experiences (Selwyn, 2019). AI in particular has supported equal opportunities in large-scale programmes through personalised learning, instant feedback, and data-driven decisions (Siemens & Long, 2011). These developments have created a paradigm shift in educational design (Zawacki-Richter et al., 2020). Accordingly, research has increasingly focused on digitalisation, AI, and interaction, as exemplified by the following studies:

- Digitalisation in distance education (Bozkurt & Zawacki-Richter, 2021; Bozkurt et al., 2015, 2019; Hakan et al., 2020; Zawacki-Richter et al., 2020)
- Social interaction and collaboration (Zawacki-Richter & Naidu, 2016)
- Web-based learning (Bozkurt et al., 2015; Zawacki-Richter & Naidu, 2016)
- Artificial intelligence (Karagözoğlu et al., 2024).

The medium-intensity group included studies on pedagogical models and teaching strategies to make online learning more effective. Problem-based learning has been widely applied to develop critical thinking and active learning in online environments (Savery, 2009). Similarly, blended learning, combining face-to-face and online education, became prominent, especially after the pandemic, supported by digitalisation infrastructure. It provides flexibility and increases satisfaction (Means et al., 2013). Distance education in clinical and medical training also gained importance, with a major transformation in health education (Ellaway, 2018). The effectiveness of clinical distance education and its impact on patient care remain subjects of ongoing research (Huang, 2021).

The low-intensity group included learning environment, MOOCs, and affective skills. MOOCs, described as a revolution in education between 2012–2015 (Daniel, 2012), initially drew interest as they expanded access and promoted equality of opportunity (Veletsianos, 2020). However, commercialisation, platform-centred structures, limited interaction, and access issues reduced interest (Cornali, 2024). Moreover, the recent research shift toward pandemic and AI topics contributed to the decline (Kurulgan, 2024).

Temporal Analysis of Topics Explored

Clearly, trends in the distance education areas being researched have changed significantly over the years. Topics such as machine learning, data analytics, emergency remote teaching (ERT), as well as motivation and success, have become increasingly prominent in recent years. On the other hand, Web-based learning,

interaction and collaboration, and course and programme, important in the past, have garnered less interest, indicating that research priorities have changed.

The contributions of AI-supported systems, such as providing individualised learning experiences, adapting content to students' needs, offering real-time feedback, and encouraging active participation in distance learning (Al-Tkhayneh et al., 2023), have revealed that AI will be increasingly used in educational environments. This finding reflects global trends towards the integration of technological transformation into education.

During the COVID-19 pandemic, distance education ensured continuity in learning and minimised losses (Yavuz et al., 2020). At almost all levels, educational activities were carried out mainly through online learning (Can, 2020). This shifted the focus of studies during and after the pandemic to temporary applications such as ERT (Hodges et al., 2020). The increase in research on motivation and success was based on two factors. First, rapid technological developments enabled online platforms and personalised learning tools, making learning more effective (Lagos-Castillo et al., 2025). Second, physical restrictions during the COVID-19 pandemic made distance education inevitable; in this context, motivation and achievement problems became central (Martin, 2023). AI-supported learning processes have attracted strong interest, especially for their potential to personalise education and increase student success (Zawacki-Richter et al., 2020). According to Teo (2019), research on motivation and achievement has increased. Similarly, bibliometric studies by Kurulgan (2024) and Karagözoğlu et al. (2024) stated that research addressing ERT, learning analytics, AI applications, and motivation has increased.

While Web-based learning initially occupied an important place in the distance education literature, its decline has shown that both research priorities and technological integration have changed (Hrastinski, 2008; Siemens, 2005). This decline was associated with the prominence of personalisation, interactivity and advanced analytical approaches; traditional Web-based methods have been replaced by more dynamic and adaptive technologies. Similar results were reported in the bibliometric analyses by Kurulgan (2024) and Karagözoğlu et al. (2024).

Conclusions and Recommendations

This research analysed studies in distance education over the last 20 years using a machine learning-based method and comprehensively evaluated trends in the field. The findings reveal the diversity of topics, their temporal changes, and the intensity of their application across disciplines.

High-intensity topics such as digitalization in distance education, social interaction and collaboration, Web-based learning, and artificial intelligence were most prominent, highlighting the growing importance of digitalisation, interaction-oriented methods, and AI applications. The COVID-19 pandemic accelerated this transformation and increased academic interest. In addition, topics like self-regulated learning and technology acceptance underline the significance of students' ability to manage learning and adopt technology in distance education.

Temporal analysis shows that while innovative areas such as artificial intelligence have gained popularity, traditional topics like Web-based learning have declined, reflecting the evolution of distance education in line with advancing technologies and changing needs.

Beyond a general focus on distance or online learning, these findings contribute directly to ODL by showing thematic shifts toward artificial intelligence, learning analytics, and affective dimensions, which relate to scalability, learner autonomy, openness, and sustainability. The taxonomy enables a theoretical reinterpretation of established ODL frameworks by highlighting how interaction, learner support, and openness are reshaped in digitally distributed environments.

Practically, the taxonomy offers guidance for program design and policy. For example, the prominence of artificial Intelligence and learning analytics suggests incorporating AI literacy, data-informed design, and ethical AI into curricula, while the decline of Web-based learning points to more adaptive, personalized models. Additionally, attention to affective skills and motivation and success can inform research priorities and funding decisions, emphasizing socio-emotional dimensions of distance education in post-pandemic contexts.

Based on the findings, recommendations are offered for researchers and policy makers. For researchers, priority should be given to studies on innovative technologies, especially AI and learning analytics, which can improve achievement, personalise learning, and enhance materials. Furthermore, qualified studies are needed on the social, psychological, and cultural effects of distance education in the post-pandemic period, particularly focusing on motivation, self-regulation, and student–teacher interactions.

For decision makers, distance education programmes should prioritise accessibility through infrastructure investments. The dissemination of personalised learning models is recommended to increase satisfaction and success, while AI-supported systems, with their potential for customised content and instant feedback, should be integrated into distance education processes.

Declaration on Generative Artificial Intelligence

The authors disclose that they employed DeepL, ChatGPT 5, and ChatGPT 5.2 (August, 2025/January, 2026) to translate and localize content. Human translators subsequently reviewed and adjusted the translations to ensure accuracy, cultural appropriateness, and contextual relevance.

References

- Aggarwal, C. C., & Zhai, C. (2012). A survey of text classification algorithms. In C. C. Aggarwal & C. Zhai (Eds.), *Mining text data* (pp. 163–222). Springer. https://doi.org/10.1007/978-1-4614-3223-4_6
- Al-Tkhayneh, K. M., Alghazo, E. M., & Tahat, D. (2023). The advantages and disadvantages of using artificial intelligence in education. *Journal of Educational and Social Research*, 13(4), 105–116. <https://doi.org/10.36941/jesr-2023-0094>
- Amoozegar, A., Khodabandelou, R., & Ale Ebrahim, N. (2018). Major trends in distance education research: A combination of bibliometric and thematic analysis. *International Journal of Information Research and Review*, 5(2), 5352–5359. <https://doi.org/10.6084/m9.figshare.6210536>
- Anderson, T. (Ed.). (2008). *The theory and practice of online learning* (2nd ed.). AU Press. <https://doi.org/10.15215/aupress/9781897425084.01>
- Aydın, C. H., Zawacki-Richter, O., & Bozkurt, A. (2020, June). A review and content analysis of the *Turkish Online Journal of Distance Education* publications between 2000 and 2015. *EDEN conference proceedings*, 1, (pp. 217–225). <https://doi.org/10.38069/edenconf-2020-ac0020>
- Aziz, S., Dowling, M., Hammami, H., & Piepenbrink, A. (2022). Machine learning in finance: A topic modeling approach. *European Financial Management*, 28(3), 744–770. <https://doi.org/10.1111/eufm.12326>
- Bates, T. (2014). *The role of technology in distance education*. Taylor & Francis. <https://doi.org/10.4324/9781315738598>
- Berge, Z. L., & Mrozowski, S. (2001). Review of research in distance education, 1990 to 1999. *American Journal of Distance Education*, 15(3), 5–19. <https://doi.org/10.1080/08923640109527090>
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the Association for Computing Machinery*, 55(4), 77–84. <https://doi.org/10.1145/2133806.2133826>
- Blyznyuk, T., Budnyk, O., & Kachak, T. (2021). Boom in distance learning during the coronavirus pandemic: Challenges and possibilities. *Journal of Vasyl Stefanyk Precarpathian National University*, 8(1), 90–98. <https://doi.org/10.15330/jpnu.8.1.90-98>
- Bozkurt, A. (2020). Coronavirus (COVID-19) pandemic process and education in the world after the pandemic evaluations: The new normal and the new educational paradigm. *Journal of Open Education Applications and Research*, 6(3), 112–142. <https://izlik.org/JA99ML86GW>
- Bozkurt, A., Kır, Ş., Şenocak, D., Elibol, S., Özer, N., Orhan, G., Alpat, M. F., Çimen, E. E., Çankırlı, E., & İbileme, A. İ. (2022). Research tendencies in the discipline of distance education (2015–2022): Examination of doctoral theses in higher education in Turkey. *Proceedings of the Fifth*

- International Open & Distance Learning Conference* (pp. 355–366).
<https://doi.org/10.5281/zenodo.7428522>
- Bozkurt, A., Kumtepe, E. G., Kumtepe, A. T., Aydın, İ. E., Bozkaya, M., & Aydın, C. H. (2015). Research trends in Turkish distance education: A content analysis of dissertations, 1986–2014. *European Journal of Open, Distance and E-learning*, 18(2), 1–21. <https://doi.org/10.1515/eurodl-2015-0010>
- Bozkurt, A., & Sharma, R. C. (2020). Emergency remote teaching in a time of global crisis due to coronavirus pandemic. *American Journal of Distance Education*, 15(1), i–vi.
<https://doi.org/10.5281/zenodo.3778083>
- Bozkurt, A., & Zawacki-Richter, O. (2021). Trends and patterns in distance education (2014–2019): A synthesis of scholarly publications and a visualization of the intellectual landscape. *International Review of Research in Open and Distributed Learning*, 22(2), 19–45.
<https://doi.org/10.19173/irrodl.v22i2.5381>
- Bozkurt, A., Zawacki-Richter, O., & Aydın, C. H. (2019). Using social network analysis to review the research in open and distance learning. *Proceedings of the Association for Educational Communications and Technology 2019 International Convention* (pp. 38–44). Las Vegas, NV, United States.
- Can, E. (2020). Coronavirus (COVID-19) pandemic and its pedagogical reflections: Open and distance education practices in Turkey. *Journal of Open Education Applications and Research*, 6(2), 11–53. <https://izlik.org/JA32KL96LX>
- Charan, J. (2014). Impact factor: Is this a true measure of quality? *International Journal of Medical Science and Public Health*, 3(3), 246–247. <https://doi.org/10.5455/ijmsph.2014.190420141>
- Cornali, F. (2024). In platform we trust(?): Concepts, problems and a case study about platform education. *European Journal of Education*, 59(4), e12702. <https://doi.org/10.1111/ejed.12702>
- Daniel, J. (2012). Making sense of MOOCs: Musings in a maze of myth, paradox and possibility. *Journal of Interactive Media in Education*, 2012(3), 18. <https://doi.org/10.5334/2012-18>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Debortoli, S., Müller, O., Junglas, I., & Vom Brocke, J. (2016). Text mining for information systems researchers: An annotated topic modeling tutorial. *Communications of the Association for Information Systems*, 39(1), 7. <https://doi.org/10.17705/1CAIS.03907>
- Delling, R. M. (1971). Grundzüge einer Wissenschaft vom Fernstudium [Fundamentals of the science of distance learning]. *Epistologidaktika*, 1, 14–28.

- Djeki, E., Dégila, J., Bondiombouy, C., & Alhassan, M. H. (2022). E-learning bibliometric analysis from 2015 to 2020. *Journal of Computers in Education*, 9(4), 727–754. <https://doi.org/10.1007/s40692-021-00218-4>
- Ellaway, R. H. (2018). Technology-enhanced learning. In T. Swanwick (Ed.), *Understanding medical education: Evidence, theory, and practice* (2nd ed., pp. 139–149). Wiley-Blackwell. <https://doi.org/10.1002/9781119373780.ch10>
- Fu, H., Xiao, Y., Mensah, I. K., & Wang, R. (2024). Exploring the configurations of learner satisfaction with MOOCs designed for computer science courses based on integrated LDA-QCA method. *Education and Information Technologies*, 29(8), 9883–9905. <https://doi.org/10.1007/s10639-023-12185-7>
- 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](https://doi.org/10.1016/S1096-7516(00)00016-6)
- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(1), 5228–5235. <https://doi.org/10.1073/pnas.0307752101>
- Gunawardena, C. N., & McIsaac, M. S. (2013). Distance education. In Jonassen, D., & Driscoll, M. (Eds.), *Handbook of research on educational communications and technology* (pp. 361–401). Routledge. <https://doi.org/10.4324/9781410609519>
- Gurcan, F., Özyurt, Ö., & Çağıltay, N. E. (2021). Investigation of emerging trends in the e-learning field using latent Dirichlet allocation. *International Review of Research in Open and Distributed Learning*, 22(2), 1–18. <https://doi.org/10.19173/irrodl.v22i2.5358>
- Hakan, C., Zawacki-Richter, O., & Bozkurt, A. (2020, June). A review and content analysis of the Turkish online journal of distance education publications between 2000 and 2015. In *EDEN Conference Proceedings*, 1, (pp. 217–225). <https://doi.org/10.38069/edenconf-2020-aco020>
- Hamal, O., El Faddouli, N. E., Harouni, M. H. A., & Lu, J. (2022). Artificial intelligence in education. *Sustainability*, 14(5), 2862. <https://doi.org/10.3390/su14052862>
- Henderson, M., Selwyn, N., & Aston, R. (2017). What works and why? Student perceptions of “useful” digital technology in university teaching and learning. *Studies in Higher Education*, 42(8), 1567–1579. <https://doi.org/10.1080/03075079.2015.1007946>
- Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. *EDUCAUSE Review*, 27(1), 1–9. <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>

- Howell, S. L., Williams, P. B., & Lindsay, N. K. (2003). Thirty-two trends affecting distance education: An informed foundation for strategic planning. *Online Journal of Distance Learning Administration*, 6(3). <https://www.westga.edu/~distance/ojdla/fall63/howell63.html>
- Hrastinski, S. (2008). What is online learner participation? A literature review. *Computers & Education*, 51(4), 1755–1765. <https://doi.org/10.1016/j.compedu.2008.05.005>
- Huang, Y. (2021). Strategies for enhancing student interaction in an online health promotion course. *College Student Journal*, 55(1), 1–6. <https://eric.ed.gov/?id=EJ1313964>
- Kaldoudi, E., Konstantinidis, S., & Bamidis, P. D. (2010). Web advances in education: Interactive, collaborative learning via Web 2.0. In M. D. Lytras, P. Ordóñez de Pablos, A. Zideman, E. Damiani, D. Avison, & J. B. E. M. L. Horner (Eds.), *Affective, interactive and cognitive methods for e-learning design: Creating an optimal education experience* (pp. 32–50). IGI Global. <https://doi.org/10.4018/978-1-60566-940-3.ch003>
- Kanishcheva, N. A. (2022, February). The role of distance learning in the development of the digital economy. *International Scientific and Practical Conference Strategy of Development of Regional Ecosystems* (pp. 198–203). Atlantis Press. <https://doi.org/10.2991/assehr.k.220208.029>
- Karagözoğlu, A. A., Abdurrezzak, S., & Doğan, Ü. (2024). A bibliometric analysis of studies on distance education. *Anadolu Journal of Educational Sciences International*, 14(1), 449–474. <https://doi.org/10.18039/ajesi.1281711>
- Karl, A., Wisnowski, J., & Rushing, W. H. (2015). A practical guide to text mining with topic extraction. *Wiley Interdisciplinary Reviews: Computational Statistics*, 7(5), 326–340. <https://doi.org/10.1002/wics.1361>
- Konrad, M. (2017). *Text mining and topic modeling toolkit*. PyPI. <https://pypi.org/project/tmtoolkit/>
- Kurulgan, M. (2024). A bibliometric analysis of research on dropout in open and distance learning. *Turkish Online Journal of Distance Education*, 25(4), 200–229. <https://doi.org/10.17718/tojde.1355394>
- Lagos-Castillo, A., Chiappe, A., Ramirez-Montoya, M. S., & Becerra Rodríguez, D. F. (2025). Mapping the intelligent classroom: Examining the emergence of personalized learning solutions in the digital age. *Contemporary Educational Technology*, 17(1), ep543. <https://doi.org/10.30935/cedtech/15617>
- Li, C., Wang, H., Zhang, Z., Sun, A., & Ma, Z. (2016, July). Topic modeling for short texts with auxiliary word embeddings. *Proceedings of the 39th International Conference on Research and Development in Information Retrieval* (pp. 165–174). Association for Computing Machinery. <https://doi.org/10.1145/2911451.2911499>

- Martin, A. J. (2023). University students' motivation and engagement during the COVID-19 pandemic: The roles of lockdown, isolation, and remote and hybrid learning. *Australian Journal of Education*, 67(2), 163–180. <https://doi.org/10.1177/00049441231179791>
- Means, B., Toyama, Y., Murphy, R., & Baki, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers College Record*, 115(3), 1–47. <https://doi.org/10.1177/016146811311500307>
- Mishra, S., Sahoo, S., & Pandey, S. (2021). Research trends in online distance learning during the COVID-19 pandemic. *Distance Education*, 42(4), 494–519. <https://doi.org/10.1080/01587919.2021.1986373>
- Moore, M. G., & Kearsley, G. (2012). *Distance education: A systems view of online learning*. Cengage.
- National Forum. (2021). *Next steps for teaching and learning: Moving forward together. Key themes in online and blended learning, 2010–2020*. <https://www.teachingandlearning.ie/wp-content/uploads/Key-themes-in-online-and-blended-learning-2010-2020.pdf>
- Oviatt, D., Burdis, J., & West, R. E. (2012). Distance education, 2000–2010. *Educational Technology*, 52(1), 44–48. <http://www.jstor.org/stable/44429990>
- Özmen, E., & Kan, A. Ü. (2021). Bibliometric analysis of theses on distance education in Turkey between 2015–2020. *Turkish Studies Educational Sciences*, 16(4), 2005–2027. <https://doi.org/10.47423/TurkishStudies.51788>
- Pai, K.-C., Kuo, B.-C., Liao, C.-H., & Liu, Y.-M. (2021). An application of Chinese dialogue-based intelligent tutoring system in remedial instruction for mathematics learning. *Educational Psychology*, 41(2), 137–152. <https://doi.org/10.1080/01443410.2020.1731427>
- Park, S. Y., Nam, M.-W., & Cha, S.-B. (2012). University students' behavioral intention to use mobile learning: Evaluating the technology acceptance model. *British Journal of Educational Technology*, 43(4), 592–605. <https://doi.org/10.1111/j.1467-8535.2011.01229.x>
- Peña-Ayala, A. (Ed.). (2014). *Educational data mining: Applications and trends*. Springer. <https://doi.org/10.1007/978-3-319-02738-8>
- Peters, O. (2014). Foreword. In O. Zawacki-Richter & T. Anderson (Eds.), *Online distance education: Towards a research agenda* (pp. ix–xii). Athabasca University Press. <http://dx.doi.org/10.15215/aupress/9781927356623.01>
- Saba, F. (2000). Research in distance education: A status report. *International Review of Research in Open and Distributed Learning*, 1(1), 1–9. <https://doi.org/10.19173/irrodl.v1i1.4>

- Savery, J. R. (2009). Problem-based approach to instruction. In C. M. Reigeluth & A. A. Carr-Chellman (Eds.), *Instructional-design theories and models: Building a common knowledge base* (Vol. III, pp. 155–178). Routledge.
- Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. John Wiley & Sons.
<https://doi.org/10.1002/9781119340829>
- Siemens, G. (2005). Connectivism: A learning theory for the digital age. *Elearnspace*.
<http://www.elearnspace.org/Articles/connectivism.htm>
- Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), 30–40. <https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education>
- Teo, T. (2019). Students and teachers' intention to use technology: Assessing their measurement equivalence and structural invariance. *Journal of Educational Computing Research*, 57(1), 201–225. <https://doi.org/10.1177/0735633117749430>
- Veletsianos, G. (2020). *Learning online: The student experience*. Johns Hopkins University Press.
- Wallach, H. M. (2006, June). Topic modeling: Beyond bag-of-words. *Proceedings of the 23rd International Conference on Machine Learning* (pp. 977–984). Association for Computing Machinery. <https://doi.org/10.1145/1143844.1143967>
- Watermeyer, R., Crick, T., Knight, C., & Goodall, J. (2021). COVID-19 and digital disruption in UK universities: Afflictions and affordances of emergency online migration. *Higher Education*, 81(3), 623–641. <https://doi.org/10.1007/s10734-020-00561-y>
- Whalley, B., France, D., Park, J., Mauchline, A., & Welsh, K. (2021). Towards flexible personalized learning and the future educational system in the fourth industrial revolution in the wake of COVID-19. *Higher Education Pedagogies*, 6(1), 79–99.
<https://doi.org/10.1080/23752696.2021.1883458>
- Xiao, J. (2023). Critical issues in open and distance education research. *International Review of Research in Open and Distributed Learning*, 24(2), 213–228. <https://doi.org/10.19173/irrodl.v24i2.6881>
- Yavuz, M., Kayalı, B., Balat, Ş., & Karaman, S. (2020). Review of distance learning applications in the universities in the COVID-19 period. *Journal of National Education*, 49(1), 129–154.
<https://doi.org/10.37669/milliegitim.784822>
- Yılmaz Özden, Ş., Hamutoğlu, N. B., Kıyıcı, M., & Bozkurt, A. (2023). Investigation of research trends in educational technologies based on scientometric method (1965–2020): A cross-comparative study between publications from the world and the UK. *Educational Technology Research and Development*, 71(4), 1421–1447. <https://doi.org/10.1007/s11423-023-10224-1>

- Yörük, T. (2024). Artificial intelligence and personalised learning in education. In S. Karataş (Ed.), *International research in the field of educational sciences XXIII*, (pp. 21–34). Eğitim Publisher.
- Zankadi, H., Idrissi, A., Daoudi, N., & Hilal, I. (2023). Identifying learners' topical interests from social media content to enrich their course preferences in MOOCs using topic modeling and NLP techniques. *Education and Information Technologies*, 28(5), 5567–5584.
<https://doi.org/10.1007/s10639-022-11373-1>
- Zawacki-Richter, O., Bäcker, E. M., & Vogt, S. (2009). Review of distance education research (2000 to 2008): Analysis of research areas, methods, and authorship patterns. *International Review of Research in Open and Distributed Learning*, 10(6), 21–50.
<https://doi.org/10.19173/irrodl.v10i6.741>
- Zawacki-Richter, O., Marin, V. I., Bond, M., & Gouverneur, F. (2020). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 17(1), 1–27. <https://doi.org/10.1186/s41239-019-0171-0>
- 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>
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2



August – 2026

Systemic Vulnerabilities in Online Education: Diagnosing and Meeting K–12 Teacher Preparation Needs

Barış Avcı and Meral Güven
Anadolu University, Eskişehir, Türkiye

Abstract

This mixed-methods study developed and evaluated a comprehensive 5-week asynchronous in-service training program to enhance K–12 teachers' online education competencies across pedagogical, social, managerial, and technical domains. Following needs assessment, 21 teachers in Türkiye participated using the Canvas learning management system (LMS). The competency-based program emphasized authentic learning experiences, requiring participants to design online lesson plans, use Web 2.0 tools, and implement the LMS in classroom contexts. Data collection employed triangulated methods including learning analytics and products, pre-/post-test assessments, and semi-structured interviews. Results demonstrated significant improvement in post-test scores ($p < .05$) and high engagement through practical activities. Learning analytics revealed varied engagement patterns, with participants completing authentic assignments. Qualitative findings highlighted gains in content development, instructional design skills, and digital competence. The study demonstrates that teachers require experiential learning as online students before becoming effective online instructors, emphasizing the necessity of systematic, needs-oriented training programs integrating theory with hands-on practice for comprehensive pedagogical transformation.

Keywords: online education, teacher training, curriculum development, K–12

Introduction

Online education, defined as learning and teaching activities occurring synchronously or asynchronously in shared digital environments regardless of physical location (Mackay, 2013), has evolved from a supplementary educational tool to a central component of contemporary schooling (Zuo et al., 2024). This technological transformation has revealed a critical disconnect wherein teachers trained exclusively in conventional classroom settings frequently lack both personal experience with online education environments and the specialized competencies required for effective digital instruction (Avcı & Güven, 2025). This disconnect represents more than a simple skills gap; it constitutes a fundamental pedagogical crisis that undermines educational quality across global educational systems.

This disconnect is exacerbated by the fact that professional development approaches remain mainly reactive rather than proactive, with teachers receiving training only after beginning online instruction rather than during initial preparation (Kamal & Illiyan, 2021). Compounding this problem, a significant research–practice divide persists, as academic researchers have traditionally not prioritized solving K–12 classroom problems (Larson & Archambault, 2019). These parallel failures create a particularly destructive feedback loop: Without research addressing practical K–12 needs, teachers must independently restructure their approaches to accommodate field-specific constraints, including program standards, course durations, and cultural terminology unique to K–12 settings (Lokey-Vega et al., 2018; Orakcioğlu, 2019). This fragmentation both perpetuates inadequate preparation and prevents the development of evidence-based practices, ensuring that each new cohort of online teachers faces the same foundational challenges as their predecessors.

The COVID-19 pandemic served as a devastating proof of concept for these systemic vulnerabilities, transforming theoretical concerns about teacher readiness into observable educational failures at scale. The crisis revealed not merely that teachers lacked specific technical skills but that the entire infrastructure for preparing and supporting online teachers was fundamentally inadequate.

Beyond Skills Training: The Experiential Foundation of Online Teaching

Effective online education demands mastery of interconnected competencies spanning technology-enhanced course design, digital content delivery, and multimodal communication with students and parents through both real-time and asynchronous platforms. However, a framing of this challenge as simply a matter of acquiring new skills fundamentally misunderstands the depth of transformation required. As Koenig (2020, para. 6) bluntly observed, “most teachers simply don’t know how to teach online. No one ever taught them how or asked them to learn.”

The problem extends beyond what teachers know to encompass what they have experienced. Research consistently demonstrates that teachers’ own experiences as online learners constitute a foundational prerequisite for developing effective online pedagogical practices (Leary et al., 2020). This experiential knowledge enables teachers to internalize the constraints, challenges, and opportunities inherent in online contexts—insights that cannot be adequately conveyed through theoretical instruction alone. The absence of this experiential foundation helps explain why pedagogical strategies effective in traditional classrooms often fail when transferred to digital environments (König et al., 2020), a phenomenon the pandemic brought into sharp relief.

Most teachers, however, enter online teaching without this critical experiential base, beginning their digital instruction careers through institutional mandate rather than deliberate professional development pathways (Tschida et al., 2017). This gap represents a systemic institutional inadequacy; teachers are being asked to guide students through learning experiences they themselves have never navigated.

These global challenges are exemplified in Türkiye's educational system, which offers a revealing case study of how systemic failures translate into tangible educational outcomes. Turkish teachers are not required to undergo specialized training for digital instruction (Bakioğlu & Çevik, 2020), revealing a critical gap in professional readiness. The consequences of this deficiency became evident during emergency remote instruction, when traditional pedagogical methods were precipitously transposed into online environments without adaptation, resulting in substantial learning deficits (Kaptanoğlu, 2021). A comprehensive analysis conducted by TEDMEM (İpek et al., 2020) further confirmed that Turkish teachers demonstrate general unpreparedness for online education delivery.

The inadequacy of teacher preparation extends beyond preservice education to in-service training initiatives. Additionally, an evaluation of in-service teacher training programs administered by the Turkish Ministry of National Education (MoNE), encompassing 1,080 courses and 174 seminars, revealed that only a minimal proportion specifically addressed online teaching competencies (Avcı & Güven, 2021). This finding aligns with broader evidence from literature reviews, university course catalogue analyses, and national education assessments, which consistently highlight pervasive inadequacy in teacher preparation for online educational contexts.

Implications and the Need for Reform

Contemporary teacher education programs demonstrably lack capacity to develop essential competencies for online instruction (Sengil-Akar & Kurtoglu-Erden, 2021), making the development of appropriate competencies an imperative rather than an option as online education continues redefining teacher roles (Basilotta-Gómez-Pablos et al., 2022).

Effective training programs grounded in systematic, evidence-based identification of essential competencies represent the cornerstone of transformation (Brenner, 2022). Yet, despite increased availability of professional development opportunities, research has inadequately addressed the foundational questions that determine program effectiveness. Critical gaps persist regarding comprehensive needs assessments (Stavermann, 2025), precise competency definitions, appropriate proficiency benchmarks for online teachers (Tian & Tian, 2025), and practical implementation strategies. Perhaps most significantly, few studies employ triangulated methodologies that integrate quantitative behavioral indicators with qualitative feedback mechanisms (Fayda-Kınık, 2021) to comprehensively evaluate program impact.

These research gaps are not merely academic concerns; they directly impede the development of preparation programs capable of breaking the cycle of inadequacy that currently characterizes online teacher education. Without rigorous evidence about what works, why it works, and how to implement effective practices at scale, educational systems will continue producing teachers unprepared for the digital environments in which increasing numbers of students learn.

Conceptual Framework of Study

Study Design and Theoretical Framework

This study employed a multiphase mixed-methods design to implement and evaluate a comprehensive in-service training program aimed at enhancing K–12 teachers' online education competencies. A multiphase mixed-methods design focuses on addressing a series of interrelated research questions aimed at achieving a research objective and provides comprehensive answers to the problem by using multiple data collection methods (Creswell, 2012). It is considered the most appropriate method for use in curriculum development and evaluation studies (Nastasi et al., 2007).

The Development and Implementation Process

The in-service training program (ITP) theoretical foundation was established through a competency-based framework developed via extensive preliminary research conducted over 1 year prior to implementation. Recognizing the inadequate preparation many teachers face and the often-overlooked needs of K–12 online teaching in scholarly inquiry, the competency identification process employed a rigorous three-round Delphi study with six experts (four in open and distance education, two in curriculum and instruction). Beginning with an open-ended inquiry that generated 200 distinct items, the pool was meticulously refined to 157 items for subsequent rounds. Through systematic consensus-building using interquartile ranges, 41 competencies achieved unanimous expert agreement and were organized across Berge's (1995) framework: pedagogical (13), social (5), managerial (14), and technical (9). This foundational work ensured alignment between training objectives and the multifaceted demands and specific challenges encountered by K–12 teachers transitioning to online educational environments.

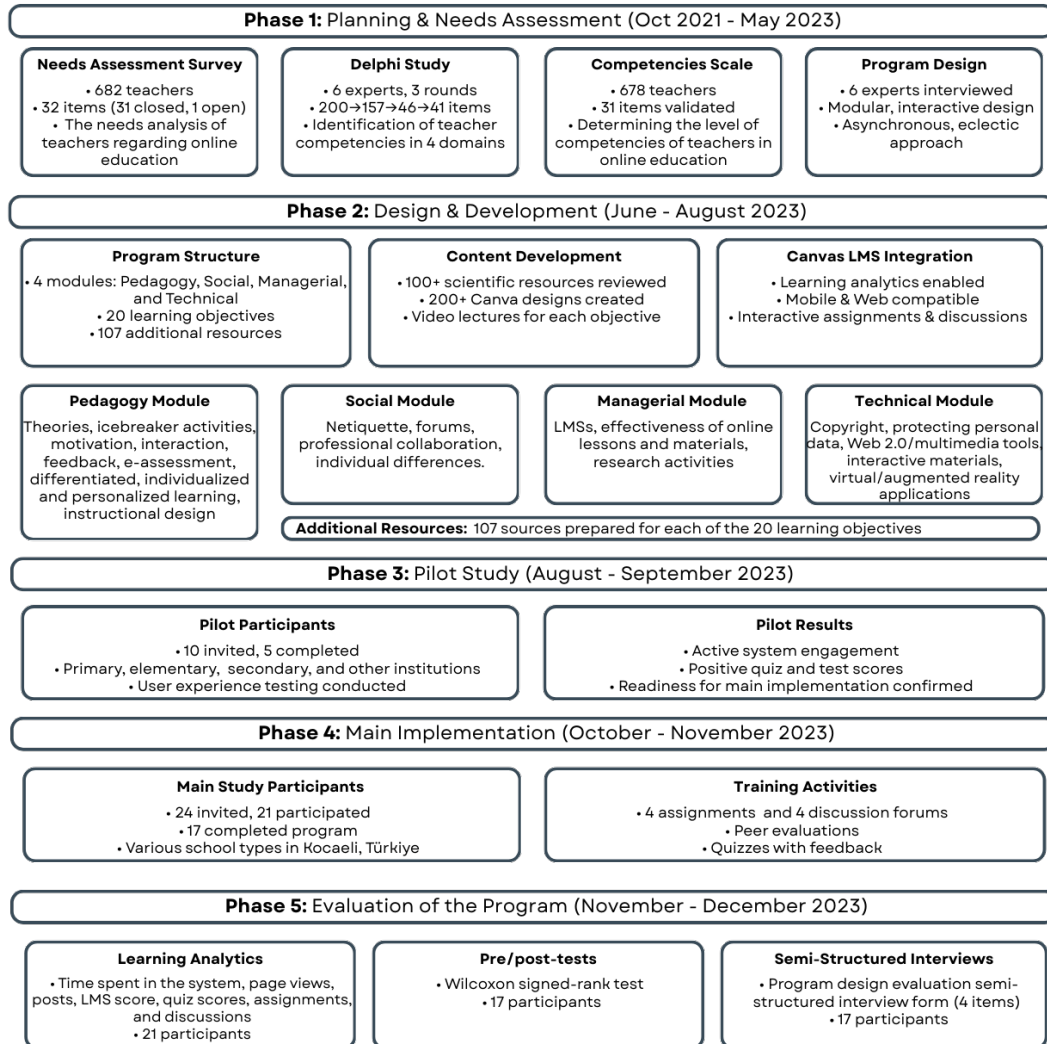
Following the establishment of this theoretical framework, the program design and implementation progressed through five distinct phases. The planning phase spanned from October 2021 to May 2023. Multiple needs assessment instruments were administered, including a survey, a Delphi study to identify key competencies, and preliminary interviews to inform program architecture. Subsequently, the content development phase ran from June through July 2023, during which program content was developed collaboratively by practicing teachers and academic experts. This process involved the review of more than 100 scholarly sources and the design of over 200 custom visuals and interactive components. The resulting content was integrated into the Canvas learning management system (LMS) between August 1 and August 6, 2023.

The instructional delivery phase employed an asynchronous format, enabling flexible access for participants. The program incorporated diverse activities (e.g., pedagogical activities, discussions, quizzes, assignments) and supplementary resources. Participants designed and implemented authentic online lessons, facilitated discussions, and received peer and expert feedback.

Finally, a comprehensive assessment framework evaluated learning outcomes through multiple methods. These included peer review for assignments and varied module-end assessments (e.g., multiple-choice, true/false, open-ended). The LMS provided automated feedback, while the instructor and expert conducted observational evaluations. Follow-up interviews were conducted to collect their insights on learning objectives, content quality, instructional strategies, and assessment. Figure 1 presents an overview of the ITP's methodological implementation stages.

Figure 1

Overview of the In-Service Training Program's Methodological Stages



Note. LMS = learning management system.

Having established these methodological stages, the research team proceeded to pilot the ITP in a real-world educational context.

Pilot Study

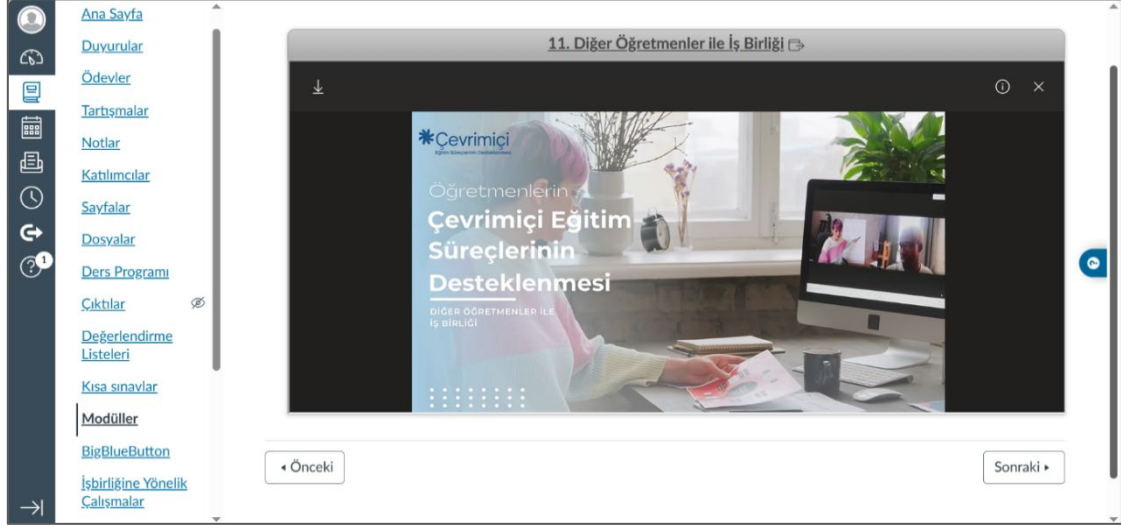
Before launching the full program, a pilot study was conducted from August 23 to September 25, 2023, with five volunteer teachers from primary, elementary, and high schools. The pilot aimed to test usability across platforms, including desktops, Android/iOS mobile browsers, and the Canvas Student app, and to evaluate participant engagement.

Analytics showed high levels of participation, and both pre- and post-tests revealed significant learning gains. Interviews conducted at the end of the pilot confirmed the program's effectiveness and usability. Encouraged by these findings, the full implementation phase ran from October 9 to November 17,

2023, delivering the asynchronous training through the same Canvas platform used during the pilot. Figure 2 shows an example of a video lesson screen.

Figure 2

Video Lesson Interface



Main Implementation

Following successful pilot study validation, the main implementation phase was executed from October 9 to November 17, 2023, maintaining identical platform specifications and instructional design components to those employed during the pilot validation period.

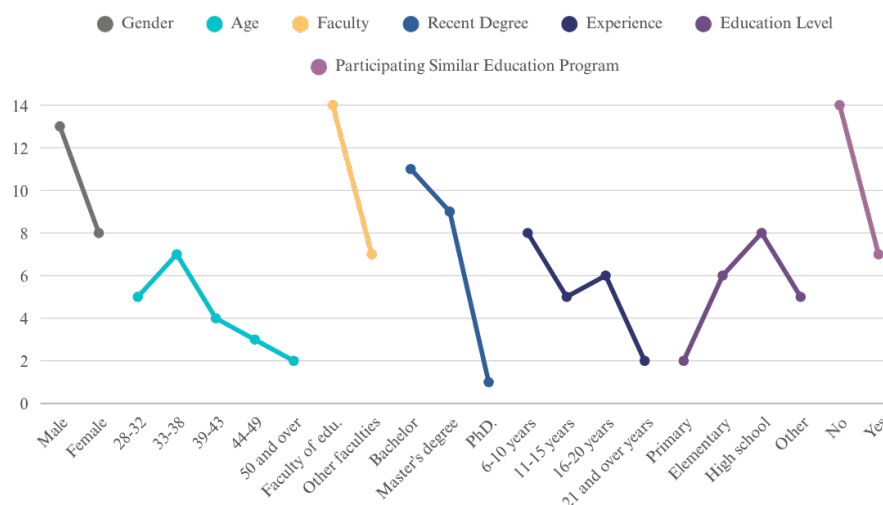
Methodology

Participants and Setting

The study was conducted in Kocaeli, Türkiye; a purposive sampling methodology was employed to ensure comprehensive representation across educational levels and academic disciplines. Initially, 21 teachers were invited to participate through established professional networks and institutional partnerships. Following completion of informed consent procedures and ethical clearance protocols, 17 teachers successfully completed the entire program implementation cycle. The sample size was intentionally determined through consultation with experts in open and distance learning (ODL) as well as curriculum and instruction. While experts recommended maintaining a group size of 10–15 participants to enable systematic tracking of all assignments and activities, timely support for emerging issues, and individualized feedback delivery, the final implementation began with 21 participants to account for the characteristically high attrition rates in online courses. This approach is consistent with the group size parameters recommended by Palloff and Pratt (2007) for asynchronous instruction.

Figure 3

Demographics of the Main Implementation Participants



As illustrated in Figure 3, the final participant cohort comprised teachers representing primary, elementary, and secondary education levels, spanning diverse academic disciplines including mathematics, natural sciences, language arts, social studies, and special education. This heterogeneous sampling approach provided comprehensive perspectives on online education challenges across varied pedagogical contexts and diverse student populations, thereby enhancing the generalizability and ecological validity of research findings.

Data Collection and Analysis

Data collection was done using a triangulated approach with three researcher-developed instruments: the teacher competencies in online education scale (pre-/post-test), the program design evaluation semi-structured interview protocol, and Canvas LMS learning analytics. Learning analytics tracked engagement metrics, system activity, content access, forum participation, and assessment performance to reflect competency development. Quantitative and qualitative data were integrated to understand program effectiveness and teacher development, with convergent findings strengthening validity and divergent findings prompting deeper exploration. Interviews captured participant perceptions, complementing behavioral data from the LMS.

Data collection occurred from February 2022 through December 2023 using multiple modalities (face-to-face and online) for surveys and interviews to ensure accessibility, while LMS analytics were recorded continuously. This triangulated methodology yielded a robust dataset for evaluating competency development and program impact.

Data Analysis Framework

Data analysis followed the mixed-methods strategies of Creswell and Plano Clark (2011). Quantitative data (pre-/post-tests, LMS metrics) were analyzed using descriptive and inferential statistics in SPSS 27. Qualitative data from interviews were thematically analyzed using MAXQDA 2020 (<https://www.maxqda.com/>), following Braun and Clarke's (2006) six-phase framework. The coding process and thematic development were guided by the semi-structured interview protocol structure and established theoretical perspectives on reflective professional practice.

Reliability and Validity Measures

Methodological rigor was maintained through reliability assessment. A self-coding reliability rate of 90% was maintained. Inter-coder reliability assessments with an independent expert achieved agreement rates of 95% (pilot) and 91% (main). These coefficients exceed Miles and Huberman's (1994) recommended 70% threshold, confirming the analytical credibility of the coding procedures. The multiphase mixed-methods design enabled validation of findings through data source triangulation, enhancing internal validity and the robustness of conclusions regarding program effectiveness and teacher competency development.

Results

The 5-week in-service training program was successfully implemented with 17 teachers completing the full sequence of instruction. This section presents findings from the comprehensive mixed-methods evaluation, integrating learning analytics, quantitative assessment data, and qualitative interview responses into a unified analytical narrative that demonstrates program effectiveness and illuminates participant experiences.

Beginning with engagement patterns, learning analytics from the Canvas LMS revealed substantial variation in participant engagement patterns across multiple dimensions (Table 1).

Table 1

Learning Analytics Summary

Participant	Time spent in LMS (hours)	Page views	Posts	LMS score (%)	Quiz score (%)
K1	02.13	166	12	65.53	62.25
K2	02.59	111	4	51.32	48.75
K3	03.10	244	9	84.21	80
K4	06.36	146	10	52.5	42
K5	01.24	137	5	71.58	68
K6	05.57	245	12	81.32	77.25
K7	01.47	81	8	56.32	53.5
K8	03.22	184	11	64.21	61
K9	02.39	164	9	55.53	52.5
K10	02.51	77	5	53.42	50.75
K11	04.48	179	8	79.47	75.5
K12	03.41	238	11	80.53	76.5
K13	04.14	203	11	92.63	88
K14	03.17	155	8	46.05	44
K15	01.38	122	7	86.67	13
K16	13.52	654	12	92.37	87.75
K17	04.15	214	9	86.32	82
K18	00.07	2	0		

Participant	Time spent in LMS (hours)	Page views	Posts	LMS score (%)	Quiz score (%)
K19	00.10	2	0		
K20	04.23	54	4	48.64	26.75
K21	00.15	44	0		
Mean	03.50	162	7,3	69.37	60.52

Note. LMS = learning management system. Empty cells indicate insufficient participant activity to generate scores.

Time spent in the system ranged from 7 minutes to 13.52 hours, with an average of 3.5 hours per participant. These figures should be interpreted cautiously due to technical limitations in Canvas’s tracking capabilities (see Conclusion). Moreover, quantitative time measurements proved less informative than qualitative engagement patterns in understanding the learning experience. For instance, while participant K16 demonstrated the highest activity levels (13.52 hours and 654 page views), participant feedback consistently emphasized that program flexibility, rather than time investment, was the critical factor in facilitating meaningful learning outcomes.

K3 articulated this flexibility as a core strength: “Online environments provide individuals with freedom in terms of place and time. In this respect, I think it is more useful and economical for adults who are aware of their learning responsibilities.” However, this same flexibility demanded self-directed learning capacity, which helps explain the natural attrition from 21 enrollees to 17 completers.

While temporal and navigational patterns illuminated how participants engaged with the program, an equally critical dimension of analysis concerned which content elements elicited the strongest engagement. Table 2 presents assignment and discussion completion rates across the four training modules.

Table 2

Number of Assignments and Discussions Submitted

Domain	No. of assignments	No. of discussions
Pedagogical	13 (Instructional design)	11 (Student interaction)
Social	9 (Initiating/sustaining discussions)	11 (Collaboration with teachers)
Managerial	11 (Using LMS)	10 (Research activities)
Technical	13 (Using Web 2.0/multimedia tools)	12 (Protection of personal data)

Note. LMS = learning management system.

The high engagement levels are notable given that the program required time-intensive tasks, such as detailed instructional design plans and the creation of Web 2.0/multimedia tool projects. This consistent effort, despite the substantial time investment, suggests that participation was driven by a strong alignment with participants’ immediate professional needs and a motivation to apply their learning, rather than by simple topic preference.

Interview responses strongly supported this connection. Participant K1 illuminated this need: “I find the content quite sufficient. ... It was an area where I felt lacking. That’s why I participated the training

program, and I'm satisfied with the result." The high completion rate for the instructional design assignment (13 participants) directly corresponded with this expressed need for practical skills. K15 captured this transformation: "I thought I was conducting online classes just by using videoconferencing. Online classes are not just about that; I improved my instructional design skills."

The instructional design assignment exemplified how the program bridged theory and practice. For instance, participant K16 prepared a 4-hour lesson plan covering all critical design stages, analyzing students, stating objectives, selecting media, ensuring participation, and evaluation, and successfully conducted the lesson in a real synchronous classroom (Figure 4).

Figure 4

Study 1—Instructional Design

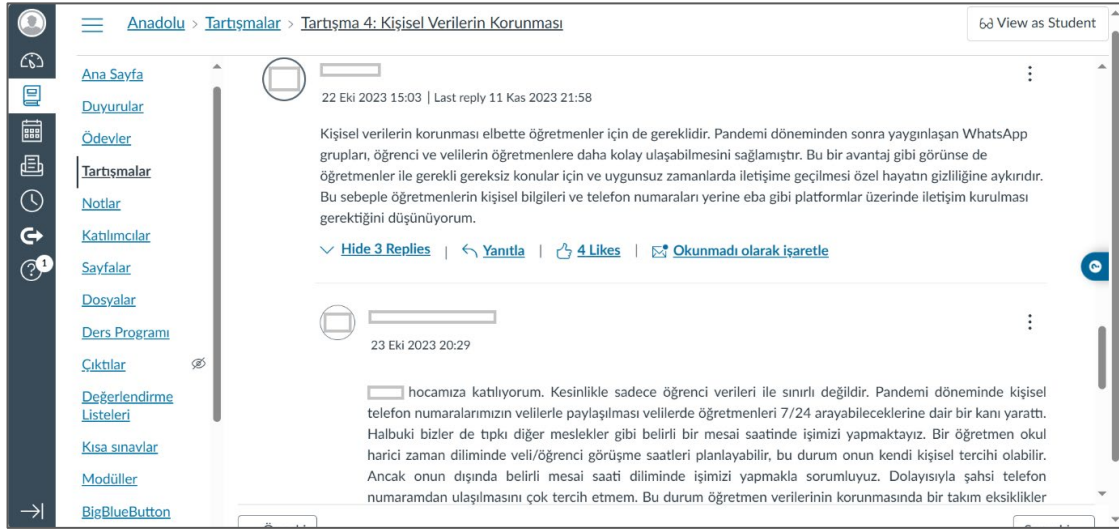
The screenshot shows a web application interface for the ASSURE Instructional Design Model. The header includes the logo 'Cevrimiçi Eğitim Süreçlerinin Desteklenmesi' and the title 'ASSURE Öğretim Tasarımı Modeli'. The form is divided into sections: 'Öğretmen:' with a text input field, 'ÖĞRENCİLERİN ANALİZİ' (Student Analysis), 'Demografik Bilgiler:' (Demographic Information) with a text input field, and 'Ön Öğrenmeler:' (Prior Learning) with a text input field. The 'Öğrenci sayısı, yaşları ve öğrenme profilleri nelerdir?' (How many students, their ages, and learning profiles are they?) question is answered with '10 öğrenci, 15-16 yaş, araştırmacı, duyarlı bireyler, düşünen bireyler, dönüşümlü düşünen bireyler.' (10 students, 15-16 years, researcher, sensitive individuals, thinking individuals, transformative thinking individuals). The 'Ön Öğrenmeler:' question is 'Öğrencilerin konuya ilişkin ön öğrenmeleri/bildikleri nelerdir?' (What are the students' prior learning/knowledge related to the topic?). The interface also shows a sidebar with 'Ödev Yorumları' (Assignment Comments) and a 'Paragraph' text editor.

Such authentic applications fostered a community of practice where participants both validated and inspired one another. K17's peer evaluation validated the quality of learning products: "It was a very effective online instructional design, thank you for your efforts." K16's own reflection revealed the motivational impact: "The applicability and transformation of the knowledge I acquired into a product motivated me personally."

Parallel to these authentic application experiences, asynchronous discussions provided additional opportunities for professional exchange. For the discussion initiation/sustaining assignment, participants were asked to initiate a discussion or participate in existing discussions on the basis of the 20 topics within the four course modules. While only 9 participants submitted the discussion initiation/sustaining assignment, resulting in lower completion metrics (see Conclusion), the 44 posts shared revealed substantive engagement with critical topics. The most discussed subjects—personal data protection, student interaction, and collaboration with other teachers—aligned with participants' emerging professional concerns rather than prescribed curriculum priorities. For the discussion initiation/sustaining assignment, participants were asked to initiate a discussion or participate in existing discussions on the basis of the 20 topics within the four modules of the program, following netiquette rules. The views of K7 and K11 on continuing the discussion are presented in Figure 5.

Figure 5

Assignment 2—Initiating/Sustaining Discussion



Note. English translation of top post: “The protection of personal data is, of course, necessary for teachers as well. WhatsApp groups, which became widespread after the pandemic, have made it easier for students and parents to reach teachers. Although this seems like an advantage, it also violates the privacy of teachers in terms of communication on irrelevant matters and at inappropriate times. For this reason, I think that communication with teachers should be established on platforms such as EBA, where teachers’ personal information and phone numbers are not disclosed.” English translation of bottom post: “I agree with my colleague [name]. It is definitely not limited to just student data. The sharing of our personal phone numbers with parents during the pandemic created an impression among parents that they could call teachers 24/7. However, just like other professions, we perform our work during specific working hours. A teacher may plan parent/student meeting hours outside of school time; this would be their own personal choice. But other than that, we are responsible for doing our job during specific working hours. Therefore, I do not much prefer being reached via my personal phone number. This indicates certain deficiencies regarding the protection of teacher data.”

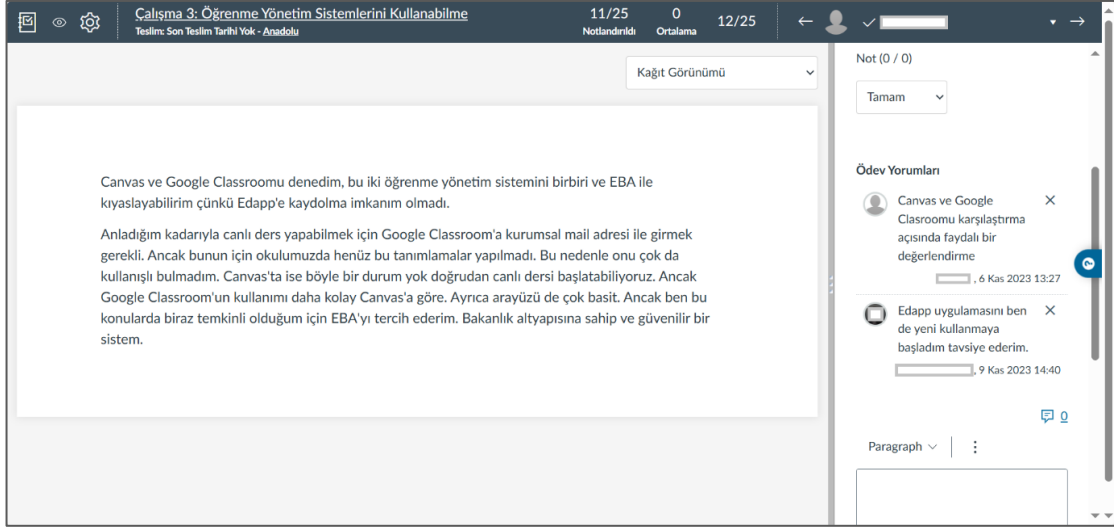
K13 provided crucial context for understanding these engagement patterns: “Classroom practices were organized in a way to support the theoretical foundations. ... In this training, participants are provided with expert support, practice, and the opportunity to communicate and interact through sharing.”

This intentional design prioritized the quality of professional dialogue over discussion frequency, as exemplified by K7’s and K11’s exchanges, which demonstrated genuine knowledge construction through peer interaction. This quality-over-quantity pattern extended to other collaborative elements. K2 recognized the program’s community-building potential: “It also offers the opportunity for collaboration, in my opinion. After all, we work in the same province, so why not carry out or collaborate on the projects we are conducting or could do together?”

This orientation toward practical collaboration was further reflected in participants’ engagement with technology-focused assignments, where they explored tools that could facilitate such professional partnerships. The assignments focused on LMS use and Web 2.0/multimedia tools generated both quantitative evidence of engagement and qualitative insights into skill development trajectories. In the assignment related to LMS use, participants were asked to sign up for Google Classroom, EdApp, and Canvas LMS and explain, with five reasons, which of the three best served their teaching phases. K15’s post, along with the opinions of 11 participants, is presented in Figure 6.

Figure 6

Assignment 3—Using Learning Management Systems



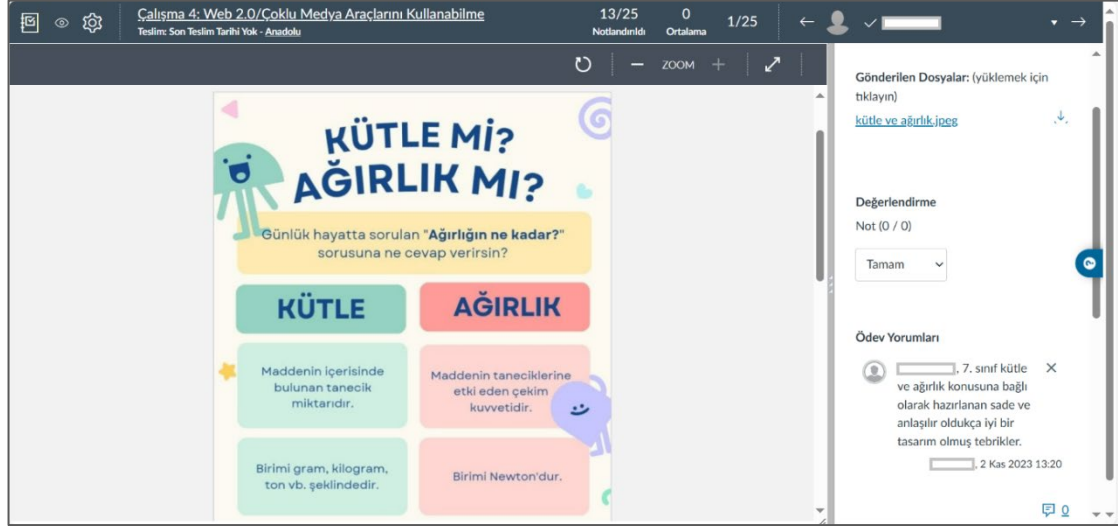
Note. English translation of main text (left): “I have tried Canvas and Google Classroom; I can compare these two learning management systems with each other and with EBA. But I did not have the opportunity to sign up for EdApp. To be able to hold live lessons, it is necessary to log in to Google Classroom with an institutional email address. However, these configurations have not yet been made at our school. For this reason, I did not find it very useful. In Canvas, there is no such situation; we can start live lessons directly. However, I find Google Classroom easier to use than Canvas. Its interface is also very simple. But since I am a bit cautious about these issues, I prefer EBA. It has the Ministry’s infrastructure and is a reliable system.” English translation of comments on right sidebar—top: “A useful evaluation regarding the comparison of Canvas and Google Classroom”; bottom: “I’m also new to EdApp. I’d recommend it.”

K15’s comparative analysis demonstrated participants’ developing capacity to critically evaluate educational technologies. This evaluative competence represented a significant shift from passive technology users to informed decision-makers. K12’s reflection on Web 2.0 tools captured this journey: “I developed my Web 2.0 tool. Acquiring skills ... to the extent that I could perform tasks in my lesson, made me happy.”

The depth of skill acquisition K12 described was exemplified in participants’ Web 2.0 assignments. The participants were asked to select one of the Web 2.0/multimedia tools or interactive material creation software and share the activities they designed using this tool. K11’s work on using Web 2.0/multimedia tools demonstrated both technical proficiency and pedagogical thoughtfulness (Figure 7).

Figure 7

Assignment 4—Using Web 2.0/Multimedia Tools



While individual assignments showcased participant growth, the underlying program design played a role in enabling such development. K8 particularly valued the program’s resource integration: “The necessary resources to make the training quality and effective were also provided. ... [We were] given various examples and additional study opportunities, QR codes/links embedded in the videos, which I find very important.” K6’s system evaluation—“I was using Google Classroom. ... However, I liked Canvas better”—illustrates how hands-on experience enabled informed preferences. K16’s emphasis on resource utility reinforced this pattern: “I frequently used the suggested digital resources during the content development process. I plan to use other suggested digital resources in my future classroom activities.”

While participants’ reflections illuminated the qualitative dimensions of their learning experiences, pre- and post-test measures provided quantitative evidence of knowledge gains. To assess learning gains, pre-test and post-test scores were analyzed using the Wilcoxon signed-rank test due to the non-normal distribution of difference scores. The results are presented in Table 3.

Table 3

Wilcoxon Signed-Rank Test Results

Post-test–pre-test	<i>N</i>	<i>M</i> rank	Sum of ranks	<i>Z</i>	<i>p</i>
Negative ranks	2	4.00	8.00	–3.245	.001
Positive ranks	15	9.67	145.00		
Ties	0				
Total	17				

The Wilcoxon signed-rank test demonstrated a statistically significant improvement in scores from pre-test ($M = 3.54$, $SD = 0.59$) to post-test ($M = 4.17$, $SD = 0.37$): $Z = -3.245$, $p = .001$. To evaluate the magnitude of this change, the effect size r was calculated using the formula $r = Z / \sqrt{N}$ (Rosenthal, 1991), yielding $r = -3.245 / \sqrt{17} = -0.787$. Following Cohen's (1988) benchmarks, this represents a large effect size ($|r| = 0.787 \geq 0.50$), indicating substantial practical significance and suggesting that the intervention produced meaningful improvements in participant performance.

While these statistical measures confirm the intervention's effectiveness, participant responses shed light on the mechanisms underlying these improvements. K1 appreciated the formative assessment design: "When an answer is incorrect, feedback pops up indicating which topic to review. I noticed my mistake, went back, and watched it again. It is a big advantage, well thought out." This feedback mechanism transformed assessment from evaluation into learning opportunity, though the analytics showed differential use.

However, the relationship between engagement patterns and learning outcomes proved more complex than surface-level metrics suggested. K15's case illustrated the limitations of quantitative metrics: Despite an LMS score of 86.67, the quiz score was only 13, suggesting either technical issues or assessment misalignment. Conversely, participants K13 (quiz score: 88) and K16 (quiz score: 87.75) demonstrated consistency across both engagement metrics and assessment outcomes. K13's transformation was particularly notable: "It helped me overcome my prejudices against online education. ... Through this training, I not only acquired many skills related to online education but also developed a deep desire to learn about the discussed topics."

Beyond individual differences in performance metrics, participants converged on identifying the human support structure that enabled their varied paths to success. Despite the program's asynchronous structure, instructor presence emerged as a critical success factor. K16 emphasized: "I believe that tracking participation in the program, grading submitted exams, providing comments on submitted work, [and] instant feedback are very important." K17 reinforced this point: "Despite the asynchronous structure of the program, one of the aspects that pleased me the most was that instructor monitored the entire process and provided guidance."

This instructor support operated within a time frame that participants judged to be well-suited to the learning goals. The 5-week program, described by K17 as "neither long enough to diminish interest nor short enough to be ineffective, but structured within an ideal time frame," was also viewed by K9 as "extremely sufficient for an asynchronous education."

However, the broader issue of workload and time commitment elicited more varied responses. The one dissenting voice on duration (finding it too long) was balanced by K11's acknowledgment of the value–effort trade-off: "The task-oriented assignments took a lot of my time, but I think they were beneficial in terms of application." This tension between time investment and perceived value appeared throughout the data, with most participants ultimately validating the program structure.

Participants' acceptance of the demanding workload reflected appreciation for the program's fundamental design principle: the integration of theoretical knowledge with practical application. K16 articulated this comprehensively: "I think that by supporting what is given in the lectures with additional examples and providing the opportunity to practice in the application part, it contributed to the teachers' competencies in both theoretical and practical areas of online education."

This integration manifested differently across participants based on their starting points and needs. K2's expectation for "a program that would be easily carried out with a practical system and understandable" was met, while K15's more nuanced appreciation of the modular resource structure—"more of these resources were presented in a modular form and in different topics"—indicated sophisticated engagement with the program architecture.

Participants' appreciation extended beyond the program itself to how it compared with alternative learning platforms. K13's comparison with other online learning platforms provided external validation:

I have also taken courses on Coursera and edX, where the instructor encourages practice and gives feedback. However, these are paid, at least for Coursera. The free ones are like YouTube videos or like a slide presentation. In this training, on the other hand, participants are provided with expert support, practice, and the opportunity to communicate and interact through sharing.

These qualitative gains, combined with the significant pre-/post-test improvements, painted a picture of comprehensive professional development that addressed multiple competency dimensions.

While these outcomes demonstrated the program's current effectiveness, participants' forward-looking recommendations revealed their deeper investment in its continued evolution. Participants' suggestions for program enhancement reflected both satisfaction with the current design and vision for future iterations. Content recommendations focused on emerging needs. K15 advocated for artificial intelligence integration: "If we know how these tools are used and their possibilities, we can easily distinguish the unethical behaviors of students"; K16 proposed subject-specific customization: "The program's content could be tailored specifically for each subject"; and K7 suggested adaptive pathways: "It might have been better if the training had been adapted to focus more or less on our specific needs."

Complementing these curricular proposals, participants offered recommendations for enhancing the program's infrastructure and collaborative elements. K15 envisioned organic program growth: "It would be a program that grows gradually, and as it grows, its content would be enriched and kept up to date." K2 recommended formalized learning groups: "In this way, the sense of community increases, and the feeling of loneliness decreases." The fact that most participants stated they would recommend the program to colleagues provided perhaps the strongest qualitative endorsement.

However, a complete program evaluation requires acknowledgment of the technical obstacles some participants encountered. Technical issues, file downloads, peer evaluation functionality, and video rewinding affected some participants but did not significantly detract from overall program effectiveness. K6's positive evaluation of the Canvas LMS, combined with the high completion rates for platform-based activities, suggest that technical barriers were minimal for most participants.

When viewed holistically, the patterns emerging across statistical measures, usage analytics, and participant narratives point to the underlying factors driving program effectiveness. The convergence of quantitative and qualitative evidence reveals a professional development program that succeeded by honoring adult learning principles: flexibility in structure, relevance to professional context, integration of theory and practice, opportunities for peer interaction, and responsive instructor support. The high completion rate, significant learning gains, varied but substantive engagement

patterns, and consistently positive participant evaluations collectively demonstrate program effectiveness.

Beneath these statistical patterns lie personal narratives of professional transformation that illuminate what the numbers represent in human terms. Most significantly, the qualitative evidence revealed transformations that quantitative metrics alone could not capture: K13's shift from prejudice to enthusiasm for online education, K15's evolution from viewing online teaching as "just video conferencing" to sophisticated instructional design, and K12's pride in newly developed Web 2.0 competencies. These narrative accounts of professional growth, supported by learning products demonstrating practical application and validated by significant statistical improvements in knowledge measures, establish that the program achieved its fundamental purpose: supporting teachers' online education processes in ways both measurable and meaningful.

Discussion

This study demonstrates that addressing K–12 online teacher preparation requires both practical program design and a theoretical understanding of professional learning in ODL contexts. The growing demand for in-service training (Koçak & Çakır, 2023) reflects teachers' evolving needs across pedagogical, social, managerial, and technical domains (Martín-Bylund & Stenliden, 2024). Responding to these needs, the developed program integrated theory and practice, emphasized real-world classroom application, and adapted to diverse contexts, distinguishing it from existing models such as AKADEMA (Anadolu University, n.d.), edX (Tsinghua University, n.d.), and Udemy (Subramonian, 2021), as well as MoNE's ÖBA platform (Turkish MoNE, n.d.) and scholar-designed programs (Kocatürk-Kapucu & Adnan, 2018). Unlike typical short, theory-driven modules, this program engaged teachers in authentic, practice-oriented tasks, including designing online lessons, integrating Web 2.0 tools, and evaluating implementations, demonstrating that experiential, contextually grounded learning is key to effective online professional development for teachers (Turley, 2022).

This reconceptualization of structure and flexibility advances understanding of how learning occurs in ODL environments. This study provides empirical support for the proposition that experiential learning in online settings is essential for developing online teaching competence. Participants' progression from viewing online teaching as "just videoconferencing" to sophisticated instructional design highlights that tacit knowledge about digital pedagogy cannot be transmitted through explicit instruction alone. Teachers who experienced online learning as students developed deeper insights into digital instructional challenges, echoing Theelen et al. (2020) on the importance of authentic learning experiences. These findings challenge deficit models that treat teacher preparation gaps as missing knowledge, suggesting instead that effective online pedagogy emerges through guided experiential transformation rather than through knowledge transmission alone.

This study operationalized this principle through a community of inquiry (CoI)–based model (Garrison et al., 1999), enabling teachers to first experience a functional CoI as learners. The ITP exemplified all three presences: Teaching presence was modelled through timely feedback and instructor monitoring while being developed through instructional design tasks; social presence was fostered via structured discussions and peer evaluations; and cognitive presence emerged through iterative engagement supported by formative feedback. This dual function of CoI as both learning

environment and outcome provides a robust framework for teacher professional development in online contexts.

Building upon this CoI foundation, the findings extend Moore's (1993) transactional distance theory by revealing a paradox: The ITP's highly structured competency framework facilitated rather than constrained dialogue and learner autonomy. While traditional theory posits that increased structure reduces dialogue and autonomy, structural clarity enabled personalized learning pathways. Wide variation in engagement patterns demonstrates that in professional development contexts, explicit frameworks may reduce psychological distance by clarifying expectations while supporting autonomous learning trajectories. This challenges conventional interpretations that treat structure and autonomy as inversely related, suggesting instead that effective ODL design requires simultaneous attention to clarity of outcomes and multiplicity of pathways. Participants achieved comparable outcomes through markedly different engagement patterns, indicating that flexibility should be conceptualized as enabling personalized trajectories within structured frameworks rather than as absence of structure. This approach aligns with research emphasizing that effective teacher professional development must be grounded in clearly defined competency frameworks rather than ad hoc training initiatives and should provide sustained support along with multiple opportunities for practice and reflection (Sims et al., 2023).

Within this model, interaction quality consistently outweighed quantity. Nine participants completed formal discussions yet generated 44 meaningful exchanges, demonstrating that in professional development contexts, the depth and relevance of dialogue matter more than frequency. Even limited but substantive interactions fostered genuine knowledge construction, suggesting that professional learners engage in selective, purposeful peer exchanges rather than frequent but superficial ones. This finding requires ODL theories to account more carefully for how professional contexts shape interaction patterns and learning outcomes, moving beyond simple metrics of participation frequency.

Finally, the study illustrates how Kolb's (2014) experiential learning cycle can operate effectively in fully asynchronous ODL contexts. Participants developed and implemented lesson plans, received peer validation, and reflected on motivational impacts, demonstrating that meaningful experiential learning does not require synchronous interaction. This finding carries important implications for cognitive load theory (Sweller, 1988), as the study revealed that authentic, complex tasks can enhance rather than hinder learning when appropriately contextualized. Specifically, when tasks align with learners' immediate professional needs, the germane cognitive load generated by relevant, complex activities supports deeper engagement and professional growth (Zeithofer et al., 2024).

Conclusion

The practical success of completion rates, large effect sizes, and positive participant evaluations provides evidence that systematic, competency-based professional development can effectively develop online teaching competencies. This effectiveness depends on several key design insights: authentic learning experiences anchored in participants' actual teaching contexts; structured architectures balancing predetermined pathways with individualized exploration; strategic instructor engagement combining timely interventions with learner autonomy; purposeful collaborative activities promoting both cognitive and social dimensions; and integration of genuine workplace tasks bridging theory with practice.

These insights extend beyond practical utility to advance ODL scholarship by refining transactional distance theory, extending the CoI framework, and elaborating experiential learning principles for digital contexts. Importantly, these advances move beyond describing what works to explaining why it works and under what conditions, providing foundations for future scholarship and practice. In doing so, the findings challenge traditional approaches by demonstrating that online education requires fundamental reconceptualization of pedagogical practice rather than mere technological adoption.

These design principles thus serve dual purposes: They explain the program’s practical success while offering theoretical insight into the mechanisms through which effective online professional learning occurs. The theoretical contributions emerge from deep qualitative analysis of the purposive sample of 17 completing participants. It must be noted that while this approach limits statistical generalizability, the findings are also contextualized by the specific demographic composition of the cohort; although recruitment extended to all schools in the province, the study predominantly attracted educators from high-performing institutions, resulting in a lack of representation from less advantaged educational contexts.

Nevertheless, this deep analysis enables refinements to transactional distance theory, extensions of the CoI framework, and elaborations of experiential and cognitive load theories in ODL contexts, all providing empirically grounded hypotheses for larger-scale validation. Beyond such validation, future theoretical development must address the epistemological challenges of ODL assessment, a necessity underscored by specific technical limitations encountered in this study. For instance, platform incompatibilities with iOS tablets prevented the recording of time spent on external videos, and the lack of device reporting made it impossible to verify the scope of this issue, ultimately resulting in the underreporting of overall engagement time. Furthermore, rigid algorithmic tracking incorrectly marked assignments as “incomplete” when participants omitted a final confirmation step despite having completed the discussion, artificially lowering completion metrics for a subset of users. These technical discrepancies exemplify the disconnect between LMS metrics and actual learning outcomes, raising fundamental questions: What constitutes evidence of learning in asynchronous digital environments? How can we distinguish meaningful engagement from superficial interaction? How might we develop learning analytics capable of capturing qualitative transformations in understanding? These epistemological dimensions represent critical frontiers for advancing ODL scholarship toward more robust theoretical frameworks that adequately capture the complexity of learning in ODL environments.

Digital Transparency Statement

Artificial intelligence assistance (Claude by Anthropic) was used in the preparation of this manuscript for the limited purposes of proofreading and checking English grammar. The authors reviewed and edited the output as needed and take full responsibility for the content.

Acknowledgment

This study is derived from Barış Avcı’s doctoral dissertation (Avcı, 2024), completed at Anadolu University, under the supervision of Meral Güven and submitted on July 6, 2024.

References

- Anadolu University. (n.d.). *Dijital eğitim ortamlarına giriş [Introduction to digital learning environments]*. AKADEMA.
<https://acik.anadolu.edu.tr/materialdetail?materialcode=10982&sourceid=1>
- Avcı, B. (2024). Öğretmenlerin çevrimiçi eğitim süreçlerinin desteklenmesine yönelik eğitim programı tasarısının hazırlanması, uygulanması ve değerlendirilmesi [Preparation, implementation and evaluation of a training program design for supporting teachers' online education processes] [Doctoral dissertation, Anadolu University]. *Council of Higher Education Thesis Center*.
https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=usXiZIM9Lpowk-YzRoAT-ypU_NRvJLyKFUxfykrTfEe7XbmwxlUas9V_54rZnnTq
- Avcı, B., & Güven, M. (2021). Öğretmenlerin çevrim içi eğitime ilişkin hizmet içi eğitim gereksinimlerinin belirlenmesi. [Identifying training needs of in-service teachers for online education]. *Dokuz Eylül Üniversitesi Buca Eğitim Fakültesi Dergisi*, (51), 345–367.
<https://doi.org/10.53444/deubefd.882866>
- Avcı, B., & Güven, M. (2025). K-12 teachers' online education competencies in Türkiye. *American Journal of Distance Education*, 1–15. <https://doi.org/10.1080/08923647.2025.2513099>
- Bakioğlu, B., & Çevik, M. (2020). COVID-19 pandemisi sürecinde fen bilimleri öğretmenlerinin uzaktan eğitime ilişkin görüşleri [Science teachers' views on distance education in the COVID-19 pandemic process]. *Turkish Studies*, 15(4), 109–129.
<https://doi.org/10.7827/TurkishStudies.43502>
- Basilotta-Gómez-Pablos, V., Matarranz, M., Casado-Aranda, L. A., & Otto, A. (2022). Teachers' digital competencies in higher education: A systematic literature review. *International Journal of Educational Technology in Higher Education*, 19(1), Article 8.
<https://doi.org/10.1186/s41239-021-00312-8>
- Berge, Z. L. (1995). The role of online instructor/facilitator. *Educational Technology*, 35(1), 20–30.
https://www.researchgate.net/publication/238348806_the_role_of_the_online_instructor_facilitator
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Brenner, C. A. (2022). Self-regulated learning, self-determination theory and teacher candidates' development of competency-based teaching practices. *Smart Learning Environments*, 9(1), Article 3. <https://doi.org/10.1186/s40561-021-00184-5>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203771587>
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson.

- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Sage Publications.
- Fayda-Kınık, F. Ş. (2021, July). *Türkiye’de “çevrimiçi eğitim” araştırmaları: güncel eğilimler üzerine bir inceleme* [“Online education” research in Turkey: A study on current trends] [Paper presentation]. Ninth International Eurasian Educational Research Congress (EJERCongress 2021), Aksaray University, Aksaray, Türkiye.
https://www.researchgate.net/publication/358346882_turkiye'de_cevrimici_egitim_arastirmalari_guncel_egilimler_uzerine_bir_inceleme
- 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](https://doi.org/10.1016/S1096-7516(00)00016-6)
- İpek, D. S., Aydın, G. Ç., Çelikdemir, K., Celep, N. D., Sunar, S., & Karip, E. (2020). *COVID-19 sürecinde eğitim: uzaktan öğrenme, sorunlar ve çözüm önerileri* [Education during the COVID-19 pandemic: Distance learning, problems and proposed solutions report]. TEDMEM.
<https://tedmem.org/storage/publications/February2023/FvZgqcTeABuqBueLWx50.pdf>
- Kamal, T., & Illiyan, A. (2021). School teachers’ perception and challenges towards online teaching during COVID-19 pandemic in India: An econometric analysis. *Asian Association of Open Universities Journal*, 16(3), 331–325. <https://doi.org/10.1108/AAOUJ-10-2021-0122>
- Kaptanoğlu, B. (2021). *Facilitating online education during the COVID-19 pandemic: Challenges & opportunities* [Master’s thesis, İstanbul Bilgi University]. Council of Higher Education Thesis Center. https://tez.yok.gov.tr/ulusaltezmerkezi/tezgoster?key=8tbpippmwv_b-irrn9yeal_stlikrvowbklkg-ffvzbkn1bsl0cry8rr9zqtgex
- Koçak, S., & Çakır, R. (2023). A qualitative study on successfully sustaining online teaching: Teachers needs in online education. *Education Mind*, 2(1), 45–55.
<https://doi.org/10.58583/pedapub.em2304>
- Kocatürk-Kapucu, N., & Adnan, M. (2018). Uzaktan öğretimde çevrimiçi eğitmen başarısının değerlendirilmesi [Assessment of e-tutor success in online professional development]. *HAYEF: Journal of Education*, 15(1), 7–20.
<https://www.proquest.com/openview/f98c4eod3fd98bada3b450deca55c9c8/1?pq-origsite=gscholar&cbl=1796367>
- Koenig, R. (2020, May 28). *Pandemic may (finally) push online education into teacher prep programs*. Ed Surge. <https://www.edsurge.com/news/2020-05-28-pandemic-mayfinally-push-online-educationinto-teacher-prep-program>
- Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT Press.
- König, J., Jäger-Biela, D. J., & Glutsch, N. (2020). Adapting to online teaching during COVID-19 school closure: Teacher education and teacher competence effects among early career

- teachers in Germany. *European Journal of Teacher Education*, 43(4), 608–622.
<https://doi.org/10.1080/02619768.2020.1809650>
- Larson, J. S., & Archambault, L. (2019). The extent of K–12 online teacher development: A disconnect between preparation and practice. In T. Heafner, R. Hartshorne, & R. Thripp (Eds.), *Handbook of research on emerging practices and methods for K–12 online and blended learning* (pp. 57–77). IGI Global. <https://doi.org/10.4018/978-1-5225-8009-6.ch003>
- Leary, H., Dopp, C., Turley, C., Cheney, M., Simmons, Z., Graham, C. R., & Hatch, R. (2020). Professional development for online teaching: A literature review. *Online Learning*, 24(4), 254–275. <https://doi.org/10.24059/olj.v24i4.2198>
- Lokey-Vega, A., Jorrín-Abellán, I., & Pourreau, L. (2018). Theoretical perspectives in K–12 online learning. In K. Kennedy & R. E. Ferdig (Eds.), *Handbook of research on K–12 online and blended learning* (pp. 65–89). ETC Press. <http://doi.org/10.1184/R1/6686810>
- Mackay, J. (2013). *Online schools*. Gale Cengage Learning.
- Martín-Bylund, A., & Stenliden, L. (2024). Closer to the senses in post-pandemic teacher training—Reclaiming the body in online educational encounters. *Education and Information Technologies*, 29, 3133–3154. <https://doi.org/10.1007/s10639-023-11952-w>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Sage Publications, Inc.
- Moore, M. G. (1993). Theory of transactional distance. In D. Keegan (Ed.), *Theoretical principles of distance education* (pp. 22–29). Routledge. <https://doi.org/10.4324/9780203983065>
- Nastasi, B. K., Hitchcock, J., Sarkar, S., Burkholder, G., Varjas, K., & Jayasena, A. (2007). Mixed methods in intervention research: Theory to adaptation. *Journal of Mixed Methods Research*, 1(2), 164–182. <https://doi.org/10.1177/1558689806298181>
- Orakcıoğlu, E. (2019). *Türkiye’de uzaktan eğitim temalı 2013–2018 yılları arasında yapılan lisansüstü tezlerin incelenmesi [An examination of postgraduate theses on distance education conducted in Türkiye between 2013 and 2018]* [Master’s thesis, Eskişehir Osmangazi University]. Council of Higher Education Thesis Center
https://tez.yok.gov.tr/ulusaltezmerkezi/tezgoster?key=mir2lxqk1dkmq9ige3pzbkvm12_wuo2v9gdxmio7ei_wnulydmxwyevabgeoqwn7
- Palloff, R. M., & Pratt, K. (2007). *Building online learning communities: Effective strategies for the virtual classroom* (2nd ed.). John Wiley & Sons.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Rev. ed.). Sage Publications, Inc. <https://doi.org/10.4135/9781412984997>
- Sengil-Akar, S., & Kurtoglu-Erden, M. (2021). Distance education experiences of secondary school math teachers during the pandemic: A narrative study. *Turkish Online Journal of Distance Education*, 22(3), 19–39. <https://doi.org/10.17718/tojde.961809>

- Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Cottingham, S., Stansfield, C., Goodrich, J., Van Herwegen, J., & Anders, J. (2023). Effective teacher professional development: New theory and a meta-analytic test. *Review of Educational Research*, 95(2), 213–254.
<https://doi.org/10.3102/00346543231217480>
- Stavermann, K. (2025). Online teacher professional development: A research synthesis on effectiveness and evaluation. *Technology, Knowledge and Learning*, 30(1), 203–240.
<https://doi.org/10.1007/s10758-024-09792-9>
- Subramonian, S. (2021, September). *Professional diploma in online teaching*. Udemy.
<https://www.udemy.com/course/fully-accredited-doctor-of-online-education-ded/>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. [https://doi.org/10.1016/0364-0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7)
- Theelen, H., van den Beemt, A., & Brok, P. den. (2020). Enhancing authentic learning experiences in teacher education through 360-degree videos and theoretical lectures: reducing preservice teachers' anxiety. *European Journal of Teacher Education*, 45(2), 230–249.
<https://doi.org/10.1080/02619768.2020.1827392>
- Tian, J., & Tian, W. (2025). Modeling K12 teachers' online teaching competency and its predictive relationship with performance-a mixed-methods study based on behavioral event interviews. *Behavioral Sciences*, 15(5), Article 628. <https://doi.org/10.3390/bs15050628>
- Tschida, C. M., Hodge, E. M., & Schmidt, S. W. (2017). Learning to teach online: Negotiating issues of platform, pedagogy, and professional development. In Information Resources Management Association (Ed.), *Medical education and ethics: Concepts, methodologies, tools, and applications* (pp. 236–257). IGI Global. <https://doi.org/10.4018/978-1-5225-0978-3.ch013>
- Tsinghua University. (n.d.). *TsinghuaX: Shifting your teaching online*. edX.
<https://www.edx.org/learn/education/tsinghua-university-shifting-your-teaching-online/>
- Turkish Ministry of National Education. (n.d.). *Eğitimler [Trainings]*. ÖBA.
<https://www.oba.gov.tr/egitim/listeleme>
- Turley, C. A. (2022). *A missing piece in the K–12 online learning puzzle: Professional development for online teachers* [Doctoral dissertation, Brigham Young University]. BYU ScholarsArchive.
<https://scholarsarchive.byu.edu/etd/9401/>
- Zeithofer, I., Zumbach, J., & Schweppe, J. (2024). Complexity affects performance, cognitive load, and awareness. *Learning and Instruction*, 94, Article 102001.
<https://doi.org/10.1016/j.learninstruc.2024.102001>
- Zuo, M., Yan, Y., Ma, Y., & Luo, H. (2024). Modeling the factors that influence schoolteachers' work engagement and continuance intention when teaching online. *Education and Information Technologies*, 29, 9091–9119. <https://doi.org/10.1007/s10639-023-12186-6>



August – 2026

Between Obligation and Aspiration: Motivation, Emotions, and Systemic Barriers in Online Professional Development for Rural Teachers

Rong Hua¹, Chunmei Yang^{2*}, and Jiawei Chen³

¹Faculty of Foreign Language Department, Hainan Normal University, Hainan, China; ²Faculty of Educational Science, Nanjing Normal University, Nanjing, China; ³Faculty of Teacher Education, Chongqing Normal University, Chongqing, China;

*Corresponding Author

Abstract

This study investigates how rural teachers in China experience online professional development (OPD), focusing on the affective, motivational, and structural dimensions of their engagement. Drawing on in-depth interviews with 30 rural primary and secondary school teachers, this study employed sentiment analysis and thematic coding to explore participants' emotional orientations, sources of motivation, and perceived barriers. While teachers generally expressed positive attitudes toward OPD, their actual experiences revealed notable tensions between optimism and disconnection. Three interrelated findings emerged: (a) a mismatch between favorable attitudes and unsatisfying learning encounters; (b) participation driven more by external mandates than by internalized professional goals; and (c) multiple constraints embedded in training design, institutional policies, sociocultural environments, and digital competence gaps. Based on these insights, the study proposes three conceptual constructs—*affective-structural dissonance*, *controlled engagement*, and *layered constraint ecology*—to explain the complex interplay between emotional readiness and systemic limitations. These findings contribute to the literature on distributed learning by emphasizing the need for motivation-sensitive and context-responsive OPD design, particularly in underserved rural settings.

Keywords: rural teachers, online professional development, motivation, self-determination theory, affective engagement, structural barriers

Introduction

The accelerated global digitalization of education has positioned online teacher training as a critical vehicle for advancing professional development (UNESCO, 2023). Emerging evidence indicates that teachers' affective experiences, learning motivation, and barriers encountered in online training environments directly shape online learning outcomes and career trajectories (Darling-Hammond et al., 2017). However, current scholarship predominantly emphasizes technological infrastructure, curriculum design, and assessment mechanisms while insufficiently addressing the affective dimensions, motivational complexities, and situated challenges that teachers face (Bozkurt & Sharma, 2021). This oversight proves particularly consequential in rural contexts, where systemic constraints, including digital resource disparities, inadequate technical support, and unaddressed personalized learning needs, severely compromise training effectiveness (Zhang & Wang, 2022). In such settings, many rural teachers engage in online professional development (OPD) not out of intrinsic aspiration but due to administrative requirements and top-down mandates. This creates a tension between external obligation and internal growth desires, a dynamic that is often overlooked in the current discourse on online teacher training.

Recent theoretical advances reveal dynamic interplays between teachers' emotional states and motivational configurations in digital learning ecosystems, factors that critically mediate training engagement and pedagogical transfer (Deci & Ryan, 2000; Sun & Rueda, 2012). Self-determination theory (SDT) provides a robust framework for examining this interplay by distinguishing between extrinsic and intrinsic motivation and explaining how institutional pressures can either support or hinder motivation internalization. However, the processes through which rural teachers cultivate sustainable affective engagement, maintain self-regulated motivation, and navigate structural barriers in OPD remain under-explored through systematic qualitative inquiry. This study addresses this critical gap by investigating how rural primary and secondary school teachers in China experience OPD programs, with specific attention to their motivation sources, emotional responses, and contextual challenges. Combining sentiment lexicon analysis with constructivist grounded theory coding, the study explores four key questions:

1. What are rural teachers' overall attitudes toward online training?
2. What are their primary motivational drivers?
3. What affective and structural barriers do they encounter?
4. How can OPD programs be improved to support more sustainable, autonomy-driven participation?

By unpacking these dynamics, the study aims to illuminate how online teacher development in resource-constrained settings can evolve from compliance-oriented implementation to motivation-sensitive and emotionally supportive practice.

Literature Review

Influencing Factors in OPD for Teachers

Multiple studies have identified a range of factors influencing its effectiveness, including learner

characteristics, platform quality, institutional support, and course design. For instance, Wang et al. (2022) argue that teachers' psychological traits, learning autonomy, and engagement attitudes significantly affect the quality of online training, especially in western China. Similarly, Kong and Kang (2022) emphasize that the effectiveness of online training is mediated by a combination of teachers' digital competence, course quality, and environmental conditions.¹ Beyond individual characteristics, the nature of learning experiences also plays a critical role. Kuo et al. (2014) found that learner satisfaction is shaped by the richness of course resources and the extent to which learning methods are autonomous or collaborative. These findings suggest that technical content alone is insufficient; the learning experience must also foster engagement and interaction to support adaptability to online formats.

Collectively, the existing research suggests that the effectiveness of OPD is shaped not only by individual teacher characteristics but also by core online and distance learning (ODL) dimensions such as interaction quality, learner support mechanisms, and course design features.

Despite the substantial body of research that has explored factors influencing online learning, most studies have predominantly focused on preservice teachers (Batmang et al., 2021; Subekti, 2020) or specific subject-area teachers (Gorozidis et al., 2020). Subekti (2020), for example, identified three primary challenges—technological infrastructure, instructor competence, and limited peer interaction—that shaped preservice teachers' perceptions of online learning. These findings were echoed by Batmang et al. (2021), who highlighted the influence of lecturers' pedagogical abilities alongside learners' self-discipline and learning passion as critical determinants of online learning success.

In-service teachers in rural settings, however, remain underrepresented in this body of literature. This imbalance limits our understanding of how ODL systems operate in rural and resource-constrained contexts, where interaction opportunities, learner support, and course design are shaped by markedly different institutional and social conditions. Gorozidis et al. (2020) conducted one of the few studies in this area, exploring how physical education teachers used a Facebook-based network to support their online professional growth. Although innovative, this study remains an exception in a field dominated by studies in urban or higher education contexts.

Recent international studies from resource-constrained contexts suggest that the challenges faced by rural teachers in OPD are not limited to China. Comparative research on teacher autonomy and agency in Brazil and Norway demonstrates that highly centralized accountability systems often restrict teachers' scope of professional decision-making, resulting in compliance-oriented participation rather than autonomous engagement (Lennert da Silva & Mølsted, 2020). Similarly, studies conducted in rural primary schools in Cambodia reveal that global policy initiatives frequently translate into surface-level participation at the local level, while structural barriers and contextual constraints generate emotional detachment and limited professional agency (Ravet & Mtika, 2024). These findings indicate that policy-driven professional development under constrained conditions may produce formal engagement without corresponding motivational or affective investment. These studies highlight a recurring challenge from an ODL perspective: Policy-driven designs may ensure participation at the surface level yet fail to provide autonomy-supportive interaction structures and learner support systems that sustain meaningful engagement.

Moreover, investigation results indicate that learning experiences, course resources, autonomous or collaborative learning methods, and time availability all influenced the adaptability of rural teachers to distance learning. Yet, these factors, which collectively impact rural teachers' participation and learning outcomes in rural online training, have not been adequately explored; a holistic analysis that integrates these contextual barriers with teachers' motivational and emotional responses is lacking.

Despite extensive research on structural and pedagogical aspects of online training, the affective and motivational dynamics shaping rural teachers' engagement remain under-explored within core ODL concerns. This oversight limits our understanding of how and why rural teachers persist in or disengage from online professional learning experiences.

Motivation in Online Learning: A Self-Determination Theory Perspective

SDT classifies human motivation into intrinsic motivation, extrinsic motivation, and amotivation (Deci & Ryan, 1985). Central to SDT are three basic psychological needs that shape motivational regulation: autonomy, competence, and relatedness. When these needs are supported, individuals are more likely to internalize external goals and sustain engagement; when frustrated, participation may become compliance-driven or emotionally disengaged. SDT provides a theoretical lens for understanding how online environments support or inhibit motivational development. In ODL contexts, this lens is particularly useful for examining how interaction design, learner support arrangements, and institutional structures shape teachers' motivational regulation. Studies have shown that intrinsic motivation is closely associated with positive emotional experiences and long-term participation in online learning (Ryan & Deci, 2020). In contrast, extrinsically motivated learners are often driven by grades, completion mandates, or external recognition.

SDT has been widely applied in the field of online learning to examine learner engagement and persistence (Sun & Rueda, 2012). Previous international studies further suggest that teacher agency plays a critical mediating role in how policy-driven professional development initiatives are enacted at the local level. Comparative research indicates that when institutional demands constrain teachers' autonomy, professional learning is more likely to be experienced as an externally regulated obligation rather than a meaningful developmental opportunity (Lennert da Silva & Mølstad, 2020). Such findings provide an important international theoretical backdrop for understanding rural teachers' motivation in OPD contexts characterized by limited choice, high accountability, and uneven support structures. In resource-constrained settings, the suppression of autonomy and relatedness may also intensify emotional burdens, leading teachers to rely on personal resilience or informal coping strategies to sustain participation (Haider et al., 2025). These conditions are likely to inhibit the internalization processes central to SDT, resulting in participation driven more by obligation than aspiration.

Recent studies in China have begun to explore how SDT applies to teacher professional development in digital contexts. Zhang and Wang (2022) found that rural teachers frequently rely on external incentives due to high cognitive demands and limited support systems. Similarly, Liang and Ding (2022) emphasized the importance of sustained intrinsic motivation as a condition for effective online learning among rural educators. Yan and Shan (2020) further argued that teacher training in the post-pandemic era must

transition from policy-driven compliance to empowerment-based models that address teachers' internal goals.

Despite growing interest in SDT-based analysis, most studies have focused on students or urban teachers. Few have examined how SDT explains rural teachers' motivational structures in online learning environments—settings often characterized by low autonomy, weak peer support, and technological challenges. Such conditions may suppress the internalization process central to SDT, leading teachers to participate out of obligation rather than aspiration.

This study adopts SDT as a theoretical framework to explore how rural teachers' intrinsic and extrinsic motivations interact with their affective experiences and the structural constraints they encounter in OPD.

Research Gap and Rationale

Previous research has predominantly examined online training effectiveness from the perspectives of learners in higher education or preservice teacher populations, often with a focus on specific subject areas (Batmang et al., 2021; Gorozidis et al., 2020; Subekti, 2020). In-service rural teachers, particularly those working in low-resource school environments, have received comparatively little attention despite being a major target group of the OPD initiatives in many countries. While some studies have identified technological, pedagogical, or institutional barriers faced by rural teachers, they often treat these factors as isolated challenges rather than as interconnected experiences that shape motivation and participation over time.

Moreover, prior studies have often treated structural, pedagogical, and emotional factors as isolated issues rather than as interconnected features of teachers' ODL experiences. Addressing these gaps requires an integrated ODL perspective that connects design conditions, motivational regulation, and affective experience.

In addition, although sentiment analysis has been applied in online course evaluations, it is still at a preliminary stage in the field of teacher professional development. Most sentiment-related studies rely on large-scale open datasets such as massive open online courses (MOOCs) or student forums, while few incorporate emotion analyses into qualitative interview data with in-service teachers. This limits our understanding of the nuanced emotional trajectories rural teachers experience in digital learning environments. Furthermore, while SDT has been widely used to examine learners' motivation, there is a lack of research integrating SDT with empirical studies of rural teachers' online learning engagement, particularly in the context of state-mandated professional development programs. As a result, little is known about how extrinsic obligations and intrinsic aspirations coexist or conflict within the affective and professional lives of these teachers.

To address these gaps, the present study draws on SDT to examine the motivational configurations, emotional experiences, and systemic barriers encountered by rural teachers in China. Combining sentiment lexicon analysis with constructivist grounded theory, the study seeks to offer a theoretically grounded, context-sensitive account of rural teachers' OPD journeys.

Method

Research Design

This study employed a mixed-methods design that integrated lexicon-based sentiment analysis to investigate rural teachers' affective experiences and motivational dynamics in OPD.

Participants

This study employed theoretical sampling to select rural primary and secondary school teachers with experience participating in the online National Teacher Training Program. Initial participants were recruited from rural schools across multiple regions to capture variation in teaching contexts and training experiences. As data collection and analysis progressed, subsequent participants were selected based on emerging analytical categories to refine and elaborate key themes.

Sampling continued until theoretical saturation was reached—that is, when additional interviews no longer generated substantively new insights. The final sample consisted of 30 rural primary and secondary school teachers from Shaanxi, Guizhou, Xinjiang, Gansu, and Hainan, representing a diverse range of geographical and institutional contexts.

Data Collection

During data collection, interviews and preliminary coding were conducted concurrently. Two researchers with educational backgrounds engaged in ongoing comparative analysis to identify emerging concepts and themes. As the interviews progressed, the research team observed that additional data primarily reiterated or elaborated existing themes rather than generating substantively new categories. To further confirm this observation, data collection was continued.

Based on the cumulative analysis, the interview data were considered to have sufficiently captured the core experiences and perspectives relevant to the research questions, and theoretical saturation was therefore determined. The final dataset consisted of 1,077 minutes of audio collected from 30 rural primary and secondary school teachers, amounting to approximately 130,000 Chinese characters of qualitative data.

Data Coding

The coding process was facilitated using NVivo 11 software. A bottom-up inductive approach was adopted, involving open coding, axial coding, and selective coding to develop hierarchical nodes (Chen, 2000). To protect participants' privacy, all participants took part in the study on a voluntary basis and provided informed consent prior to data collection. Personal identifiers were removed, and pseudonyms (e.g., T+ Number) were used in all transcripts and reports. All data were used solely for research purposes and handled in accordance with standard ethical guidelines to ensure confidentiality.

Coding was conducted by a research team of five members, including three doctoral students in educational technology and two senior experts. The coding framework was developed iteratively through training, pilot coding, and two rounds of team discussion based on the research questions. To ensure coding consistency, two interview transcripts were first pilot-coded in NVivo, and discrepancies were discussed to refine coding decisions. Formal coding was then carried out with ongoing memo writing and iterative refinement. Inter-

coder reliability was assessed through independent coding and yielded a Cohen's kappa coefficient of 0.907, indicating high consistency.

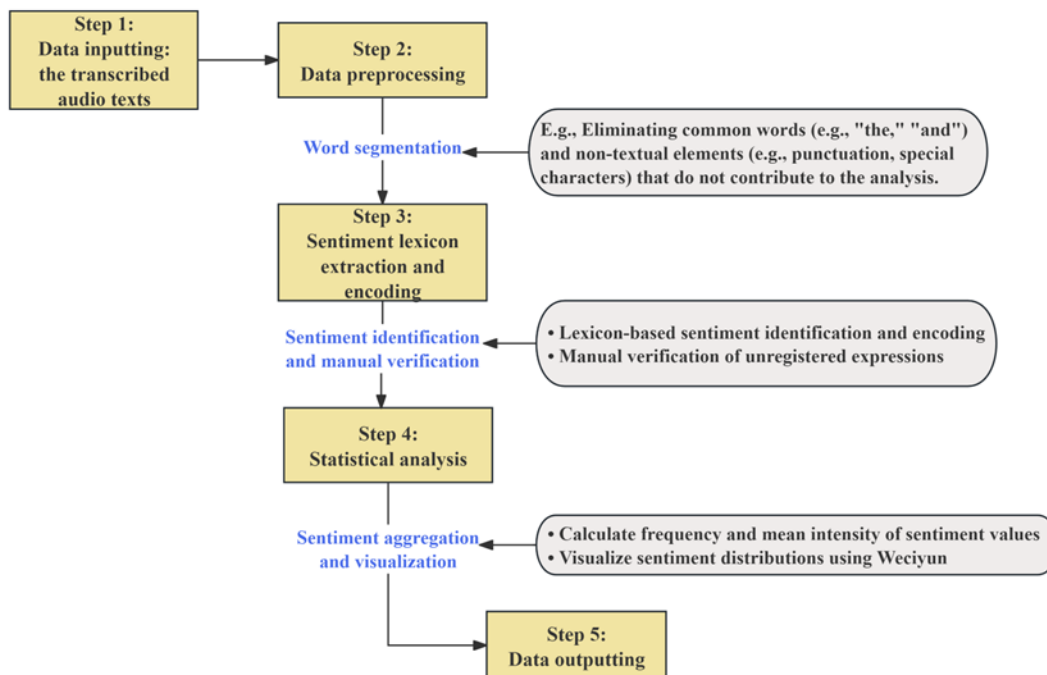
Data Analysis

To ensure the methodological validity, this study employed the Dalian University of Technology Chinese Sentiment Lexicon (DUTIR), which has been widely used in Chinese text-mining research and demonstrates high inter-annotator reliability (Cohen's $\kappa > 0.86$). DUTIR is particularly suitable for analyzing Chinese interview data, as it incorporates context-sensitive affective expressions commonly used in everyday and professional discourse, thereby reducing semantic bias associated with translation-based lexicons (Xu et al., 2008).

Interview transcripts were first preprocessed through word segmentation, and non-textual elements were removed. Sentiment-bearing words were then identified and encoded based on DUTIR-defined polarity and intensity values, with sentiment scores automatically generated by the lexicon. Colloquial or unregistered expressions frequently used by rural teachers were manually reviewed and supplemented following the same coding principles. Sentiment scores were then aggregated to examine frequency distributions and mean affective intensity. Rather than serving as a standalone analytic technique, sentiment analysis was used to identify overall affective tendencies and to triangulate qualitative themes derived from grounded theory analysis (Pang & Lee, 2008). The detailed implementation process is outlined in Figure 1.

Figure 1

Sentiment Analysis Implementation Process



Qualitative data were analyzed using NVivo 11 software, following a grounded theory approach involving open, axial, and selective coding (Chen, 2000). Coding was conducted by a five-member research team comprising three graduate students and two senior experts in educational technology. Prior to formal analysis, all coders received training and collaboratively developed an initial coding framework based on the research questions. Twenty-seven transcripts were first subjected to open coding to generate preliminary categories related to motivation and perceived barriers (Tables 1 and 2), which were then synthesized into higher-order themes through axial coding. Selective coding was subsequently used to integrate core categories into a cohesive analytical framework (Figure 2). Coding consistency was ensured through pilot coding of two transcripts and inter-coder reliability testing, yielding a Cohen's kappa coefficient of 0.907, indicating high reliability.

Table 1

Rural Teachers' Online Training Motivations

Open coding	Reference point	Example quotation	Axial coding	Selective coding	Reference point
Task requirements	32	"We participated because we were required to by our superiors, like leaders. It is a task."	External requirements		
Administrative demands	90	"Additionally, since I am a school administrator, this learning is indispensable for my role."			
Career advancement	19	"This [the online training] will also be beneficial when undergoing future job title evaluations or promotions."			
Resource needs	12	"Online training lets us watch recorded classes by experts and top teachers, and learn from their demo lessons. These help us improve our teaching ideas and methods."	External need	Extrinsic motivation	174
Era-driven	15	"Times are changing, and I believe I should keep up with the trends and become a lifelong learner as a teacher, especially in this area."			
Research demands	6	"Every year, our school requires us to submit research topics, but as rural teachers, we lack guidance and			

Open coding	Reference point	Example quotation	Axial coding	Selective coding	Reference point
		experience in this area. Therefore, we want to learn more about how to successfully submit research proposals.”			
Enhance teaching effectiveness	3	“On a personal level, improving my knowledge and skills can enhance my future teaching.”			
Personal growth	17	“It’s been nearly 10 years since I graduated, and to stop learning could hinder my future development.”	Sense of achievement		
Problem-solving	13	“By participating in online training, such as listening to expert teachers’ classes, and comparing these with my own teaching practices, I can continuously improve my methods and address real classroom issues.”		Intrinsic motivation	52
Passion	5	“I feel that my strong interest in teaching fuels my desire to learn more.”	Interest		
Grand vision	14	“I aim to acquire management techniques that will enable me to contribute to the mid- to long-term strategic planning for our school’s growth and development.”	Value		

Table 2

Rural Teachers’ Online Training Barriers

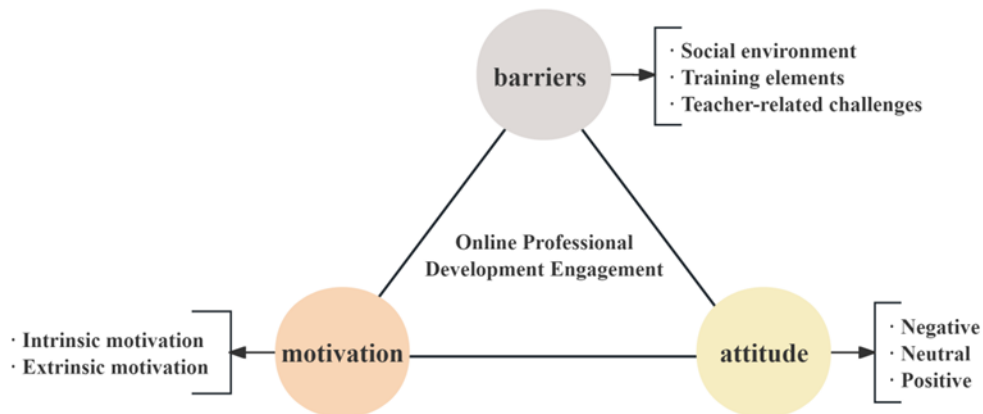
Open coding	Example quotation	Axial coding	Selective coding	Reference point
Lack of professional guidance and interaction	“When I had questions, I wanted to find someone to explain where I went wrong, but there was no one available to answer in the online training.”	Supporting training services	Training elements	184

Open coding	Example quotation	Axial coding	Selective coding	Reference point
Practical training on software and hardware	"We often administer exams to students and wanted to use the Baiyi assessment software, but after learning it, we still find it difficult to operate."			
Lack of supervision	"The online learning atmosphere is not as strong, and there is no supervision, which makes studying more tiring."			
Lack of communication with other learners	"During online training sessions, you may add other teachers to your contact list, but in most cases, you don't usually follow up or stay in touch afterward."			
Lack of evaluation	"When it comes to evaluations, I feel that most of the training sessions I've attended didn't include evaluations."			
Validity of evaluation criteria	"Being evaluated through a final paper at the end of an online course is quite challenging for me."			
Lack of practicality	"Some of the content covered may not be particularly applicable to rural teachers, and its practicality is relatively weak."			
Difficult training content	"There is a lot of theoretical content in online courses, which feels somewhat abstract and difficult to understand."	Training content and schedule		
Too general	"The downside is that online training is mainly designed for the general audience, not tailored to individual needs."			
Scheduling conflicts	"The training sessions are scheduled to start at 9 a.m., which coincides with our school's class time, causing a conflict with our work."			
Unavailable classroom software/hardware	"Some tools may not be usable due to the school's hardware limitations."	School environment	Social environment	62
Other trivial matters	"Besides teaching, some teachers also handle other roles, such as teaching multiple			

Open coding	Example quotation	Axial coding	Selective coding	Reference point
	subjects, managing logistics, overseeing safety, maintaining school websites, and organizing events.”			
Having elderly dependents and children	“For example, during summer training, I was mostly at home. With other family members around, like children, there were quite a lot of distractions.”	Family environment		
Loneliness	“Since online learning is done through a screen, it often lacks the immersive atmosphere of in-person learning. I felt lonely.”			
Guilt	“The requirements for teachers’ self-discipline are low. I felt guilty as I did not learn all the courses. Sometimes, I did other things. Actually, you know, the computer is the learner.”	Teachers’ emotional experience		
Frustration	“At the beginning of the training, when multiple training tasks piled up all at once, I felt extremely frustrated.”		Teacher-related challenges	89
Lack of basic digital skills	“I really want to learn about multimedia and information technology, but honestly, I feel I’m not very good at these areas.”	Teachers’ digital competence		
Difficulties integrating technology and teaching	“As a rural art teacher, I feel that my ability to integrate modern online resources with art teaching is relatively weak.”			

Figure 2

Analytical Framework for Rural Teachers' Online Learning Attitudes, Motivations, and Influencing Factors



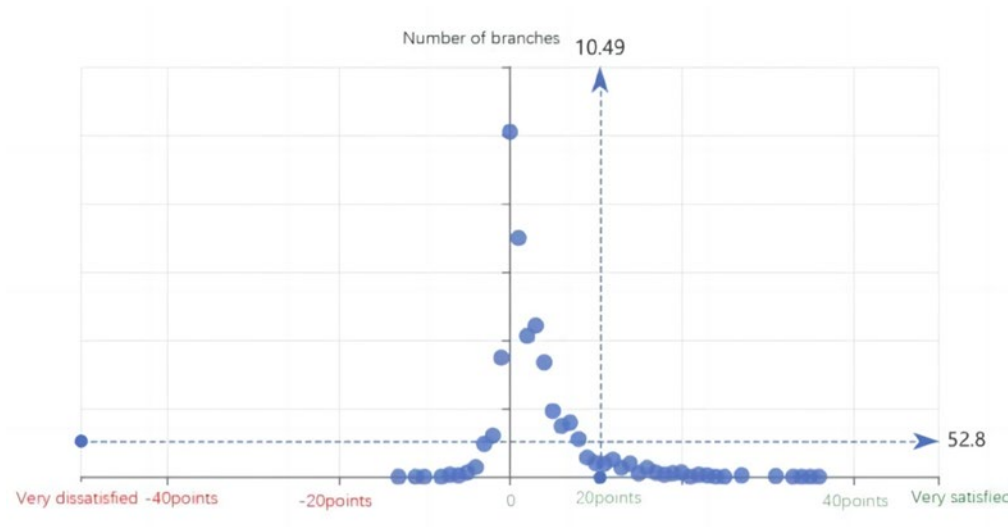
Results

Positive Perceptions, Contradictory Experiences

Lexicon-based sentiment analysis revealed that rural teachers' overall emotional orientation toward OPD was generally positive. Positive sentiment accounted for 45.4% of all coded segments, compared with 27.57% negative and 27.03% neutral sentiment. The mean sentiment score was 10.49 (see Figure 3), exceeding the zero baseline, indicating an overall favorable emotional response to online training (Lin et al., 2008). Frequently occurring positive keywords included “learning,” “improvement,” “gain,” and “satisfaction,” which were later echoed in qualitative accounts describing access to expert resources and perceived opportunities for professional updating, suggesting that participants recognized the convenience and potential for instructional enhancement afforded by the OPD.

Figure 3

Distribution of Emotional Values and Quantities



Despite the generally positive affective orientation, a substantial portion of the interview data expressed dissatisfaction and frustration. Negative sentiment accounted for 27.57% of coded statements, with high-frequency keywords including “lack,” “problems,” and “insufficient.” These terms reflected teachers’ dissatisfaction with training effectiveness, guidance, interaction, time allocation, and practical relevance. This discrepancy between optimistic expectations and suboptimal experiences echoes the aspiration–reality gap identified in technology-mediated professional development (Zheng et al., 2020).

Critically, the negative perceptions reflected in high-frequency keywords such as “lack” and “insufficient” were consistently echoed in qualitative narratives describing deficiencies in training design, interaction opportunities, and institutional arrangements, pointing to limitations in the design and implementation of the online learning system, which is consistent with international findings on rural online education.

This tension indicates a misalignment between teachers’ initial expectations and their actual training experiences. Sentiment analysis revealed a surface-level optimism that coexisted with deeper emotional dissatisfaction, particularly in relation to content relevance, interaction quality, and logistical feasibility, which constrained meaningful engagement with online learning.

Although rural teachers expressed generally positive attitudes toward OPD, their reported experiences revealed notable affective dissonance, suggesting a misalignment between program design and the contextualized needs of rural educators.

External Mandates Over Internal Drive: A Motivation Profile

Thematic analysis of interview data revealed that rural teachers’ participation in the OPD was predominantly driven by extrinsic motivations, with intrinsic motivation comparatively weak. Extrinsic motivation accounted for 174 coding instances across all 30 participants, compared with 52 instances of intrinsic motivation, indicating a pattern of compliance-oriented rather than autonomous engagement

(Table 1).

A large proportion of participants indicated that their involvement in the OPD was due to administrative directives or performance evaluation requirements. For example, one teacher stated, “We are expected to complete these trainings because they’re part of our job assessments.” Others described the training as “an obligatory task assigned by superiors” or emphasized its role in promotion or title evaluation. These comments reflect a compliance-based motivational structure, consistent with SDT’s notion of externally regulated behavior (Deci & Ryan, 2000).

Beyond institutional mandates, participants’ motivations were often linked to access to external resources and professional visibility. Teachers reported learning from model courses, watching expert demonstration lessons, and accessing standardized teaching materials. As one participant noted, “It’s not about interest—it’s about survival. Online training is our only way to learn from experts since we don’t have those resources locally.” These accounts highlight the resource-compensatory role of online training in geographically isolated rural contexts.

In contrast, intrinsically motivated responses were relatively rare and primarily focused on long-term professional growth or addressing instructional challenges. A few participants mentioned their desire to improve teaching effectiveness or keep up with educational trends. For instance, one teacher remarked, “I want to understand how others manage their classrooms so I can refine my own practices.” Another noted, “I enjoy teaching, so I’m curious to learn more—even if it’s outside of requirements.” These voices, however, were significantly less frequent and often overshadowed by externally imposed goals.

The imbalance between external obligation and internal aspiration reflects a low-autonomy motivational ecology in which teachers engage in the OPD not as self-driven professionals but as institutional actors responding to policy pressure. This motivational configuration helps explain why generally positive or neutral affective orientations identified in the sentiment analysis do not translate into sustained or autonomous engagement.

This configuration poses risks for sustained engagement and meaningful learning transfer, as policy-driven, bureaucratic framings of professional development may undermine internalization, reflection, and teacher agency.

Barriers to Meaningful Participation: Structural, Emotional, and Institutional Dimensions

Thematic analysis identified three interrelated dimensions of barriers shaping rural teachers’ participation in OPD. These obstacles can be categorized into three overarching dimensions: (a) training-related design and delivery issues, (b) sociocultural and institutional constraints, and (c) individual emotional and digital challenges.

The most frequently cited barriers were related to training content, format, and scheduling (184 coded references). Many teachers reported that online training content was overly theoretical or abstract or insufficiently contextualized. As one teacher put it, “It felt like the courses were designed for urban schools or general audiences, not for the needs of rural teachers like us.” Others emphasized the lack of interactivity

and practical application: “Prerecorded videos have no feedback. You just watch and click. That’s it.” In addition, rigid schedules often conflicted with regular teaching hours, leaving teachers unable to fully engage: “Training sessions often take place during our working time, so we either skip them or multitask.”

A second layer of barriers emerged from the sociocultural and institutional environments in which teachers work (62 references). These included insufficient infrastructure in schools, multiple administrative responsibilities, and family-related distractions. One teacher shared, “During summer training, I was home with my children and parents and it was hard to focus at all.” Another noted, “Besides teaching, I also have to manage safety, logistics, and tech support at school.” Such overlapping responsibilities and lack of dedicated time or space for learning made sustained participation difficult.

Finally, teachers described a range of emotional and competence-related challenges (89 references), including feelings of loneliness, guilt, and technological inadequacy. Some teachers expressed frustration at their own digital limitations: “I know I need to learn multimedia tools, but I just don’t know how.” Others reported low emotional engagement due to the impersonal nature of the learning environment: “I felt invisible, like I wasn’t really part of anything.” These experiences align with broader concerns around emotional presence and self-regulation in online learning, particularly in under-supported professional contexts.

Collectively, these barriers constitute a layered constraint ecology in which structural, social, and personal factors interact to suppress meaningful engagement and help explain the neutral and negative affect identified in the sentiment analysis. From an SDT perspective, these ecology constraints are related to autonomy, competence, and relatedness, increasing the risk that OPD participation becomes compliance-oriented rather than growth-oriented (Deci & Ryan, 2000).

Discussion

Affective-Structural Dissonance

While the sentiment analysis indicated rural teachers’ generally positive affective orientation toward OPD, qualitative data revealed a contradiction between teachers’ optimistic expectations and their frustrated training experiences. We conceptualize this tension as affective-structural dissonance, reflecting a misalignment between emotional expectations and the structural conditions of learning, whereby aggregate positive affect conceals deeper frustrations related to training design and institutional arrangements.

This dissonance was most evident in neutral and negative expressions related to training design, relevance, and interactivity. Although OPD was often described as “useful in theory,” it was simultaneously characterized as irrelevant or disconnected from practice, reflecting emotional disengagement and signaling deeper structural misalignments between policy-driven training and teachers’ lived needs in resource-constrained contexts.

This finding echoes Zheng et al.’s (2020) identification of the aspiration–reality gap in technology-mediated professional development, where teachers’ expectations of learning benefits are frequently undercut by insufficient support or contextual misalignment. It also resonates with Borup et al.’s (2020) emphasis on the role of social presence in online learning engagement, something our participants consistently found

lacking. Moreover, the sense of technostress, time pressure, and fragmented attention described in our data aligns with Veletsianos's (2020) critique of OPD under constrained conditions.

As a conceptual contribution, affective-structural dissonance highlights the gap between teachers' emotional readiness for and institutional capacity of OPD, shifting attention from technological access to the emotional and contextual alignment between training environments and teacher needs, with implications for more motivation-sensitive professional development design.

Controlled Engagement

Findings indicate that rural teachers' engagement in OPD is largely shaped by externally regulated motives, such as administrative mandates and evaluation requirements, leaving limited space for autonomous interest. As one teacher succinctly stated, "We join these training because we have to—not because we feel inspired." This helps explain why positive or neutral affect tendencies do not translate into sustained or autonomous engagement.

This pattern further reflects what SDT terms *controlled regulation*, a form of extrinsic motivation in which behavior is enacted to satisfy external demands or avoid negative consequences rather than to fulfill personally endorsed goals (Deci & Ryan, 2000). Many participants described the OPD as a "task," a "requirement," or a "non-negotiable part of appraisal," underscoring the instrumental role that institutional expectations play in shaping participation.

Controlled engagement eroded teachers' autonomy, undermining their capacity to internalize learning goals and sustain meaningful participation. Teachers frequently reported having no choice over course content, delivery format, or training schedules. This lack of control restricted their ability to connect the training to their classroom realities or professional identities. Several interviewees admitted to multitasking during sessions or completing requirements superficially, reinforcing the gap between presence and authentic learning. These findings are consistent with prior work (Zhang & Wang, 2022), noting that rural teachers often treat OPD as a compliance-oriented task due to policy saturation and rigid bureaucratic structures. Recent research in China further supports this tendency. Liang and Ding (2022) argued that rural teachers' online learning remains highly dependent on external supervision and lacks personalized guidance, making internal motivation difficult to sustain. Gorozidis et al. (2020), examining teacher networks from an SDT perspective, also found that online participation driven by external accountability alone failed to generate lasting motivation or community belonging. These challenges are compounded in development-oriented systems where digital training is often standardized, assessment-driven, and unresponsive to local realities (Bozkurt & Sharma, 2021). Furthermore, even when online training environments are accessible, surface-level activity may not reflect authentic engagement. As Martin and Bolliger (2018) emphasize, high enrollment or log-on rates do not equate to emotional or cognitive investment, particularly in contexts where participation is driven by obligation rather than aspiration.

The prevalence of controlled engagement reflects a motivational ecology in which external compliance outweighs internal agency, raising concerns about the sustainability and pedagogical impact of OPD initiatives that achieve high participation without meaningful engagement.

Layered Constraint Ecology

Beyond individual motivation, this study reveals that rural teachers' participation in OPD is shaped by a layered ecology of structural and emotional constraints. Thematic analysis identified three interrelated levels of barriers: (a) training design and delivery limitations, (b) sociocultural and institutional constraints, and (c) personal emotional and digital challenges. These barriers do not operate in isolation; rather, they interact to form a constraint ecology that suppresses engagement and diminishes learning outcomes.

At the instructional level, teachers cited abstract content, poor interactivity, inflexible schedules, and irrelevant assessment mechanisms as major challenges. These frustrations are consistent with previous findings that online training for rural teachers often lacks contextualization and user-centeredness (Kong & Kang, 2022; Zhang & Wang, 2022). At the sociocultural level, participants reported workload overload, infrastructure shortages, and family-related distractions, challenges that have been shown to significantly affect rural teachers' ability to engage in online learning (Liang & Ding, 2022). At the personal level, teachers described feelings of guilt, loneliness, and technological inadequacy. These emotional states reflect broader concerns about self-regulation in online learning and the psychological distance that can emerge in low-support environments (Rovai, 2002).

These findings suggest that barriers to the OPD are ecological rather than isolated. From an SDT perspective, this ecology obstructs the satisfaction of autonomy, competence, and relatedness. Accordingly, the neutral and negative affective expressions identified in the sentiment analysis can be understood as emergent outcomes of this layered constraint ecology rather than as isolated emotional reactions.

This layered constraint ecology explains why even motivated rural teachers struggle to sustain meaningful engagement, as accumulating structural, social, and emotional barriers weaken the internalization of professional goals and call for OPD designs that better account for teachers' learning contexts.

To sum, although this study is situated in the Chinese rural context, the challenges identified are not unique to China. Prior research on rural and resource-constrained OPD in other regions has similarly documented policy-driven participation, limited teacher autonomy, and emotional disengagement (Borup et al., 2020; Bozkurt & Sharma, 2021). Within established ODL frameworks, these challenges have often been discussed in terms of weakened social presence and constrained learner autonomy.

From this perspective, the three conceptual constructs proposed in this study—*affective-structural dissonance*, *controlled engagement*, and *layered constraint ecology*—can be understood as analytical lenses for interpreting how structural conditions, motivational regulation, and affective experiences interact in rural ODL settings across developing contexts. Specifically, *affective-structural dissonance* deepens discussions of social presence by revealing how emotional expectations remain unmet when institutional and interactional structures fail to support meaningful engagement, while *controlled engagement* extends debates on learner autonomy by illustrating how externally mandated participation limits the internalization of learning goals despite high levels of formal involvement.

Rather than being context-specific labels, these constructs highlight shared mechanisms that shape teachers' engagement in OPD under constrained conditions, offering a theoretically grounded framework for interpreting rural ODL challenges beyond the Chinese context.

Conclusion

Using a mixed-methods approach combining sentiment analysis and qualitative interviews, this study found that although rural teachers expressed generally positive attitudes toward OPD, their lived experiences revealed emotional and structural tensions characterized by surface-level optimism, externally driven participation, and multilayered design, institutional, and personal barriers.

This study makes three conceptual contributions. First, affective-structural dissonance captures the misalignment between teachers' emotional readiness and the institutional conditions of OPD. Second, controlled engagement describes participation driven primarily by external mandates rather than internalized motivation. Third, layered constraint ecology explains how interconnected pedagogical, sociocultural, and emotional barriers jointly suppress sustained engagement. Together, these constructs extend ODL frameworks by highlighting the interaction of affect, motivation, and structure in rural professional development.

From a practical perspective, these findings suggest that rural-appropriate OPD systems should move beyond a narrow focus on access and content provision. Design efforts need to account for the emotional and contextual compatibility of learning environments. Specifically, OPD initiatives should support teachers' autonomy through greater choice and flexibility, enhance competence through context-sensitive content and guidance, and foster relatedness through meaningful interaction and social presence. At the policy level, this implies a shift away from purely compliance-driven participation toward more developmental and teacher-centered approaches. Platform designers are encouraged to integrate interactional and affective affordances rather than relying solely on standardized delivery models. Teacher educators, in turn, play a critical role in supporting the internalization of professional learning goals and helping teachers connect online training with their professional identities.

Several limitations should be acknowledged. This study is situated within the Chinese rural context, which may constrain the generalizability of the findings to other national or institutional settings. In addition, the cross-sectional nature of the data limits insight into how teachers' motivations and emotional experiences evolve over time. Future research could address these limitations by adopting longitudinal designs, conducting cross-context or cross-national comparisons, and examining how targeted design interventions may reduce affective-structural dissonance and foster more autonomous forms of engagement.

Future research should further build on the conceptual framework proposed in this study. Longitudinal studies are needed to examine how rural teachers' affective experiences and motivational regulation evolve over time in OPD contexts. Comparative research across different rural or resource-constrained settings may help assess the transferability of the proposed constructs, including affective-structural dissonance, controlled engagement, and layered constraint ecology. In addition, design-based research could explore how targeted interventions in platform design, instructional support, or policy arrangements may reduce affective-structural dissonance and foster more autonomous and sustained forms of teacher engagement.

References

- Batmang, B., Sultan, M., Azis, A., & Gunawan, F. (2021). Perceptions of pre-service teachers on online learning during the COVID-19 pandemic. *International Journal of Education in Mathematics, Science and Technology*, 9(3), 449–461. <https://doi.org/10.46328/ijemst.1595>
- Borup, J., Graham, C. R., West, R. E., Archambault, L., & Spring, K. J. (2020). Academic communities of engagement: An expansive lens for examining support structures in blended and online learning. *Educational Technology Research and Development*, 68, 807–832. <https://doi.org/10.1007/s11423-020-09744-x>
- Bozkurt, A., & Sharma, R. C. (2021). On the verge of a new renaissance: Care and empathy oriented, human-centered pandemic pedagogy. *Asian Journal of Distance Education*, 16(1), i–vii. <https://www.asianjde.com/ojs/index.php/AsianJDE/article/view/576>
- Chen, X. (2000). *Research methods and social science research*. Educational Science Publishing House.
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute. <https://doi.org/10.54300/122.311>
- Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, 19(2), 109–134. [https://doi.org/10.1016/0092-6566\(85\)90023-6](https://doi.org/10.1016/0092-6566(85)90023-6)
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01
- Gorozidis, G. S., Tzioumakis, Y. S., Krommidas, C., & Papaioannou, A. G. (2020). Facebook group PETCoN (Physical Education Teacher Collaborative Network). An innovative approach to PE teacher in-service training: A self-determination theory perspective. *Teaching and Teacher Education*, 96, Article 103184. <https://doi.org/10.1016/j.tate.2020.103184>
- Haider, S., Moin, L., & Haider, H. (2025). Sustaining professional competence in resource-constrained contexts: Perspectives of university teachers in Balochistan. *Journal of Political Stability Archive*, 3(4), 612–628. <https://doi.org/10.63468/jpsa.3.4.37>
- Kong, X., & Kang, H. (2022). Research on influencing factors and strategies for the effectiveness of remote training for rural teachers: An analysis based on grounded theory. *Educational Academic Monthly*, 363(10), 106–112.
- Kuo, Y.-C., Walker, A. E., Schroder, K. E. E., & Belland, B. R. (2014). Interaction, Internet self-efficacy, and self-regulated learning as predictors of student satisfaction in online education courses. *The Internet and Higher Education*, 20, 35–50. <https://doi.org/10.1016/j.iheduc.2013.10.001>

- Lennert da Silva, A. L., & Mølstad, C. E. (2020). Teacher autonomy and teacher agency: A comparative study in Brazilian and Norwegian lower secondary education. *The Curriculum Journal*, 31(1), 115–131. <https://doi.org/10.1002/curj.3>
- Liang, L., & Ding, S. (2022). Research on the construction of a learning support system for rural teachers' online open courses based on the "U-G-S" tripartite collaboration. *Digital Education*, 8(1), 54–60.
- Lin, Y. M., Lin, G. Y., & Laffey, J. M. (2008). Building a social and motivational framework for understanding satisfaction in online learning. *Journal of Educational Computing Research*, 38(1), 1-27. <https://doi.org/10.2190/EC.38.1.a>
- Martin, F., & Bolliger, D. U. (2018). Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment. *Online Learning*, 22(1), 205–222. <https://doi.org/10.24059/olj.v22i1.1092>
- Pang, B., & Lee, L. (2008). *Opinion mining and sentiment analysis*. Now Publishers Inc.
- Ravet, J., & Mtika, P. (2024). Educational inclusion in resource-constrained contexts: A study of rural primary schools in Cambodia. *International Journal of Inclusive Education*, 28(1), 16–37. <https://doi.org/10.1080/13603116.2021.1916104>
- Rovai, A. P. (2002). Building sense of community at a distance. *The International Review of Research in Open and Distributed Learning*, 3(1). <https://doi.org/10.19173/irrodl.v3i1.79>
- Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, Article 101860. <https://doi.org/10.1016/j.cedpsych.2020.101860>
- Subekti, A. S. (2020). COVID-19-triggered online learning implementation: Pre-service English teachers' beliefs. *Journal of English Language, Literature, and Teaching*, 4(3), Article 3. <https://doi.org/10.31002/metathesis.v4i3.2591>
- Sun, J. C. Y., & Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education. *British Journal of Educational Technology*, 43(2), 191–204. <https://doi.org/10.1111/j.1467-8535.2010.01157.x>
- UNESCO. (2023). Global education monitoring report 2023: Technology in education: A tool on whose terms? <https://unesdoc.unesco.org/ark:/48223/pf0000385723>
- Veletsianos, G. (2020). *Learning online: The student experience*. Johns Hopkins University Press. <https://muse.jhu.edu/book/73824>
- Wang, G., Li, J., & Lai, C. (2022). A meta-analysis of factors influencing the quality of online teacher training in China. *Distance Education Journal*, 40(4), 82–91.

- Xu, L., Lin, H., & Zhao, J. (2008). Construction and analysis of an emotional corpus. *Journal of Chinese Information Processing*, 01, 116-122. <https://doi.org/10.3969/j.issn.1003-0077.2008.01.019>
- Yan, H., & Shan, J. (2020). From training to empowerment: Blueprint construction for teacher professional development in the post-pandemic era. *E-Learning Research*, 41(6), 13–19.
- Zhang, W., & Wang, Y. (2022). Digital divide and rural teachers' online professional development: Challenges and coping strategies. *Journal of Educational Technology & Society*, 25(3), 109–123.
- Zheng, B., Lin, C. H., & Kwon, J. B. (2020). The impact of learner-, instructor-, and course-level factors on online learning. *Computers & Education*, 150, Article 103851. <https://doi.org/10.1016/j.compedu.2020.103851>



August – 2026

An Exploratory Factor Analysis of the Online Information Literacy Self-Efficacy (OILS) Instrument

Wei Zakharov¹, Anne Traynor², and Clarence Maybee¹

¹Purdue University, Libraries and School of Information Studies, West Lafayette, USA; ²Purdue University, College of Education, Educational Psychology & Research Methodology, West Lafayette, USA

Abstract

The purpose of this study was to develop and validate an effective instrument to measure undergraduate students' self-efficacy in engaging with information for online coursework. Recognizing the increasing importance of information literacy in online learning environments, the Online Information Literacy Self-Efficacy (OILS) instrument was designed to assess key competencies associated with students' abilities to engage with information effectively. The validity and reliability of the OILS instrument were examined using exploratory factor analysis (EFA) and reliability analysis. The instrument initially comprised 30 items, grounded in two authoritative frameworks: the Association of College and Research Libraries (ACRL) Framework for Information Literacy and the Association of American Colleges and Universities (AAC&U) Information Literacy VALUE Rubric. Survey responses from 259 undergraduate students enrolled in online courses at a large Midwestern US R1 institution were analyzed using EFA and reliability analysis. EFA results revealed a four-factor structure explaining substantial variance in the item response patterns: scoping research topics, obtaining information, evaluating information quality and producing research documents, and crediting sources. These four factors demonstrated high reliability, with omega coefficients ranging from 0.891 to 0.960. One item was recommended for removal due to cross-loadings and skewness. Exploratory measurement invariance analyses suggested that the factor structure was comparable for students with and without prior online learning experience. The findings support the OILS instrument's utility in measuring students' confidence in applying information literacy skills in online learning contexts, offering valuable insights for educators and librarians to enhance instructional strategies and support student success.

Keywords: information literacy, self-efficacy, online learning, factor analysis, undergraduate education

Introduction

In today's increasingly digitized academic landscape, equipping students in higher education with the skills necessary to engage critically and effectively with information in online coursework is essential. Information literacy—defined as the ability to use information critically to learn and solve problems—has therefore emerged as a key competency in supporting students' academic success and lifelong learning (Bradley, 2013). As online learning platforms become the norm, students constantly face challenges in locating highly relevant information, discerning credible sources, synthesizing information, and applying it to their academic tasks.

An essential concept related to information literacy in online settings is self-efficacy, which significantly influences students' motivation and engagement. Self-efficacy, rooted in Bandura's (1997) social cognitive theory, refers to an individual's belief in their capacity to execute behaviors necessary to achieve specific performance outcomes. Honicke and Broadbent's (2016) systematic review demonstrated that higher self-efficacy empowers students to approach challenges with confidence, viewing obstacles as opportunities for learning rather than threats. This mindset can lead to greater persistence in problem-solving and more proactive engagement in learning activities, which are crucial in an online education context. Studies have shown that higher self-efficacy is linked to increased motivation, persistence, and overall academic performance (Adigue et al., 2023; Honicke & Broadbent, 2016). In the context of online learning, where students navigate information-rich but unstructured environments, self-efficacy for information literacy becomes a critical predictor of their ability to perform information-related tasks. Consequently, it is imperative to both assess and support the development of students' self-efficacy in online information literacy throughout their online learning experience.

Technology-enhanced online education also provides various opportunities for developing information literacy skills for online courses. Online learning platforms can leverage interactive tools and resources to promote active learning and collaboration among students. More recently, generative artificial intelligence (AI) tools have become embedded in many of these platforms, further mediating how students search for, evaluate, synthesize, and produce information (Giannakos et al., 2024; Kassorla et al., 2024; UNESCO, 2023). For instance, structured online modules emphasizing the critical evaluation of information sources, including both human-authored and AI-generated content, can be implemented, thereby enhancing the learning experience (Hovious, 2024; Johnston, 2010; McClellan, 2016). Furthermore, instructors can incorporate assessments that evaluate not only content mastery but also students' information literacy skills, which remain essential for exercising judgment, credibility assessment, and ethical information use in AI-augmented learning contexts, ensuring a comprehensive approach to education that prepares students for real-world challenges (Kassorla et al., 2024; McGrew et al., 2018; Weaver, 2024).

The purpose of this study is to develop an effective instrument to measure undergraduate students' perceived confidence in engaging with information for online coursework. Specifically, this study employed exploratory factor analysis (EFA) to empirically test the instrument and examine its psychometric properties. The study has two specific research questions:

1. Which set of items, determined through psychometric analyses, should be included in the final instrument to assess self-efficacy in online learning information literacy?

2. What is the reliability and validity evidence of the developed instrument?

Conceptual Framework

The conceptual framework guiding this study is informed by two authoritative resources in the field of information literacy. The first is the Association of College and Research Libraries (ACRL) Framework for Information Literacy for Higher Education (ACRL, 2015). This framework outlines six core concepts that shape students' understanding and practice of information literacy in academic contexts:

- Authority is constructed and contextual: Recognizes that information authority varies based on context and audience.
- Information creation as a process: Emphasizes the iterative nature of information production.
- Information has value: Highlights the ethical and economic considerations tied to information use.
- Research as inquiry: Frames research as an ongoing, iterative process of questioning.
- Scholarship as conversation: Promotes the view of academic work as a dialogue among scholars.
- Searching as strategic exploration: Encourages strategic and adaptable information-seeking behaviors.

Collectively, this framework establishes a holistic view of information literacy as a set of interconnected, critical-thinking practices essential for student success in higher education.

Complementing the ACRL framework is the Association of American Colleges and Universities' (AAC&U) Information Literacy VALUE Rubric (AAC&U, 2009), which provides a measurable, skills-based approach to assessing information literacy competencies. The VALUE Rubric includes five key dimensions:

- Determine the extent of information needed: Focuses on defining the scope of research inquiries.
- Access the needed information: Ensures students can efficiently retrieve relevant sources.
- Evaluate information and its sources critically: Encourages discernment and reliability checks.
- Use information effectively to accomplish a specific purpose: Emphasizes the application of information toward academic goals.
- Access and use information ethically and legally: Underscores the importance of academic integrity and compliance with legal guidelines.

Together, these two foundational frameworks provide both the theoretical and practical bases for the development and validation of the Online Information Literacy Self-Efficacy (OILS) instrument factors used in this study.

Methods

Participants

The study sample consisted of undergraduate students enrolled in 40 online courses across 73 sections at a Midwestern US R1 university. A total of 1,832 undergraduate students were invited to participate. Institutional review board approval was obtained prior to data collection. All the online courses' instructors were informed via email regarding the purpose of the study. Subsequently, all students received an email invitation to participate, followed by two reminder emails, which were distributed by the university's institutional registrar. Twenty percent of the students who completed the survey were randomly drawn and each received an incentive of a \$5 Amazon e-gift card.

A total of 354 students submitted survey responses. Of these, 259 students completed at least part of the information literacy scale and were included in the final analysis. Among these, 215 were domestic students and 43 were international students (one participant declined to respond about their national origin). In terms of the academic levels, 5 students had freshman status, 53 had sophomore status, 90 had junior status, and 111 students had senior status. Additionally, 67 students reported that they were enrolled in an online class for the first time, while 192 students had taken at least one online class previously.

Instrumentation

Based on the two framework sources, the project team initially defined the construct of information literacy for online learning to consist of five main ability components: (a) determining the extent of information needed to complete a research assignment, (b) accessing the information needed, (c) evaluating obtained information critically, (d) accessing and crediting sources, and (e) producing the assignment submission. Two librarian authors independently reviewed the draft item set for content alignment and clarity, settling on 30 items to include in the initial version of the information literacy instrument. The items directed participants to self-rate various aspects of their information literacy on a 5-point Likert scale: 1 = almost always true, 2 = often true, 3 = sometimes true, 4 = usually not true, and 5 = almost never true. The complete text of the items is provided in Appendix A. The survey instrument contained an initial section requesting basic demographic information, followed by two psychometric scales: the information literacy self-efficacy scale composed of 30 items, and validation scale consisting of 8 multiple-choice information literacy knowledge items (Appendix A). The survey was administered using an online platform (Qualtrics).

Analytic Methods

To analyze evidence about the meaning of total scores from the scale, we conducted traditional item analysis, EFA, and measurement invariance analysis. We first computed item analysis statistics: the mean, standard deviation, and skewness for each item's rating scale responses. We used these statistics to identify any items that showed little variation in the population, had rating scale points that were too difficult or easy to endorse, or had a badly skewed response distribution (which would violate a key assumption of EFA, the next step).

After examining item descriptive statistics as a screening step, we applied EFA to the item response data. The goals of EFA were to identify and describe any latent factor(s) formed by patterns in the item response data and to compare the empirical latent factor structure to our proposed five-factor measurement model.

Because any traits underlying responses to the information literacy self-efficacy items seemed likely to be interrelated, an oblique rotation, geomin, was used to find a simple-structure solution for each specified number of factors. Model parameters were estimated using maximum likelihood. We used criteria suggested by Little (2013) for the root mean square error of approximation (RMSEA) and the Tucker–Lewis index (TLI) to judge model fit. Then, we selected the most plausible measurement model for the observed data considering theory about information literacy, statistical model fit indices, and the principle of parsimony.

Once we had mapped items to underlying latent factors, we estimated the reliability of the total score on each factor using coefficient omega (McDonald, 1970). Coefficient omega is a suitable score reliability coefficient when factor loadings vary across items. To address concerns that the underlying factor structure might differ between students taking their first online course and students who had previously taken online courses, we examined the factor structure separately within these two groups. Because the scale did not yet have a well-established factor structure (i.e., replicated in an independent sample), we again used EFA modeling, now within these two subgroups of respondents, to produce a tentative assessment of the scale's measurement invariance across these respondent groups. Finally, we computed the correlation between each latent factor underlying the information literacy self-efficacy scale, and the total score on the eight information literacy knowledge items, as a form of external evidence for the meaning of scores from the information literacy self-efficacy scale. All analyses were conducted using Mplus software.

Results

Item Analysis

Descriptive statistics for responses to the 30 information literacy self-efficacy rating scale items are reported in Table 1. The mean response of these questions ranges from 1.582 to 2.413. Given the rating scale labels, the more strongly respondents tended to agree with an item statement, the lower its mean value would be. The lowest mean value (1.582) was observed for the item about citing all reference sources. The highest mean (2.413) was for the item about searching using keywords and Boolean logic. The range of mean values suggests respondents generally reported being able to carry out the items' information literacy behaviors on most, or at least some, occasions. Items had similar standard deviation values, and no item showed a low standard deviation that might raise questions about that item's utility for measuring information literacy. Items from the "accessing information needed" ability domain, some of which referred to library resources, displayed slightly higher response variability than other items. One item (10) about seeking "information on the Web" displayed notably high skewness (1.53), meaning the respondents tended to select answers from one end, without much variability or discrimination. The rate of missing data did not exceed 7% on any item.

Table 1

Descriptive Statistics for Survey Items

Item	<i>M</i>	<i>SD</i>	Skewness
1. Determine the information I need	1.804	0.77	0.55
2. Determine scope of research questions	1.976	0.83	0.49
3. Determine types of information sources	1.870	0.84	0.72
4. Conduct searches to frame topic	1.925	0.88	0.59
5. Use search box on library home page	2.162	1.25	0.88
6. Access electronic sources through library	2.340	1.21	0.65
7. Access topic-relevant databases	2.274	1.12	0.69
8. Determine where and how to find needed information	2.151	1.05	0.86
9. Search using keywords and Boolean logic	2.413	1.21	0.59
10. Seek information on the Web	1.640	0.91	1.53
11. Use Internet search tools (Google Scholar, etc.)	1.636	0.87	1.22
12. Recognize need to go beyond Internet resources	2.143	1.15	0.70
13. Evaluate currency of information	1.859	0.89	0.91
14. Evaluate relevance of information	1.795	0.84	0.95
15. Evaluate author's credentials	2.016	1.01	0.67
16. Evaluate accuracy of information	2.024	0.92	0.59
17. Evaluate intended purpose of information	2.008	0.97	0.79
18. Assess quality of Internet information resources	1.858	0.81	0.61
19. Avoid fake information on the Web	1.811	0.87	0.84
20. Improve quality of my information—seeking process	1.886	0.89	0.83
21. Synthesize information from a variety of sources	1.874	0.91	0.87
22. Use bibliographic reference management tools	2.251	1.15	0.65
23. Write documents	1.788	0.87	0.88
24. Create academic presentations	1.681	0.83	1.06
25. Disseminate information using social media	1.917	0.91	0.76
26. Identify information as published, confidential, or proprietary	1.788	0.87	0.93
27. Cite all my sources	1.582	0.79	1.26
28. Use quotations properly	1.707	0.84	1.07
29. Avoid plagiarism	1.584	0.78	1.27
30. Create a bibliography	1.770	0.93	1.20

Note. The text of the scale items is abbreviated for display purposes.

Exploratory Factor Analysis

Because the design specifications for the instrument consisted of five categories of information literacy skills, we examined results for EFA models containing between one and seven underlying latent factors. (For the most complex hypothesized model, with seven factors, the estimation process did not converge.) One factor in the six-factor model had no items that loaded highly ($> .30$) on it, so the loading pattern for the six-factor model was not substantively interpretable. The five-factor model contained four well-defined factors and a fifth factor on which only two items loaded highly. These two items, about ability to “seek information on the Web” (10) and to “use Internet search tools” (11), seemed to be measuring highly related behaviors. Because the response distribution of the item about seeking information on the Web also displayed high skewness, indicating that this item might be too easy in the target respondent population, we suggested eliminating that item from future versions of the instrument. Eliminating that item produced a more parsimonious four-factor EFA model with fit that we judged to be adequate but not excellent (see Table 2). Little (2013, pp. 109, 115) suggested that an RMSEA value of .08 indicates “mediocre” to “acceptable” fit and a TLI value of .89 indicates “mediocre” fit. As previously, models with five or six factors displayed factor loading patterns that were not interpretable. Models with fewer than four factors fit the data poorly. (All results are available from the corresponding author.) Table 3 displays the factor loadings for the four-factor model. The four factors in the preferred model are interpreted as (a) scoping the research topic and identifying needed information, (b) obtaining information via library Web pages and other sources, (c) evaluating information quality and producing research documents, and (d) crediting sources.

Table 2

Fit Statistics for Six Exploratory Factor Analysis Models

Model	χ^2	RMSEA	TLI	<i>df</i>
One-factor	2,009.049*	0.129	0.705	377
Two-factor	1,372.604*	0.106	0.800	349
Three-factor	1,059.569*	0.094	0.844	322
Four-factor	785.386*	0.080	0.887	296
Five-factor	620.669*	0.071	0.912	271
Six-factor	505.445*	0.064	0.929	247

Note. RMSEA = root mean square error of approximation; TLI= Tucker–Lewis Index.

* Chi-square model fit test statistic is statistically significant ($p < .001$).

Table 3

Factor Loadings for the Four-Factor EFA Model

Item	Factor 1	Factor 2	Factor 3	Factor 4
1. Determine the information I need	0.746*	0.094	0.040	0.031
2. Determine scope of research questions	0.900*	0.025	0.038	-0.163
3. Determine types of information sources	0.811*	0.020	-0.011	0.088
4. Conduct searches to frame topic	0.749*	-0.053	0.014	0.088
5. Use search box on library home page	-0.065	0.885*	-0.061	0.187
6. Access electronic sources through library	-0.052	0.921*	0.015	0.056
7. Access topic-relevant databases	0.089	0.744*	0.127	0.003
8. Determine where and how to find needed information	0.141	0.539*	0.281	-0.018
9. Search using keywords and Boolean logic	0.098	0.544*	0.287*	-0.157*
10. Use Internet search tools (Google Scholar, etc.)	0.121	0.083	0.355*	0.316 *
11. Recognize need to go beyond Internet resources	0.033	0.300*	0.367*	0.088
12. Evaluate currency of information	0.105	0.049	0.616*	0.125
13. Evaluate relevance of information	0.183	0.008	0.614*	0.095
14. Evaluate author's credentials	-0.132	0.131	0.860*	-0.083
15. Evaluate accuracy of information	-0.055	-0.011	0.916*	-0.071
16. Evaluate intended purpose of information	0.079	0.074	0.723*	-0.058
17. Assess quality of Internet information resources	-0.042	0.004	0.773*	0.097
18. Avoid fake information on the Web	-0.070	0.009	0.653*	0.214
19. Improve quality of my information-seeking process	0.140	0.023	0.665*	0.018
20. Synthesize information from a variety of sources	0.144	0.039	0.537*	0.195
21. Use bibliographic reference management tools	-0.021	0.039	0.604*	-0.021
22. Write documents	0.146	-0.013	0.515*	0.248
23. Create academic presentations	-0.001	-0.059	0.552*	0.296
24. Disseminate information using social media	0.102	-0.087	0.657*	0.083
25. Identify information as published, confidential, or proprietary	0.061	-0.014	0.589*	0.195
26. Cite all my sources	0.070	0.062	0.038	0.797*

Item	Factor 1	Factor 2	Factor 3	Factor 4
27. Use quotations properly	0.043	0.063	0.055	0.793*
28. Avoid plagiarism	−0.030	−0.159*	0.428*	0.533*
29. Create a bibliography	0.054	0.074	0.056	0.664*

Note. $N = 259$. Boldface indicates the factor that a certain item is judged to be primarily measuring.

*Factor loading is significantly different from zero ($p < .05$).

Factor Total Score Reliability

We also estimated the omega reliability coefficient for observed total scores on each factor in the best EFA model, the four-factor model. The total score reliability coefficient for Factor 1 (Scoping the research topic) was 0.902, for Factor 2 (Obtaining information) was 0.917, for Factor 3 (Evaluating information quality and producing research documents) was 0.960, and for Factor 4 (Crediting sources) was 0.891. Omega reliability coefficients ≥ 0.80 may be treated as “sufficiently high” for score use in most social science research (Raykov & Marcoulides, 2011, p. 169). DeVellis (2017) suggested that when score reliability values exceed 0.90, reducing the number of items could be considered.

Measurement Invariance Analysis

We also attempted to test measurement invariance of the information literacy self-efficacy scale between students who were taking an online course for the first time (67 students) and students who had previously taken one or more online courses (192 students). The sample size of the smaller group was marginal for factor analysis, so we conceptualized this analysis as highly exploratory. We tested “configural” invariance, first, constraining a multiple-group confirmatory factor analysis model so that both groups had the same four-factor structure but possibly different factor loadings. The chi-square test of model fit indicated that the configural invariance model did not fit the data— $\chi^2(742) = 1370.83, p < .0001$ —but fit indices suggested it was marginally adequate: TLI = 0.855; RMSEA = 0.081. We proceeded to test a “scalar” or “strong” invariance model with the same four-factor structure, constraining factor loadings, and item intercepts to be equal across groups. The fit of the scalar invariance model was no worse than that of the configural invariance model: $\chi^2(50) = 63.76, p = .091$. We interpret these results as suggesting that the information literacy self-efficacy items may be measuring the same underlying traits in the same metric, regardless of whether students have previously taken online courses or not, but we also recommend that this analysis should be repeated in a larger sample from the same population to provide stronger evidence.

Correlations Between Self-Reported Information Literacy Abilities and Information Literacy Knowledge

To analyze external evidence for the meaning of scores from our information literacy self-efficacy instrument, we computed correlations between factors in the self-efficacy and total scores on the eight-item information literacy knowledge test. Because lower scores on the information literacy self-efficacy scale indicated more consistent ability to carry out information literacy activities, we reverse-coded the information literacy knowledge test scores so that, similarly, lower values corresponded to higher information literacy knowledge. The correlation of the information literacy test scores with Factor 1 (Scoping the research topic) was 0.122; with Factor 2 (Obtaining information), 0.134; with Factor 3

(Evaluating information quality and producing research documents), 0.158; and with Factor 4 (Crediting sources), 0.185. Although they are in the expected positive direction, these correlations are fairly small, indicating perhaps modest evidence of students' ability to accurately self-report on their own information literacy. We expect the magnitude of the correlation may have been attenuated to some extent due to limited reliability of the information literacy knowledge test scores, which consisted of only eight binary-scored items.

Future Research

Although the instrument demonstrated promising preliminary results, additional confirmatory validation is advised to further establish its reliability and verify the underlying dimensional structure. It is also recommended that future studies employ confirmatory factor analysis to assess the predictive validity of the OILS instrument, as the current study primarily used EFA and reliability analysis to examine the instrument's psychometric properties.

In addition, future research should examine how emerging generative AI tools intersect with students' information literacy self-efficacy. While the OILS instrument captures foundational competencies such as evaluating information quality, ethical use, and attribution, AI-mediated information practices introduce new dimensions of judgment, transparency, and responsibility that warrant explicit investigation. Subsequent studies could extend the instrument by incorporating items that assess students' confidence in critically engaging with AI-generated content, disclosing AI use, and integrating AI tools responsibly into academic work.

Conclusions

The Online Information Literacy Self-Efficacy (OILS) instrument, as detailed in Appendix C, presents a robust, nuanced measure of students' perceived abilities in information literacy within online learning environments. Through empirical testing using exploratory factor analysis and reliability analysis, the instrument's four-factor structure—scoping research topics, obtaining information, evaluating information quality and producing research documents, and crediting sources—was confirmed, with strong reliability coefficients. These findings highlight the critical role of self-efficacy in online information literacy and provide educators and librarians with a practical tool to assess and support student readiness in navigating complex information landscapes. Future research is encouraged to examine other types of validity evidence, such as confirmatory factor analyses and further predictive score correlations, ensuring the instrument's applicability across diverse student populations and learning contexts.

References

- Adigue, A. P., Tumacder, J. R. O., & Urbano, J. M. (2023). Self-efficacy and learning strategies in the context of online learning. In *Proceedings of the 3rd International Conference on Education and Technology (ICETECH 2022)* (pp. 547–564). Atlantis Press. https://doi.org/10.2991/978-2-38476-056-5_55
- Association of American Colleges and Universities. (2009). *Information literacy VALUE rubric*. <https://www.aacu.org/value/rubrics/information-literacy>
- Association of College and Research Libraries. (2015). *Framework for information literacy for higher education*. American Library Association. <https://www.ala.org/acrl/standards/ilframework>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Bradley, C. (2013). Information literacy in the programmatic university accreditation standards of select professions in Canada, the United States, the United Kingdom, and Australia. *Journal of Information Literacy*, 7(1), 44–68. <https://doi.org/10.11645/7.1.1785>
- DeVellis, R. F. (2017). *Scale development: Theory and applications*. SAGE Publications.
- Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernandez-Leo, D., Järvelä, S., Mavrikis, M., & Rienties, B. (2024). The promise and challenges of generative AI in education. *Behaviour & Information Technology*, 44(11), 2518–2544. <https://doi.org/10.1080/0144929X.2024.2394886>
- Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review*, 17, 63–84. <https://doi.org/10.1016/j.edurev.2015.11.002>
- Hovious, A. (2024). Information creation as an AI prompt: Implications for the ACRL framework. *Kansas Library Association College and University Libraries Section Proceedings*, 14(1), Article 5. <https://doi.org/10.4148/2160-942X.1094>
- Johnston, N. (2010). Is an online learning module an effective way to develop information literacy skills? *Australian Academic & Research Libraries*, 41(3), 207–218. <https://doi.org/10.1080/00048623.2010.10721464>
- Kassorla, M., Georgieva, M., & Papini, A. (2024). *AI literacy in teaching and learning: A durable framework for higher education*. EDUCAUSE. <https://www.educause.edu/content/2024/ai-literacy-in-teaching-and-learning/executive-summary>
- Little, T. D. (2013). *Longitudinal structural equation modeling*. Guilford Press.

- McClellan, S. (2016). Teaching critical thinking skills through commonly used resources in course-embedded online modules. *College & Undergraduate Libraries*, 23(3), 313–327. <https://doi.org/10.1080/10691316.2014.987416>
- McDonald, R. P. (1970). The theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology*, 23(1), 1–21. <https://psycnet.apa.org/doi/10.1111/j.2044-8317.1970.tb00432.x>
- McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & Research in Social Education*, 46(2), 165–193. <https://doi.org/10.1080/00933104.2017.1416320>
- Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. Routledge. <https://doi.org/10.4324/9780203841624>
- UNESCO. (2023, September 7). *Guidance for generative AI in education and research*. <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>
- Weaver, K. D. (2024). The artificial intelligence disclosure (AID) framework: An introduction. *College & Research Libraries News*, 85(10), 407–411. <https://doi.org/10.5860/crln.85.10.407>

Appendix A

Initial Online Information Literacy Self-Efficacy (OILS) Instrument

I feel confident and competent to ...	Almost always true	Often true	Sometimes true	Usually not true	Almost never true

Determine the extent of information needed—4 items

Determine the information I need

If needed, determine the scope of the research questions

Determine types of information sources

Conduct needed preliminary information searches to help frame my research topic

Access the needed information (through academic libraries vs. on the Web)—8 items

Use the search box on an academic library's home page

Access any electronic information such as specific journal articles through academic libraries

Access databases appropriate for my topic

Decide/determine where and how to find the types of information I need

Develop effective search strategies by using keywords and Boolean logic

Seek information on the Web

Use Internet search tools (Google Scholar, etc.)

Recognize the need to go beyond Internet resources

Evaluate information critically—7 items

Evaluate currency—the timeliness of the information

Evaluate relevancy—the importance of the information for my needs

Evaluate author's credentials (author's name, contact information, and expertise identifiable)

Evaluate accuracy—the reliability,

I feel confident and competent to ...	Almost always true	Often true	Sometimes true	Usually not true	Almost never true
truthfulness, and correctness of the content					
Evaluate purpose—the reason the information exists and for whom the information is intended					
Assess the quality of Internet information resources					
Avoid fake information on the Web					
Effectiveness to accomplish a research assignment in an online course—6 items					
Criticize the quality of my information-seeking process					
Summarize and synthesize information from a variety of types of information sources					
Use bibliographic references management tools (Endnote, Zotero, etc.)					
Write documents (discussions, reports, reflections, research papers, etc.)					
Create academic presentations (PowerPoint, digital storytelling, etc.)					
Disseminate information on the Internet (social media)					
Crediting sources—5 items					
Identify published, confidential, or proprietary information					
Cite all my sources in my research assignment					
Use quotations properly in my research assignment					
Avoid plagiarism					
Create a bibliography					

Appendix B

Eight Multiple-Choice Information Literacy Knowledge Items

Correct answers are marked in *italics*.

Q1. When performing research, one important part of forming a search strategy is to:

- A. Locate books using the library's online catalog
- B. Search a computer database for articles
- C. Analyze your topic to identify alternative keywords/concepts*
- D. Check the Internet for background information on your topic

Q2. The best sites always come up first in a search result. (Yes; No)

Q3. Read the following statement and answer the question below:

Statement: "Describe the effects of automobile emissions on air quality."

Which of the following keyword examples may yield the best results for the statement above in an electronic database search? (The * indicates multiple suffix possibilities.)

- A. automobile* and effects
- B. automobile* and emission* and air quality*
- C. effects and emission* and quality
- D. emission* and air quality

Q4. To refine your Web search results, you should:

- A. Use the advanced search feature*
- B. Use only one search engine
- C. Limit your results to the first 10 pages
- D. Use the built-in Help*

Q5. Which of the following is a characteristic of a scholarly journal?

- A. The presence of glossy photos and advertisements
- B. Articles are written by people from a variety of fields
- C. Footnotes and bibliographies are regularly used to cite sources*
- D. Topics covered are not confined to one major field of study

Q6. When evaluating resources, which one is the least important to consider?

- A. Document format*
- B. Publishing dates
- C. Publisher
- D. Author

Q7. Read the following statement and answer the question below:

Statement: "Describe the effects of automobile emissions on air quality."

Which source would most likely provide you with objective information for the main concepts in the

statement?

- A. A personal interview with an influential lobbyist
- B. A website that advocates clean air
- C. The latest annual report from a major automobile manufacturer
- D. *A study featured in a peer-reviewed periodical*

Q8. When evaluating resources, which of the following statements are not correct?

- A. Check the author's authority to write on the topic
- B. *Use only the most current resources*
- C. Read the mission statement of a Web page
- D. Find out if others have cited the resource
- E. Use resources that include different viewpoints

Appendix C

Online Information Literacy Self-Efficacy (OILS) Instrument

I feel confident and competent to ...	Almost always true	Often true	Sometimes true	Usually not true	Almost never true
Scoping the research topic and identifying needed information—4 items					
Determine the information I need					
If needed, determine the scope of the research questions					
Determine types of information sources					
Conduct needed preliminary information searches to help frame my research topic					
Obtaining the needed information (through academic libraries vs. on the Web)—5 items					
Use the search box on an academic library's home page					
Access any electronic information through academic libraries such as specific journal articles					
Access databases appropriate for my topic					
Decide/determine where and how to find the types of information I need					
Develop effective search strategies by using keywords and Boolean logic					
Evaluating information critically and creating research documents—16 items					
Use Internet search tools (Google Scholar, etc.)					
Recognize need to go beyond Internet resources					
Evaluate currency—the timeliness of the information					
Evaluate relevancy—the importance of the information for your needs					
Evaluate author's credentials (author's name, contact information, and expertise identifiable)					
Evaluate accuracy—the reliability, truthfulness, and correctness of the content					

I feel confident and competent to ...	Almost always true	Often true	Sometimes true	Usually not true	Almost never true
Evaluate purpose—the reason the information exists and for whom the information is intended					
Assess the quality of Internet information resources					
Avoid fake information on the Web					
Criticize the quality of my information- seeking process					
Summarize and synthesize information from a variety of types of information sources					
Use bibliographic references management tools (Endnote, Zotero, etc.)					
Write documents (discussions, reports, reflections, research papers, etc.)					
Create academic presentations (PowerPoint, digital storytelling, etc.)					
Disseminate information on the Internet (social media)					
Identify published, confidential, or proprietary information					
Crediting sources—4 items					
Cite all my sources in my research assignment					
Use quotations properly in my research assignment					
Avoid plagiarism					
Create a bibliography					



August – 2026

Predicting Student Outcomes in Open High Schools Using Educational Data Mining

Ahmet Polat^{1*} and Mehmet Barış Horzum²

¹Ministry of National Education, Sivas, Türkiye; ²Department of Computer Education and Instructional Technology, Faculty of Education, Sakarya University, Adapazarı, Türkiye; *Corresponding Author

Abstract

Open high schools fulfill a critical function by offering flexible educational pathways for students encountering diverse socioeconomic and personal challenges. Nevertheless, escalating dropout rates represent a significant concern, necessitating robust predictive models to facilitate early interventions. This study employed educational data mining (EDM) techniques to analyze the academic trajectories of 484,158 students enrolled in Turkish open high schools. We evaluated multiple classification algorithms—J48, decision tree, k-nearest neighbors (kNN), naïve Bayes, and random forest—across four distinct data preprocessing scenarios to predict student status (dropout/delayed graduation/graduation). The J48 algorithm demonstrated superior performance, achieving an accuracy of 80.47% and a kappa statistic of 0.61. Key findings reveal that academic and administrative features, notably total credit accumulation and initial enrollment type, are more important predictors than demographic variables. This research provides empirically grounded insights for the early identification of at-risk students. It offers data-driven recommendations to enhance student retention policies within open high school systems, contributing a large-scale, multi-class prediction analysis within a unique, under-researched national distance education context.

Keywords: educational data mining, dropout prediction, open high school, machine learning, distance education, student retention, early warning system

Introduction

Education is fundamental to human development, influencing individual opportunities and societal progress. The United Nations' (UN) 2030 Sustainable Development Goals reflect this priority, targeting universal primary and secondary access, vocational skills development, and equitable tertiary education (United Nations Department of Economic and Social Affairs [UNDESA], 2015), while the UN Human Development Index measures educational progress through average and expected years of schooling (United Nations Development Program [UNDP], 2021). As a UN member, Türkiye pursues these educational aims, yet faces persistent challenges in achieving these targets.

Equitable access to quality education remains a significant challenge in Türkiye (Ministry of National Education [MoNE], 2015). Although overall schooling rates vary across educational levels—68% in preschool, 91% in primary, 84% in secondary, and 44% in higher education (MoNE, 2019)—secondary education, in particular, exhibits notable regional disparities. For instance, the secondary school enrollment rate is 91% in western Anatolia, compared to 72% for males and 68% for females in southeastern Anatolia (MoNE, 2019). Furthermore, among adults aged 25–34, nearly half have not completed secondary education, with a significant gender gap in average years of schooling: 7.3 years for women versus 9 years for men (Organisation for Economic Co-operation and Development [OECD], 2018; UNDP, 2021).

Addressing educational challenges, Türkiye extended compulsory education to 8 years (Türkiye Cumhuriyeti Resmî Gazete [TC Resmî Gazete], 1997), then 12 years (TC Resmî Gazete, 2012). National development plans promoted open and distance education for equality and lifelong learning (TC Resmî Gazete, 2006, 2013). Subsequently, open high school (OHS) enrollment grew from 14% (940,268) of all secondary students in 2012 to 26% (1,554,938) in 2017 (MoNE, 2012, 2017). However, dropout rates increased dramatically; by 2017, OHS dropouts ($N = 2,143,674$) reached 137% of active enrollment (Polat, 2021), a persistence problem recognized in MoNE strategic planning (2015).

Secondary education dropout impacts national development through diminished human capital (Rosenzweig, 2010; Rumberger, 2020), reduced productivity (Aydın et al., 2012; Çalışkan et al., 2013), and increased social costs in crime, health, and life expectancy (Ereş, 2005; Kızmaz, 2004; Luy et al., 2019). Young adults without high school diplomas face higher unemployment rates (OECD, 2021). Parental education also significantly influences academic outcomes across generations (MoNE, 2022). Student dropout within Turkish OHSs represents an escalating problem affecting a student demographic facing socioeconomic and personal barriers to traditional schooling and reducing OHS dropout rates is thus crucial for both individual success and the country's socioeconomic advancement.

Open High Schools in Türkiye

Türkiye's first OHS (in Turkish, açık öğretim lisesi or AÖL) was established in 1992. Subsequently, the vocational open high school (in Turkish, mesleki açık öğretim lisesi or MAÖL) was established in 2006, followed by the open education imam hatip high school (açık öğretim imam hatip lisesi or AÖİHL) in 2016. While these institutions were initially part of AÖL, they later became separate entities. As of the 2020/2021 academic year, the total number of actively enrolled students across these OHSs reached 1,452,331 (MoNE, 2021).

Despite the size of the student population, research on the OHS dropout problem remains limited. Early studies focused on operational aspects, learning environments, materials, and support services

(Demiray & Sağlık, 2003). Later research examined reasons for transferring to open schools (Adıgüzel, 2016; Çuhadar Öncü, 2017; Şahin, 2017), student-perceived obstacles (Çiçek, 2005; Şahin, 2017; Sipahi, 2019; Soylu, 2014), academic achievement (Özkahveci, 2001), service effectiveness (Bedel, 2006; Sarıhan, 2010; Şentürk, 2009; Yavuz, 2014), student perspectives (Demirtaş et al., 2017; Dere, 2002), and individual differences (Randler et al., 2014).

However, much of this existing research is descriptive and localized, relies on small sample sizes (often focusing on MAÖL students due to face-to-face components), and lacks representativeness concerning the vast and diverse population enrolled across Türkiye. Consequently, there is limited understanding of the factors predicting persistence or dropout on a national scale, hindering the development of evidence-based interventions. While engaging large, representative samples present logistical challenges (labor, time, cost), the increasing availability of rich administrative data within student information systems offers a valuable opportunity. Educational data mining (EDM) techniques can leverage this data to uncover complex patterns and predictors of student outcomes that might otherwise remain obscured.

Educational Data Mining

Information technology advancements enable educational institutions to collect comprehensive student data through sophisticated management systems, creating valuable analytical opportunities. Data mining, extracting implicit and potentially useful information from large datasets (Han et al., 2011), provides tools for addressing complex educational challenges. EDM applies these methods to large-scale educational data, illuminating student behaviors and learning environments (International Educational Data Mining Society, n.d.). EDM applications primarily focus on predicting academic performance, identifying at-risk students, and understanding dropout factors (Romero & Ventura, 2010).

EDM holds promise for anticipating educational needs, personalizing learning experiences, and ultimately improving student outcomes. Prediction studies, a core component of EDM, assist educators in identifying critical success factors and their interrelationships, thereby highlighting areas for targeted intervention and systemic improvement. Within the Turkish context, the concurrent rise in demand for OHSs and the alarming increase in dropout rates create a compelling case for applying EDM. The existing research gap concerning dropout predictors in these specific institutions necessitates the use of data mining methods to uncover potentially complex patterns associated with success, delayed graduation, and dropout, with the ultimate goal of enhancing completion rates.

Identifying students at risk of dropout is a crucial prerequisite for implementing timely and effective interventions (Dekker et al., 2009; Heppen & Therriault, 2008; Márquez-Vera et al., 2016; Tanner, 2003). Dropout is a multifaceted phenomenon influenced by individual, academic, family, school, and community factors (Rumberger & Lim, 2008). However, collecting comprehensive data, particularly psychosocial or detailed background information, via traditional methods such as surveys is often impractical in large-scale distance education settings like Turkish OHSs. Therefore, leveraging readily available administrative data from student information systems presents a pragmatic and efficient approach to investigating dropout predictors, conserving resources while enabling analysis at scale.

A review of the EDM literature focused on predicting student performance and dropout reveals a predominant focus on higher education settings (Berens et al., 2019; Delen, 2010; Djulovic & Li, 2013; Kovacic, 2010; Lassibille & Navarro-Gomez, 2008). Studies examining dropout at the high school level

are comparatively scarce (Lee & Chung, 2019; Márquez-Vera et al., 2013; Şara et al., 2015). Existing research spans diverse national contexts, educational systems (face-to-face, distance, e-learning), and data sources. While common predictors such as academic achievement (GPA, prior grades, exam scores), personal characteristics (gender, age, disability), and sometimes psychometric traits (motivation) emerge, the heterogeneity in study contexts, methodologies, sample sizes, and data types limits the generalizability of findings. This variability underscores the necessity for context-specific research.

This study is conceptually grounded in established models of student departure, primarily drawing upon Tinto's (1975) model of student integration, and Bean and Metzner's (1985) model tailored for non-traditional students. Tinto's framework emphasizes the critical roles of academic and social integration in fostering persistence. Bean and Metzner's model, particularly relevant to the OHS population, which often includes non-traditional learners, additionally highlights the influence of external environmental factors (e.g., employment, family responsibilities) alongside academic variables and background characteristics. This theoretical lens guides our variable selection, positing that a combination of factors—including academic progress (e.g., credits earned), background characteristics (e.g., enrollment type, age), and potentially external factors reflected in administrative data (e.g., employment status)—interact to shape student persistence within the unique context of OHSs.

Given research gaps—the lack of national-scale studies in the Turkish OHS context, the limited focus on the dropout problem in open secondary education, and the need for nuanced outcome prediction—an EDM study investigating dropout in this specific setting is warranted. Such research is essential not only for addressing existing knowledge gaps but also for guiding future research directions and informing the development of targeted, evidence-based policies and interventions aimed at improving student retention. This study directly addressed these gaps by applying predictive modeling to a comprehensive, nationwide dataset of Turkish OHS students, employing a multi-class outcome variable for a more nuanced and realistic representation of learner trajectories, and identifying the most important predictors specific to this unique distance secondary education environment.

Research Questions

This study aimed to predict the status (graduation, active/delayed graduation, or dropout) of students enrolled in Turkish OHSs at the end of their standard 4-year education period (defined as the official program duration starting from the student's initial enrollment date), using EDM methods applied to data available within the institutional student information system. In pursuit of this aim, the study posed the following research questions:

RQ1: What is the classification performance of various machine learning models developed to predict the graduation, dropout, and active/delayed status of OHS students?

RQ2: Which student features available in the administrative dataset are the most important predictors of these outcomes?

Methodology

Employing EDM techniques to analyze educational datasets, this study followed the Cross Industry Standard Process for Data Mining (CRISP-DM) model (Chapman et al., 2000), a structured framework

guiding the analysis from business understanding through deployment (Figure 1). This systematic methodology facilitated a rigorous approach to the data analysis process.

Data Analysis

Following the CRISP-DM framework, all data processing and modeling were carried out using RapidMiner Studio VM software (Version 9.0), a comprehensive data science platform facilitating visual workflow design and automation across the data science lifecycle (Altair Engineering, 2026).

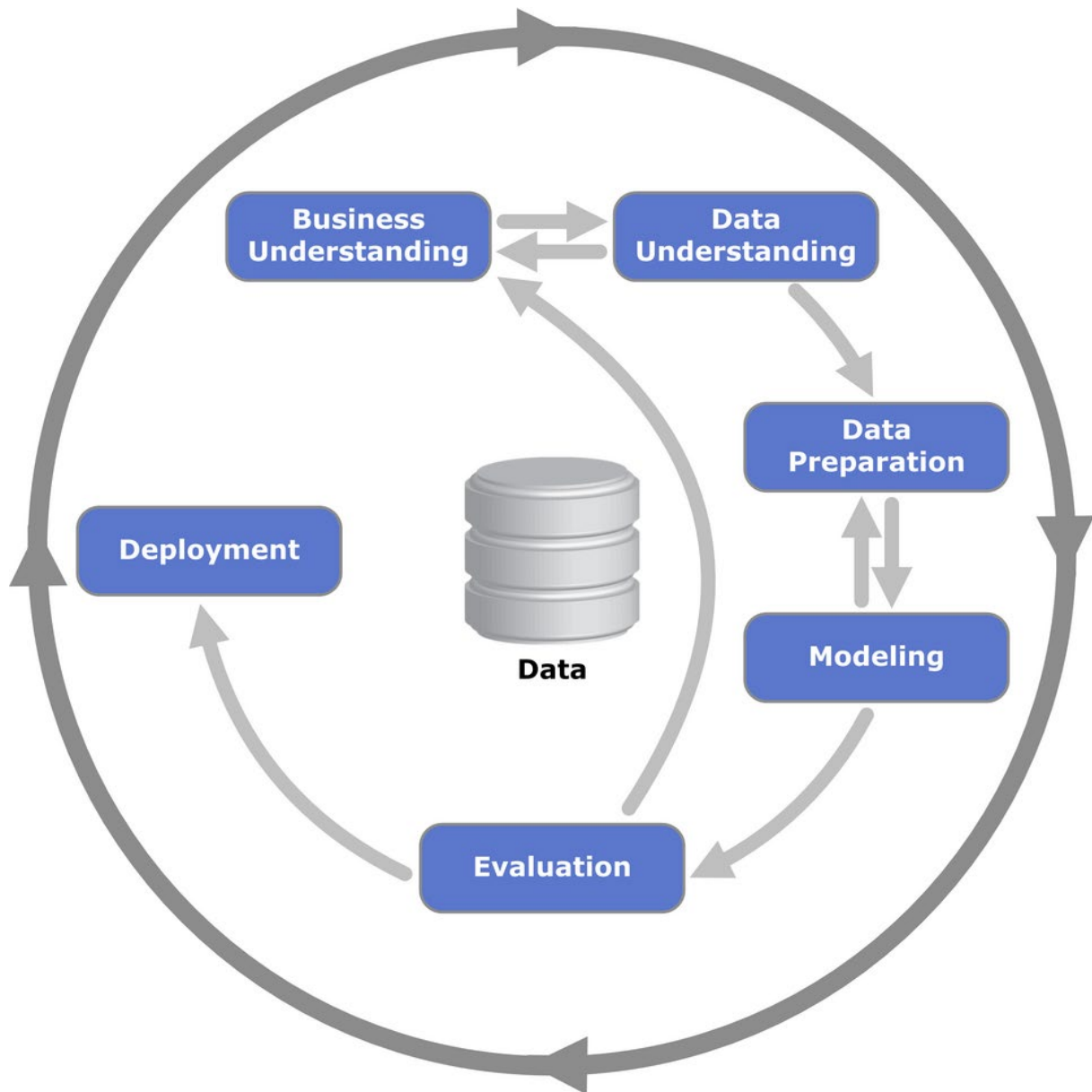
Business Understanding

The core objective was predicting student status after the standard 4-year timeframe, with three target categories:

- graduated: students who completed all requirements and received their diploma within or by the end of the 4-year timeframe.
- active/delayed graduation: students who remained enrolled beyond the standard 4 years, indicating delayed progress towards graduation.
- dropout: students who discontinued their studies and were no longer enrolled, without having graduated.

Figure 1

CRISP-DM Model



Note. From *CRISP-DM Process Diagram*, by K. Jensen, 2012, Wikimedia Commons (<https://commons.wikimedia.org/w/index.php?curid=24930610>). CC BY-SA 3.0.

Data Understanding

The dataset was obtained with official permission from the MoNE Information Processing Department (Official Letter No: 65968543-605.01-E.86777), adhering to MoNE regulations. It comprised anonymized records for students who first enrolled in one of the three types of OHSs (AÖL, MAÖL, or AÖİHL) during the 2013 academic year. This year was chosen as it marked the implementation of 12-year compulsory education and saw a significant increase in OHS enrollment. The final analysis dataset, after handling exclusions for specific non-representative cases (e.g., deceased, returned to formal education), consisted of 484,158 student records.

The raw data contained various administrative, demographic, and academic features. Exploratory analysis using RapidMiner identified the potentially relevant attributes listed in Table 1.

Table 1

Features Initially Available in the Dataset

Category	Feature	Description
Administrative	School type	Type of OHS (AÖL/MAÖL/AÖİHL)
	Enrollment type	Educational status at enrollment (e.g., primary/secondary graduate, transfer from traditional HS)
	Field/department	Registered field or department (if applicable)
Demographic	Date of birth	Student's date of birth
	Gender	Student's gender
	Address city	City of residence
	Address district	District of residence
	Employment status	Employment status (employed/unemployed, type of employment)
	Disability status	Presence and type of disability
	Special status	Special circumstances (e.g., detained, child of a martyr/veteran, under court order)
	Military service	Military service status (relevant primarily for males)
Academic	Status	Target variable: Student's status (graduated, active, dropout)
	Previous system	The class transition system of the school student transferred from
	Previous school	Type of school transferred from (if applicable)
	Total credit count	Total course credits earned by the student

Note. OHS = open high school; AÖL = açık öğretim lisesi [open high school]; MAÖL = mesleki açık öğretim lisesi [vocational open high school]; AÖİHL = açık öğretim imam hatip lisesi [imam hatip open high school].

The majority of students (80%) were enrolled in AÖL, followed by MAÖL (13%) and AÖİHL (7%). Gender distribution varied by school type: we found approximately 60% were male in the AÖL, 62% were male in the MAÖL, whereas 63% were female in the AÖİHL.

Data Preparation

Several steps were taken to prepare the data for modeling:

- An age feature at the time of enrollment (calculated from date of birth) was added, aligning with theoretical propositions regarding the impact of age on student persistence.
- A region feature was derived using the European Union's nomenclature of territorial units for statistics (NUTS) level 1 classification (Eurostat, 2026) based on the address city to capture regional socioeconomic variations that might influence educational outcomes.
- The status feature was designated as the target variable. The distribution across the three classes was: graduated 11.1%, active/delayed graduation 26.2%, and dropout 62.7%.

The dataset exhibited significant class imbalance. We deliberately avoided resampling techniques to preserve the authentic data distribution reflecting the real-world phenomenon. Our strategy focused on using algorithms relatively robust to imbalance and employing evaluation metrics (precision, recall, F-measure, kappa) less sensitive to imbalance than simple accuracy (He & Garcia, 2009).

To systematically evaluate data representation impact on model performance, we implemented four progressive preprocessing stages: stage 1: minimal transformation; stage 2: removal of high-missingness features; stage 3: simplification of nominal features; and stage 4: discretization of continuous variables.

Modeling

Five classification algorithms were selected for this study: decision tree, J48, k-nearest neighbors (kNN), naïve Bayes, and random forest. These algorithms were chosen because they represent different methodological approaches to classification and have demonstrated effectiveness in EDM research (Kılınç, 2015; Shahiri & Husain, 2015; Umar, 2016; Yukselturk et al., 2014; Yurdakul, 2015). Testing multiple algorithms allowed us to identify which approach performed best for predicting student outcomes in the Turkish OHS context.

The selected algorithms operate using the following distinct methodologies:

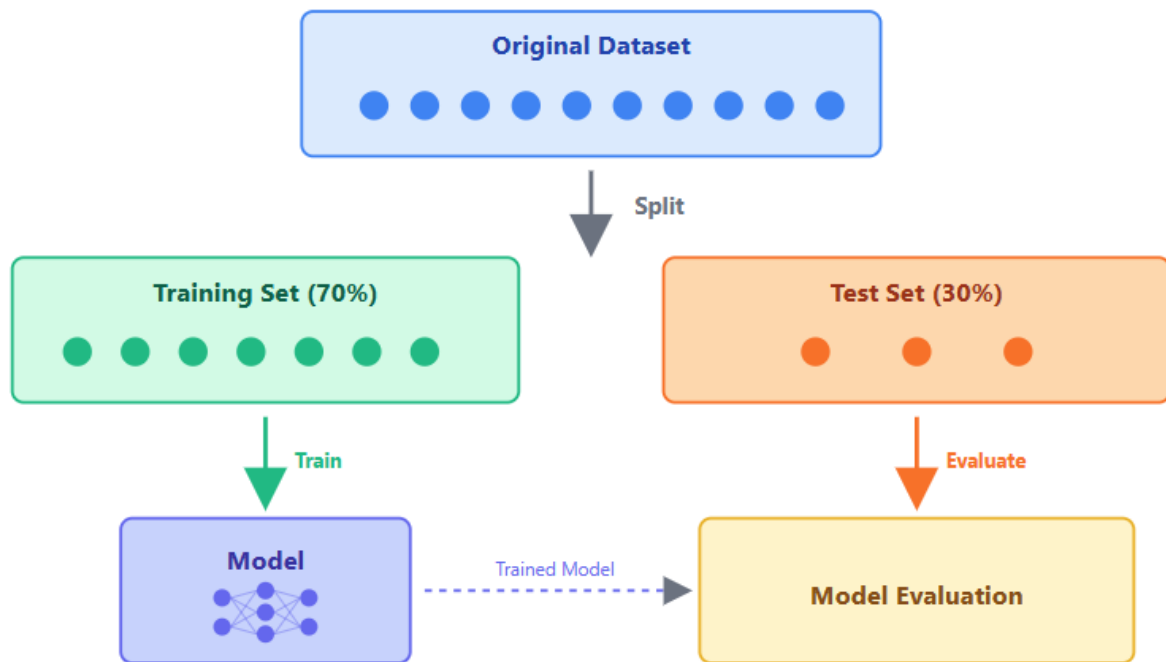
- decision tree: constructs a flowchart-like model where branches represent decision nodes derived from specific student characteristics, such as credit accumulation.
- J48: an advanced decision tree that improves prediction by automatically pruning (trimming) unnecessary branches to prevent overfitting, when a model learns training data too specifically and performs poorly on new data. J48 works by examining student features, selecting the feature that best separates students into outcome groups, creating branches, and pruning branches that don't improve accuracy.
- random forest: generates an ensemble of decision trees and aggregates their predictions through a voting process to increase accuracy.
- k-nearest neighbors (kNN): uses instance-based learning, classifying students by identifying the prevailing outcome among similar data points.

- naïve Bayes: applies probabilistic principles to predict outcomes based on attribute likelihoods and the assumption of feature independence.

For model training and testing, the holdout method (70% training, 30% testing), depicted in Figure 2, was employed with stratified sampling, which builds random subsets to ensure the class distribution is the same as in the whole dataset (Altair Engineering, 2026).

Figure 2

Holdout Method Used for Model Training and Testing



All algorithms were run using their default parameter settings (without fine-tuning) within RapidMiner Studio (Altair Engineering, 2026). This approach was chosen to establish a baseline comparison of algorithm performance under standard conditions across different preprocessing stages (Romero & Ventura, 2020), rather than optimizing each algorithm individually, which was considered beyond the scope of this initial investigation but represents an avenue for future refinement (Witten et al., 2016).

Evaluation

Model performance was assessed using standard metrics derived from the confusion matrix on the test set (Hossin & Sulaiman, 2015): correct classification rate (CCR/accuracy), precision, recall (sensitivity), F-measure (F1-score), and kappa statistic (Cohen, 1960; Landis & Koch, 1977). Model performance metrics are explained in Figure 3, which also shows the confusion matrix from which we derived these metrics. (Further detail about this matrix is presented in Table 2).

Figure 3

Model Performance Metrics

Accuracy (CCR)

Percentage of all correct predictions

Example: 80 correct out of 100 students = 80% accuracy

Precision

When model predicts an outcome, how often is it correct?

Example: Predicted "Graduated" 20 times, 18 actually graduated = 90% precision

Recall (Sensitivity)

Of all students with an outcome, how many did model identify?

Example: 25 actually graduated, model found 20 of them = 80% recall

F-Measure (F1-Score)

Balance between Precision and Recall (harmonic mean)

Single metric combining both: 85% F-measure = good balance

Kappa Statistic

Performance compared to random guessing

Interpretation Scale:

≤ 0.00: No better than chance

0.01-0.20: Slight agreement

0.21-0.40: Fair agreement

0.41-0.60: Moderate agreement

0.61-0.80: Substantial agreement

0.81-1.00: Almost perfect

Confusion Matrix Example:

		Predicted			
		Dropout	Graduated	Active	
Actual	Dropout	82,182	50	18,101	Correct predictions Incorrect predictions
	Graduated	634	15,459	2,112	
	Active	6,651	306	17,123	

Using multiple metrics provides a comprehensive view of model performance. A model might achieve high CCR but low recall for a particular class, meaning it is correct most of the time but misses many cases of that outcome. In the Turkish OHS context, the ideal model would be one that not only achieves high overall CCR but also reliably identifies at-risk students (high recall for the dropout class) without excessive false alarms (high precision). By evaluating models across multiple metrics, we were able to identify which algorithm provided the best balance of performance characteristics for the specific goal of predicting student outcomes.

Deployment

This phase involved reporting the findings, discussing their implications, and formulating recommendations, as presented below.

Findings

The classification analyses were performed iteratively across the four preprocessing stages:

- stage 1: initial modeling using most available features with minimal transformation (baseline).
- stage 2: removal of features with high percentages of missing values or limited applicability (e.g., the features previous school, previous system, and military service status) to create a more complete and potentially robust dataset.
- stage 3: transformation and simplification of nominal features. For example, employment status and disability status were binarized (yes/no). Categories within enrollment type were consolidated. The region feature replaced detailed location data. This stage aimed to reduce complexity and potentially improve model generalizability.
- stage 4: discretization of continuous variables (age, total credit count) using different binning strategies (frequency-based, range-based) and varying numbers of bins (2, 4, 8). This explored whether converting continuous features into categorical ones improved performance for certain algorithms.

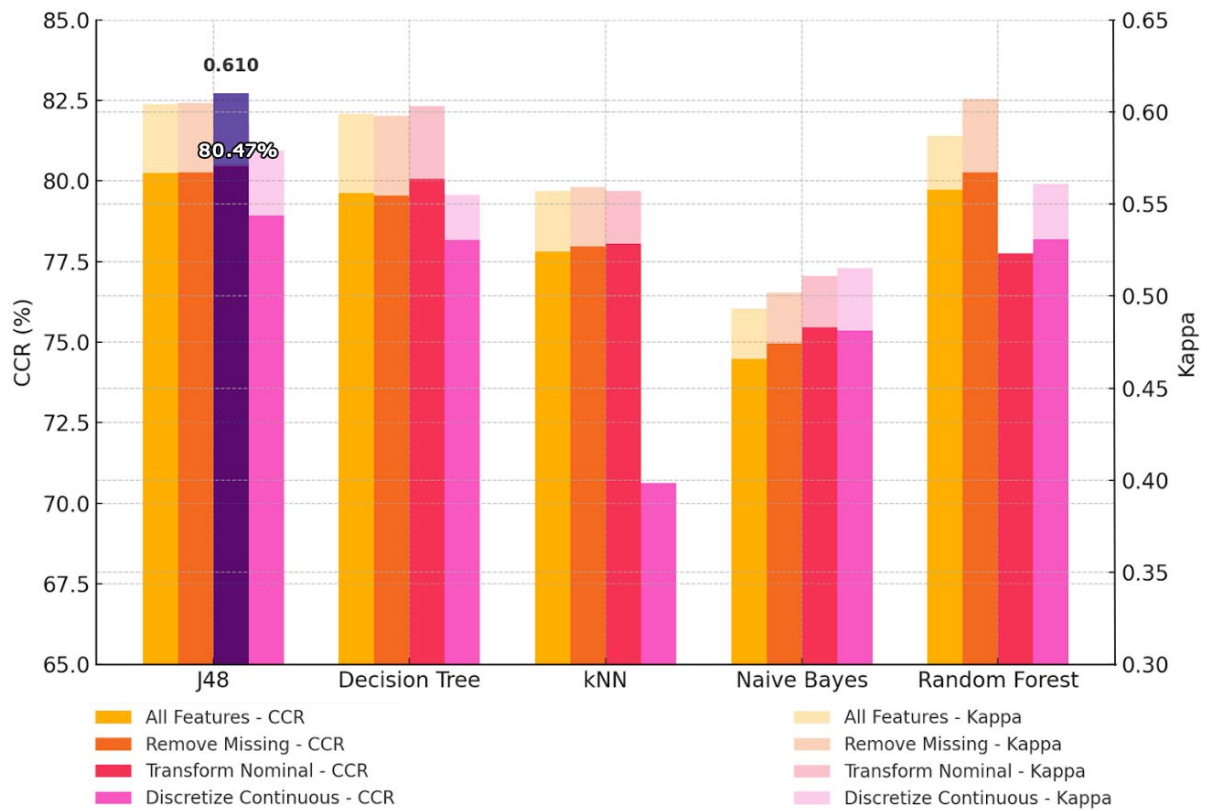
The performance of the five selected algorithms was evaluated at each stage using the test dataset.

Model Performance Comparison

The CCR and kappa metrics for the models developed using the aforementioned datasets are presented in Figure 4.

Figure 4

Comparison of Algorithm Performance Across Preprocessing Stages



Note. The purple bar highlights the most accurate algorithm, J48, and its performance metrics. CCR = correct classification rate.

Analysis across preprocessing stages showed accuracy (CCR) ranging from 73.36% to 80.47%. The highest performance was consistently achieved after the third preprocessing stage (simplifying nominal features). Within this stage, the J48 algorithm yielded the best results, achieving a CCR of 80.47% and a Kappa statistic of 0.61. This kappa value indicates substantial agreement (Landis & Koch, 1977), suggesting a robust model performing significantly better than chance.

Detailed Performance of the Best Model (J48—Stage 3)

The confusion matrix for the J48 model, the best performing on the test set (shown earlier in Figure 3), is presented in Table 2.

Table 2

Confusion Matrix for the J48 Model (N Instances in Test Set)

	Predicted			Total actual	Recall, %
	Dropout	Graduated	Active		
Actual					
Dropout	82.182	50	18.101	100.333	91.5
Graduated	634	15.459	2.112	18.205	97.7
Active	6.651	306	17.123	24.080	45.8
Total Predicted	89.467	15.815	37.336	142.618	
Precision, %	90.1	88.2	63.5		

The model demonstrates excellent performance for dropout and graduated classes (recall > 91.5%). Performance is notably lower for the active/delayed graduation class, particularly recall (45.8%). This difficulty in identifying students persisting beyond four years warrants further discussion.

The overall performance metrics for the best J48 model are shown in Table 3.

Table 3

Overall Performance Metrics for the Best J48 Model

Algorithm	Weighted avg recall, %	Weighted avg precision, %	Weighted avg F- measure, %	CCR, %	Kappa
J48	78.48	79.31	78.89	80.47	0.61

Note. CCR = correct classification rate.

These balanced overall metrics indicate a robust model that performed consistently across different facets of classification, minimizing bias toward predicting one class at the expense of others (Ben-David, 2008). This level of performance aligns well with benchmarks often considered effective for practical application in educational settings (Baker & Inventado, 2014; Romero & Ventura, 2010).

Feature Importance

Feature importance analysis for the best J48 model (Stage 3 data) revealed the relative importance of predictors (Figure 5).

Figure 5

Feature Importance Weights for the Best J48 Model

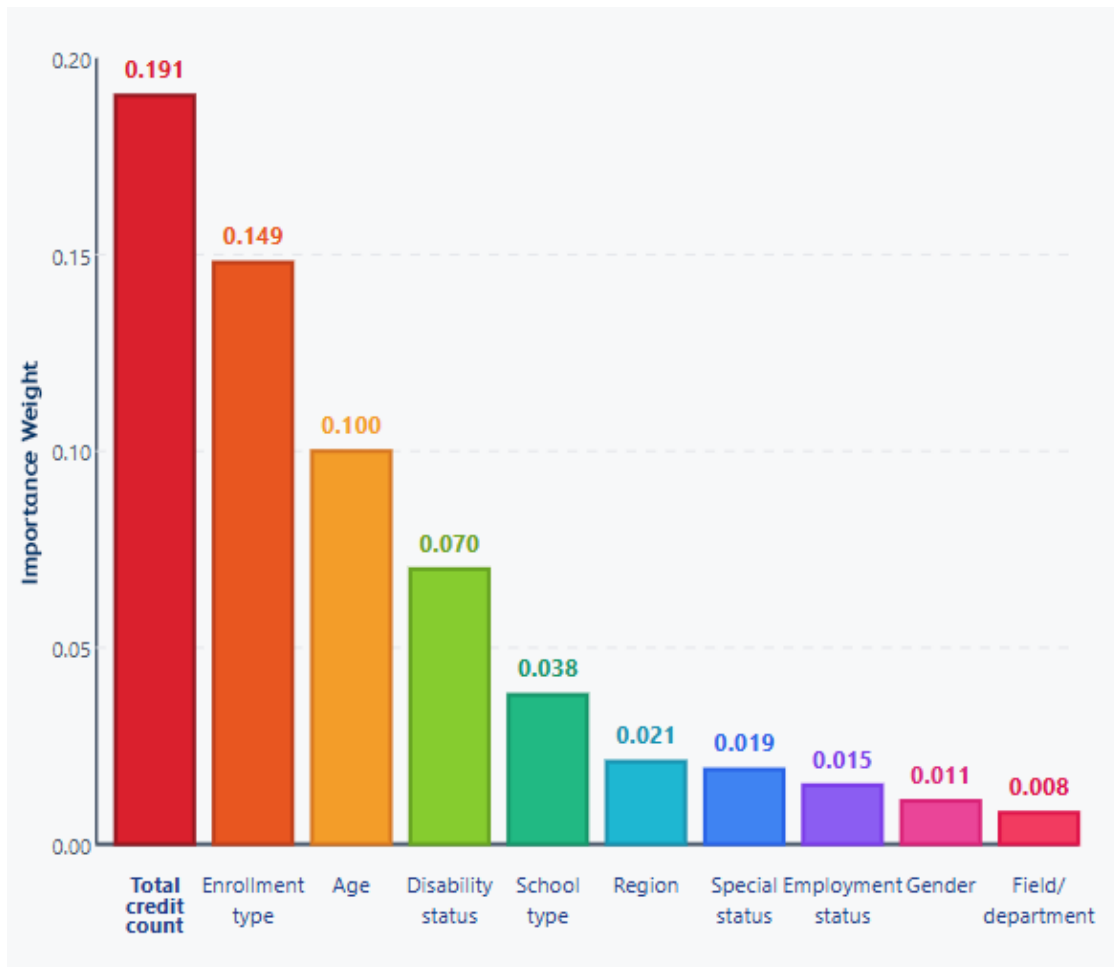


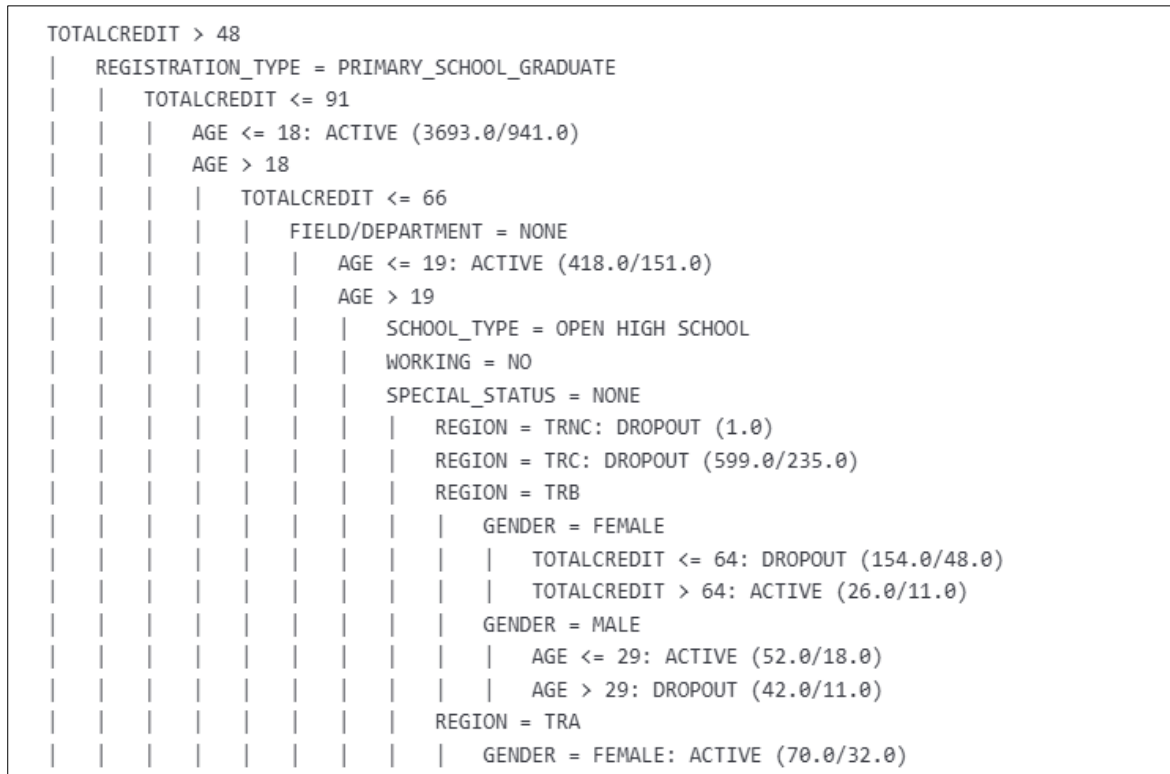
Figure 5 presents the feature importance weights, revealing each predictor's contribution to model performance. Academic progress indicators (total credit count, enrollment type) demonstrate greater predictive power than demographic variables.

Interpretation of Decision Rules

The J48 algorithm generated a complex decision tree (2,832 branches, 2,045 leaves) expressible as IF-THEN rules. A section from the tree structure of the J48 model is given in Figure 6.

Figure 6

A Section From the Tree Structure of the J48 Model



While the full tree is impractical to display, representative rules highlight the model's logic as shown in Figure 7.

Figure 7

IF-THEN Decision Rule Examples of the J48 Model



These examples show the model combining multiple features, with total credit count often serving as a primary splitting criterion, reflecting its high importance.

Conclusion and Discussion

This study employed EDM classification analysis on a large dataset ($N = 484,158$) to predict the end-of-term status of students in Turkish OHSs. Departing from common binary classifications, this research's three-class approach (graduated, active, dropped out), despite complexities such as imbalance, offers a more nuanced representation of learner trajectories, contributing to an underdeveloped EDM area (Martins et al., 2023).

Following the CRISP-DM framework, we applied data transformation and manual feature selection to a dataset containing demographic, academic, and administrative features. Multiple classification algorithms were evaluated, acknowledging that no single algorithm is universally superior. The J48 algorithm emerged as the most effective, achieving an 80.47% correct classification rate (CCR) and a kappa value of 0.61, accurately predicting the status for four out of five students in the test set. This performance is notable given the three-class complexity and surpasses many binary classification studies. While J48 often performs well in EDM (Al-Barrak & Al-Razgan, 2016; Dekker et al., 2009; Kabakchieva, 2013; Kumar & Vijayalakshmi, 2011; Mishra et al., 2014; Sumitha & Vinothkumar, 2016), other algorithms such as random forest (Şara et al., 2015), approximate nearest neighbor (Ibrahim & Rusli, 2007), support vector machine (Delen, 2010), and k nearest neighbors (Yukselturk et al., 2014) have shown better results in specific contexts, likely due to differences in dataset size, feature structures, feature variety, and class distributions.

Through the lens of established student persistence theories, total credit count emerged as the most important predictor in this study, strongly aligning with Tinto's (1975) emphasis on academic integration. In OHSs, where traditional social integration is minimal, academic progress is the primary persistence anchor, likely reflecting achievement, motivation, and self-regulation crucial for distance learning (Zimmerman, 2002). This resonates with recent research showing academic progression's key role in distance learning retention (Colpo et al., 2024; Hannaford, 2021; Perez et al., 2018) and highlights its greater importance when social integration is limited. Shahiri and Husain (2015) also confirmed cumulative GPA (related to credits) as a reliable success indicator.

Enrollment type, the second key predictor, supports findings by Allensworth and Easton (2007) and Nagy and Molontay (2018) and aligns with Bean and Metzner's (1985) model emphasizing background and external factors for non-traditional students. This likely proxies for unobserved variables such as prior academic preparedness (students transferring after failure face higher risks; Bedel, 2013; Jimerson & Ferguson, 2007), socioeconomic status, and initial adaptation to distance learning. Its significance highlights the impact of pre-entry conditions (Astin, 1985) and life-course factors, especially for transfers from conventional schools due to academic challenges, facing academic challenges, and higher dropout risks (Bedel, 2013; Jimerson & Ferguson, 2007).

Age, disability status, region, employment, and gender were also important predictors, confirming learner characteristics' importance in distance education (Bozkurt et al., 2015; Durak et al., 2017). Intriguingly, academic/administrative variables (credits, enrollment type, age) were more predictive than demographics (gender, employment). This might suggest that in this large-scale distance setting, individual academic momentum and entry characteristics exert a more direct influence than demographic attributes, whose impact might be mediated differently than in face-to-face settings. It challenges assumptions about universal predictors, suggesting that in this context, Tinto's academic integration and Bean and Metzner's pre-entry/background factors are particularly influential due to reduced social integration and heightened demands for individual drive.

The model struggled to predict the active/delayed class (recall = 45.8%), due to group heterogeneity and the static nature of administrative data. Static data lacks dynamic indicators (engagement, momentum, intentions) needed to distinguish slow progressors from potential dropouts. Dynamic engagement data or qualitative methods could better illuminate this group's specific challenges and motivations (cf. Bowers, 2010).

Limitations include relying solely on administrative data, omitting variables such as family background, prior achievement (e.g., middle school GPA), psychometric factors, dynamic engagement metrics (e.g., login frequency), and nuanced socioeconomic indicators. Including such data improved accuracy elsewhere (Márquez-Vera et al., 2013; Şara et al., 2015) and could better differentiate risk levels, especially for the active/delayed-graduation group. Self-reported data reliability (e.g., employment) is another limitation. Despite limitations, this study contributes to a large-scale, multi-class predictive analysis within Türkiye's open secondary education system. It empirically shows academic progress and initial enrollment conditions outweigh demographics here, refining persistence theories (Bean & Metzner, 1985; Tinto, 1975) for large-scale distance learning. The findings offer an evidence base for targeted, data-driven interventions and policies, moving beyond generic assumptions. Future research should integrate comprehensive, dynamic data, explore real-time prediction, and use clustering to better understand student subgroups.

Recommendations

Based on the findings and discussion, the following recommendations are proposed:

For Practitioners

The implementation of data-driven early warning systems (EWS) is strongly recommended to identify at-risk students using key predictors such as low credit accumulation rates, high-risk enrollment types, and specific age profiles, with these systems integrated into existing infrastructure while addressing technical needs, data integration, staff training, and ethical guidelines. Targeted interventions should be developed based on specific predictors: for students with low credit accumulation, institutions should offer proactive academic counseling, study skills workshops, and review sessions for failed courses; for high-risk enrollment types (particularly transfers due to failure), mandatory enhanced orientation programs, mentor assignments, and early support access are essential. Academic momentum should be closely monitored by tracking credit accumulation rates over time, triggering interventions when progress slows relative to enrollment duration, and implementing system alerts for below-threshold credit counts.

Institutions should enhance onboarding and communication by tailoring activities to enrollment type and identified risks, using targeted communications for important reminders and motivational feedback. Systematic investigation of barriers through periodic surveys and focus groups, especially targeting at-risk students, can help understand specific obstacles and develop strategies to address issues such as exam absenteeism. Practitioners should advocate for enriching enrollment data with crucial information such as prior academic achievement, detailed socioeconomic indicators, and distance education readiness to improve predictive accuracy. Where available, course management system integration should be explored to leverage interaction data for finer-grained analysis of student progress and risk identification.

For Researchers

Future research should integrate diverse data sources while examining heterogeneity within student groups, especially the active/delayed graduation cohort, through mixed-methods approaches. Longitudinal modeling techniques such as survival analysis would track evolving risks, while advanced machine learning and explainable artificial intelligence could improve prediction accuracy without sacrificing interpretability (Er, 2023; Goran et al., 2024). Model refinement through hyperparameter optimization and class imbalance techniques offers performance enhancement opportunities as demonstrated by recent studies (Althaqafi et al., 2025; Sahlaoui et al., 2024). Implementation of early warning systems requires rigorous evaluation of their impact on retention outcomes. Cross-context comparisons across educational systems would identify universal versus context-specific predictors, while clustering analyses could reveal distinct student subgroups. Continued exploration of diverse algorithms and validation methods remains essential for robust research comparisons.

Limitations

This study's limitations include restricted feature availability from the administrative student information system, omitting potentially influential variables such as prior academic achievement, detailed socioeconomic factors, and behavioral indicators. These omissions may constrain the models' ability to fully capture persistence factors and affect predictor interpretation. Additionally, computational constraints with our large dataset ($N = 484,158$) necessitated using stratified holdout

validation rather than k-fold cross-validation. These limitations suggest opportunities for future research leveraging more comprehensive datasets and enhanced computational resources.

Acknowledgements

This study is derived from the PhD dissertation by the first author, Ahmet Polat, submitted to Sakarya University, Institute of Educational Sciences, titled “Examining Dropout and Graduation Status of Open High School Students Using Educational Data Mining”, supervised by the second author, Professor Mehmet Barış Horzum.

References

- Adıgüzel, M. U. (2016). *Öğrencilerin mesleki açık öğretim lisesini seçme nedenleri (Kayseri ili örneği)* [The reasons for selecting vocational open education high school by students (Kayseri sample)] (Thesis No: 446354) [Master's thesis, Erciyes University]. YÖK Thesis Center. <https://tinyurl.com/38wupbwe>
- Al-Barrak, M. A., & Al-Razgan, M. (2016). Predicting students final GPA using decision trees: A case study. *International Journal of Information and Education Technology*, 6(7), 528–533. <https://doi.org/10.7763/IJiet.2016.V6.745>
- Allensworth, E. M., & Easton, J. Q. (2007). *What matters for staying on-track and graduating in Chicago public high schools: A close look at course grades, failures, and attendance in the freshman year*. Consortium on Chicago School Research. <https://consortium.uchicago.edu/sites/default/files/2018-10/07%20What%20Matters%20Final.pdf>
- Altair Engineering (2026). *RapidMiner* (Version 9) [Computer software]. <https://docs.rapidminer.com/9.0/studio/getting-started/index.html>
- Althaqafi, T., Saleem, F., & AL-Ghamdi, A. A. M. (2025). Enhancing student performance prediction: The role of class imbalance handling in machine learning models. *Discover Computing*, 28(1), Article 79. <https://doi.org/10.1007/s10791-025-09576-4>
- Astin, A. W. (1985). *Achieving educational excellence*. Jossey-Bass.
- Aydın, A., Sarier, Y., & Uysal, Ş. (2012). Sosyoekonomik ve sosyokültürel değişkenler açısından PISA matematik sonuçlarının karşılaştırılması [The comparative assessment of the results of PISA mathematical literacy in terms of socioeconomic and sociocultural variables]. *Eğitim ve Bilim*, 37(164), 20–29. <https://doi.org/10.15390/ES.2012.1024>
- Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 61–75). Springer. https://doi.org/10.1007/978-1-4614-3305-7_4
- Bean, J. P., & Metzner, B. S. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research*, 55(4), 485–540. <https://doi.org/10.3102/00346543055004485>
- Bedel, A. (2013). Sınıf tekrarı yapan ve yapmayan öğrencilerin akademik güdülenme ve kaygı düzeylerinin karşılaştırılması [Comparison of academic motivation and anxiety levels of students in classes repeated and not repeated]. *Milli Eğitim Dergisi*, 43(200), 111–122. <https://dergipark.org.tr/tr/download/article-file/457551>
- Bedel, Y. (2006). *Açıköğretim lisesi mesleki açıköğretim programı akademik danışmanlık hizmetlerinin etkililiğinin değerlendirilmesi* [Evaluation of the effectiveness of open education high school vocational curriculum academic advisory services] (Thesis No: 204612) [Master's thesis, Ankara University]. YÖK Thesis Center. <https://tinyurl.com/4r72kf9v>

- Ben-David, A. (2008). Comparison of classification accuracy using Cohen's weighted kappa. *Expert Systems with Applications*, 34(2), 825–832. <https://doi.org/10.1016/j.eswa.2006.10.022>
- Berens, J., Schneider, K., Gortz, S., Oster, S., & Burghoff, J. (2019). Early detection of students at risk—Predicting student dropouts using administrative student data from German universities and machine learning methods. *Journal of Educational Data Mining*, 11(3), 1–41. <https://doi.org/10.5281/zenodo.3594771>
- Bowers, A. J. (2010). Analyzing the longitudinal K-12 grading histories of entire cohorts of students: Grades, data driven decision making, dropping out and hierarchical cluster analysis. *Practical Assessment, Research, and Evaluation*, 15(1), Article 7. <https://doi.org/10.7275/r4zq-9c31>
- Bozkurt, A., Akgun-Ozbek, E., Yilmazel, S., Erdogdu, E., Ucar, H., Guler, E., Sezgin, S., Karadeniz, A., Sen-Ersoy, N., Goksel-Canbek, N., Dincer, G. D., Ari, S., & Aydin, C. H. (2015). Trends in distance education research: A content analysis of journals 2009–2013. *The International Review of Research in Open and Distributed Learning*, 16(1), 330–363. <https://doi.org/10.19173/irrodl.v16i1.1953>
- Çalışkan, Ş., Karabacak, M., & Meçik, O. (2013). Türkiye’de eğitim-ekonomik büyüme ilişkisi: 1923–2011 (Kantitatif bir yaklaşım) [Relationship between education and economic growth in Turkey: 1923–2011 (A quantitative approach)]. *Yönetim Bilimleri Dergisi*, 11(21), 29–48. <https://dergipark.org.tr/tr/download/article-file/46273>
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0 Step-by-step data mining guide*. SPSS Inc. <https://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/14.2/es/CRISP-DM.pdf>
- Çiçek, S. (2005). *Kız meslek lisesi açık lise programı uygulamalarında karşılaşılan sorunlar ve çözüm önerileri (Elâzığ, Malatya ve Diyarbakır İlleri örneği)* [Problems in practice in girls' open vocational secondary school and recommendations for solutions (Elazığ, Malatya, and Diyarbakır samples)] (Thesis No: 188405) [Master's thesis, Fırat University]. YÖK Thesis Center. <https://tinyurl.com/4j6wep9t>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Colpo, M. P., Primo, T. T., de Aguiar, M. S., & Cechinel, C. (2024). Educational data mining for dropout prediction: Trends, opportunities, and challenges. *Revista Brasileira de Informática na Educação*, 32, 220–256. <https://doi.org/10.5753/rbie.2024.3559>
- Çuhadar Öncü, E. (2017). *Ortaöğretim öğrencilerinin açık öğretim lisesine geçiş nedenlerine ilişkin öğrenci, okul yöneticisi, öğretmen ve veli görüşleri* [Opinions of students, principals, teachers, and parents about factors leading to high school students enrolling in open high school] (Thesis No: 468228) [Master's thesis, Ankara University]. YÖK Thesis Center. <https://tinyurl.com/mtdzbyje>

- Dekker, G. W., Pechenizkiy, M., & Vleeshouwers, J. M. (2009). Predicting students drop out: A case study. In T. Barnes, M. Desmarais, C. Romero, & S. Ventura (Eds.), *Proceedings of the International Working Group on Educational Data Mining* (pp. 41–50). International Working Group on Educational Data Mining.
<https://www.educationaldatamining.org/EDM2009/uploads/proceedings/dekker.pdf>
- Delen, D. (2010). A comparative analysis of machine learning techniques for student retention management. *Decision Support Systems*, 49(4), 498–506.
<https://doi.org/10.1016/j.dss.2010.06.003>
- Demiray, U., & Sağlık, M. (2003). Açıköğretim fakültesi ve Açıköğretim lisesi uygulamalarını içeren araştırmalara ilişkin bir değerlendirme [General evaluation of the research on open education faculty and open high school applications]. *The Turkish Online Journal of Educational Technology*, 2(4), 50–59. <https://tojet.net/articles/v2i4/248.pdf>
- Demirtaş, Z., Tutkun, Ö. F., & Arslan, A. (2017). Mesleki açık öğretim lisesi (MAÖL) öğrencilerinin mesleki eğitime yönelik görüşleri [Vocational open education high school students' opinions on vocational education]. *PESA Uluslararası Sosyal Araştırmalar Dergisi*, 3(4), 231–240.
<https://dergipark.org.tr/tr/download/article-file/454515>
- Dere, S. (2002). Açıköğretim lisesine kayıtlı engelli öğrencilerin sisteme ilişkin görüş ve beklentilerinin değerlendirilmesi [The evaluation of expectations and opinions of disabled students who are enrolled at open high school] (Thesis No: 117551) [Master's thesis, Ankara University]. YÖK Thesis Center. <https://tez.yok.gov.tr>
- Djulovic, A., & Li, D. (2013). Towards freshman retention prediction: A comparative study. *International Journal of Information and Education Technology*, 3, 494–500.
<https://www.ijiet.org/papers/324-KO45.pdf>
- Durak, G., Çankaya, S., Yünkül, E., Urfa, M., Topraklıklioğlu, K., Arda, Y., & İnam, N. (2017). Trends in distance education: A content analysis of master's thesis. *TOJET: The Turkish Online Journal of Educational Technology*, 16(1), 203–218.
<https://files.eric.ed.gov/fulltext/EJ1124887.pdf>
- Er, E. (2023). An explainable machine learning approach to predicting and understanding dropouts in MOOCs. *Kastamonu Education Journal*, 31(1), 143–154.
<https://doi.org/10.24106/kefdergi.1246458>
- Ereş, F. (2005). Eğitimin sosyal faydaları: Türkiye-AB karşılaştırması [Social contribution of education in Turkey: Comparison of Turkey and European Union]. *Milli Eğitim Dergisi*, 33(167). https://dhgm.meb.gov.tr/yayinlar/dergiler/Milli_Egitim_Dergisi/167/orta3-eres.htm
- Eurostat. (2026). *NUTS—Nomenclature of territorial units and statistical regions—Level 1*. European Union. <https://ec.europa.eu/eurostat/documents/345175/17779945/2021-NUTS-1-map.pdf/od90fc54-903c-347b-8a12-aa6ced7a4d25?t=1698686269555>

- Goran, R., Jovanovic, L., Bacanin, N., Stanković, M. S., Simic, V., Antonijevic, M., & Zivkovic, M. (2024). Identifying and understanding student dropouts using metaheuristic optimized classifiers and explainable artificial intelligence techniques. *IEEE Access*, 12, 122377–122400. <https://doi.org/10.1109/ACCESS.2024.3446653>
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining concepts and techniques* (3rd ed.). Morgan Kaufmann. <https://doi.org/10.1016/C2009-0-61819-5>
- Hannaford, L., Cheng, X., & Kunes-Connell, M. (2021). Predicting nursing baccalaureate program graduates using machine learning models: A quantitative research study. *Nurse Education Today*, 99, Article 104784. <https://doi.org/10.1016/j.nedt.2021.104784>
- He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263–1284. <https://doi.org/10.1109/TKDE.2008.239>
- Heppen, J. B., & Therriault, S. B. (2008). *Developing early warning systems to identify potential high school dropouts* [Issue Brief]. National High School Center. <https://files.eric.ed.gov/fulltext/ED521558.pdf>
- Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2), 1–11. <https://doi.org/10.5121/ijdkp.2015.5201>
- Ibrahim, Z., & Rusli, D. (2007, September 5). *Predicting students' academic performance: Comparing artificial neural network, decision tree and linear regression* [Presentation]. 21st Annual SAS Malaysia Forum, Kuala Lumpur. https://www.researchgate.net/publication/228894873_Predicting_Students'_Academic_Performance_Comparing_Artificial_Neural_Network_Decision_Tree_and_Linear_Regression
- International Educational Data Mining Society. (n.d.). *Home*. <https://educationaldatamining.org/>
- Jensen, K. (2012). *CRISP-DM process diagram*. Wikimedia Commons. <https://commons.wikimedia.org/w/index.php?curid=24930610>
- Jimerson, S. R., & Ferguson, P. (2007). A longitudinal study of grade retention: Academic and behavioral outcomes of retained students through adolescence. *School Psychology Quarterly*, 22(3), 314–339. <https://doi.org/10.1037/1045-3830.22.3.314>
- Kabakchieva, D. (2013). Predicting student performance by using data mining methods for classification. *Cybernetics and Information Technologies*, 13(1), 61–72. <https://doi.org/10.2478/cait-2013-0006>
- Kılınç, Ç. (2015). *Üniversite öğrenci başarısı üzerine etki eden faktörlerin veri madenciliği yöntemleri ile incelenmesi* [Examining the effects on university student success by data mining techniques] (Thesis No: 415460) [Master's thesis, Eskişehir Osmangazi University]. YÖK Thesis Center. <https://tinyurl.com/2yvzw3xu>

- Kızmaz, Z. (2004). Öğrenim düzeyi ve suç: Suç-okul ilişkisi üzerine sosyolojik bir araştırma [Educational level and crime: A sociological research on the relation between school and crime]. *Firat University Journal of Social Science*, 14(2), 291–319.
<https://dergipark.org.tr/en/download/article-file/71991>
- Kovacic, Z. J. (2010). Early prediction of student success: Mining students enrolment data. In *Proceedings of Informing Science & IT Education Conference (InSITE) 2010* (pp. 647–665). Informing Science Institute. <https://doi.org/10.28945/1281>
- Kumar, S. A., & Vijayalakshmi, M. N. (2011). Efficiency of decision trees in predicting student's academic performance. In D. C. Wyld & M. Wozniak (Eds.), *First International Conference on Computer Science, Engineering and Applications: CCSEA proceedings* (pp. 335–343). AIRCC Publishing. <https://airccj.org/CSCP/vol1/csit1230.pdf>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Lassibille, G., & Navarro-Gomez, L. (2008). Why do higher education students drop out? Evidence from Spain. *Education Economics*, 16(1), 89–105.
<https://doi.org/10.1080/09645290701523267>
- Lee, S., & Chung, J. Y. (2019). The machine learning-based dropout early warning system for improving the performance of dropout prediction. *Applied Sciences*, 9(15), Article 3093. <https://doi.org/10.3390/app9153093>
- Luy, M., Zannella, M., Wegner-Siegmundt, C., Minagawa, Y., Lutz, W., & Caselli, G. (2019). The impact of increasing education levels on rising life expectancy: A decomposition analysis for Italy, Denmark, and the USA. *Genus*, 75, Article 11. <https://doi.org/10.1186/s41118-019-0055-0>
- Márquez-Vera, C., Cano, A., Romero, C., Noaman, A. Y. M., Mousa Fardoun, H., & Ventura, S. (2016). Early dropout prediction using data mining: A case study with high school students. *Expert Systems*, 33(1), 107–124. <https://doi.org/10.1111/exsy.12135>
- Márquez-Vera, C., Morales, C. R., & Soto, S. V. (2013). Predicting school failure and dropout by using data mining techniques. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, 8(1), 7–14. <https://doi.org/10.1109/RITA.2013.2244695>
- Martins, M. V., Baptista, L., Machado, J., & Realinho, V. (2023). Multi-class phased prediction of academic performance and dropout in higher education. *Applied Sciences*, 13(8), Article 4702. <https://doi.org/10.3390/app13084702>
- Mishra, T., Kumar, D., & Gupta, S. (2014). Mining students' data for prediction performance. In J. P. Mittal (Chair), *Proceedings: Fourth International Conference on Advanced Computing & Communication Technologies* (pp. 255–262). IEEE. <https://doi.org/10.1109/ACCT.2014.105>
- Ministry of National Education. (2012). *Millî eğitim İstatistikleri: Örgün eğitim 2011–2012* [National education statistics: Formal education 2011-2012].

https://sgb.meb.gov.tr/meb_iys_dosyalar/2025_06/02135804_meb_istatistikleri_orgun_egitim_2011_2012.pdf

Ministry of National Education. (2015). *Millî Eğitim Bakanlığı: 2015–2019 stratejik planı* [Ministry of National Education: 2015-2019 Strategic plan].

http://sgb.meb.gov.tr/meb_iys_dosyalar/2015_09/10052958_10.09.2015sp17.15imzasz.pdf

Ministry of National Education. (2017). *Millî eğitim İstatistikleri: Örgün eğitim 2016/17* [National education statistics: Formal education 2016/17].

https://sgb.meb.gov.tr/meb_iys_dosyalar/2017_09/08151328_meb_istatistikleri_orgun_egitim_2016_2017.pdf

Ministry of National Education. (2019). *Millî eğitim İstatistikleri: Örgün eğitim 2018/19* [National education statistics: Formal education 2018/19].

http://sgb.meb.gov.tr/meb_iys_dosyalar/2019_09/30102730_meb_istatistikleri_orgun_egitim_2018_2019.pdf

Ministry of National Education. (2021). *Millî eğitim İstatistikleri: Örgün eğitim 2020/21* [National education statistics: Formal education 2020/21].

https://sgb.meb.gov.tr/meb_iys_dosyalar/2021_09/10141326_meb_istatistikleri_orgun_egitim_2020_2021.pdf

Ministry of National Education. (2022). *Ortaöğretim kurumlarına ilişkin merkezi sınav* [The central exam for secondary education institutions].

https://cdn.eba.gov.tr/icerik/2022/06/2022_LGS_rapor.pdf

Nagy, M., & Molontay, R. (2018). Predicting dropout in higher education based on secondary school performance. In *2018 IEEE 22nd International Conference on Intelligent Engineering Systems (INES)* (pp. 000389–000394). IEEE.

<https://doi.org/10.1109/INES.2018.8523888>

Organisation for Economic Co-operation and Development (2018). *Education at a glance 2018: OECD indicators*. OECD.

<https://doi.org/10.1787/eag-2018-en>

Organisation for Economic Co-operation and Development. (2021). *Education at a glance 2021: OECD indicators*. OECD.

<https://doi.org/10.1787/b35a14e5-en>

Özkahveci, Ö. (2001). *Açıköğretim lisesi mesleki açık öğretim programı öğrencileri ile kız meslek liseleri öğrencilerinin akademik başarılarının karşılaştırılması* [Comparison of academic achievement of open high school vocational education program students and girls' vocational high school students] (Thesis No: 117810) [Master's thesis, Gazi University]. YÖK Thesis Center. <https://tez.yok.gov.tr>

Perez, B., Castellanos, C., & Correal, D. (2018). Predicting student drop-out rates using data mining techniques: A case study. In A. Orjuela-Canon, J. Figueroa-Garcia, & J. Arias-Londono (Eds.), *Applications of computational intelligence, ColCACI 2018* (pp. 111–125). Springer.

https://doi.org/10.1007/978-3-030-03023-0_10

- Polat, A. (2021). *Açık öğretim liseleri öğrencilerinin okul terki ve mezuniyet durumlarının eğitsel veri madenciliği ile incelenmesi* [Examining dropout and graduation status of open high school students using educational data mining] (Thesis No: 681548) [Doctoral dissertation, Sakarya University]. YÖK Thesis Center. <https://tinyurl.com/22k97c8s>
- Randler, C., Horzum, M. B., & Vollmer, C. (2014). The influence of personality and chronotype on distance learning willingness and anxiety among vocational high school students in Turkey. *The International Review of Research in Open and Distributed Learning*, 15(6), 93–110. <https://doi.org/10.19173/irrodl.v15i6.1928>
- Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6), 601–618. <https://doi.org/10.1109/TSMCC.2010.2053532>
- Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), Article e1355. <https://doi.org/10.1002/widm.1355>
- Rosenzweig, M. R. (2010). Microeconomic approaches to development: Schooling, learning, and growth. *Journal of Economic Perspectives*, 24(3), 81–96. <https://doi.org/10.1257/jep.24.3.81>
- Rumberger, R. W. (2020). The economics of high school dropouts. In S. Bradley & C. Green (Eds.), *The economics of education: A comprehensive overview* (2nd ed., pp. 149–158). Academic Press. <https://doi.org/10.1016/B978-0-12-815391-8.00012-4>
- Rumberger, R. W., & Lim, S. A. (2008). *Why students drop out of school: A review of 25 years of research*. California Dropout Research Project. <https://ej.issuelab.org/resources/11658/11658.pdf>
- Şahin, B. (2017). *Açık öğretim lisesi öğrenci ve mezunlarının katılım örüntüleri (Ankara İli Çankaya İlçesi örneği)* [Participation patterns of open high school students and graduates (Ankara city Çankaya district sample)] (Thesis No: 468273) [Doctoral dissertation, Ankara University]. YÖK Thesis Center. <https://tinyurl.com/4kk3s7jx>
- Sahlaoui, H., Alaoui, E. A. A., Agoujl, S., & Nayyar, A. (2024). An empirical assessment of SMOTE variants techniques and interpretation methods in improving the accuracy and the interpretability of student performance models. *Education and Information Technologies*, 29(5), 5447–5483. <https://doi.org/10.1007/s10639-023-12007-w>
- Şara, N.-B., Halland, R., Igel, C., & Alstrup, S. (2015). High-school dropout prediction using machine learning: A Danish large-scale study. In M. Verleysen (Ed.), *Proceedings. ESANN 2015: 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning* (pp. 319–324). ESANN. <https://core.ac.uk/download/pdf/269280541.pdf>
- Sarıhan, Ş. (2010). *Mesleki açık öğretim lisesindeki öğrenci hizmetlerinin etkililiği konusunda öğrenci görüşleri (Ankara ili örneği)* [Student opinions on effectiveness of student service at open vocational high school (Ankara sample)] (Thesis No: 279946) [Master's thesis, Ankara University]. YÖK Thesis Center. <https://tinyurl.com/bjupprjn>

- Şentürk, E. (2009). *Mesleki açık öğretim lisesi bilişim teknolojileri alanında verilen eğitimin etkililiğine yönelik öğrenci görüşlerinin değerlendirilmesi* [Evaluation of students' opinions towards effectiveness of education in vocational open high school's information and communication technologies department] (Thesis No: 235277) [Master's thesis, Uludağ University]. YÖK Thesis Center. <https://tinyurl.com/dsrpwx4r>
- Shahiri, A. M., & Husain, W. (2015). A review on predicting student's performance using data mining techniques. *Procedia Computer Science*, 72, 414–422. <https://doi.org/10.1016/j.procs.2015.12.157>
- Sipahi, K. B. (2019). *Mesleki açık öğretim lisesi öğrencilerinin uzaktan eğitime yönelik algıladıkları engeller ile tutumları arasındaki ilişkinin incelenmesi* [Researching the relationship between barriers and attitude perceived by vocational open high school students towards distance education] (Thesis No: 538965) [Master's thesis, Sakarya University]. YÖK Thesis Center. <https://tinyurl.com/5n8u23a4>
- Soylu, S. (2014). *Mesleki açık öğretim lisesi muhasebe eğitiminde karşılaşılan sorunlar ve çözüm önerileri* [Problems encountered in accounting education at vocational open high school and proposed solutions] (Thesis No: 366242) [Master's thesis, Gazi University]. YÖK Thesis Center. <https://tinyurl.com/c6x5czw9>
- Sumitha, R., & Vinothkumar, E. S. (2016). Prediction of students outcome using data mining techniques. *International Journal of Scientific Engineering and Applied Science (IJSEAS)*, 2(6), 132–139. <https://ijseas.com/volume2/v2i6/ijseas20160615.pdf>
- Tanner, D. E., Newbold, B. L., & Johnson, D. B. (2003). *Academic achievement as a dropout predictor* (ED478173). ERIC. <https://files.eric.ed.gov/fulltext/ED478173.pdf>
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89–125. <https://doi.org/10.3102/00346543045001089>
- Türkiye Cumhuriyeti Resmî Gazete [Republic of Türkiye official gazette], No. 23084. (1997, August 18). Law no. 4306 (pp. 2–6). <https://www.resmigazete.gov.tr/arsiv/23084.pdf>
- Türkiye Cumhuriyeti Resmî Gazete [Republic of Türkiye official gazette], No. 26215. (2006, July 1). Dokuzuncu kalkınma planı [Ninth development plan]. <https://www.resmigazete.gov.tr/eskiler/2006/07/20060701M1-2.pdf>
- Türkiye Cumhuriyeti Resmî Gazete [Republic of Türkiye official gazette], No. 28261. (2012, April 11). Law no. 6287. <http://www.resmigazete.gov.tr/eskiler/2012/04/20120411-8.htm>
- Türkiye Cumhuriyeti Resmî Gazete [Republic of Türkiye official gazette], No. 28699. (2013, July 6). Onuncu kalkınma planı [Tenth development plan]. <https://www.resmigazete.gov.tr/eskiler/2013/07/20130706M1-1-1.doc>
- Umar, U. (2016). *Evaluation of school administrators by using data mining techniques* (Thesis No: 436503) [Master's thesis, Fırat University]. YÖK Thesis Center. <https://tinyurl.com/3scyjuji>

- United Nations Department of Economic and Social Affairs. (2015). *Transforming our world: The 2030 Agenda for Sustainable Development*. <https://sdgs.un.org/2030agenda>
- United Nations Development Program. (2021). *Human development report 2020: The next frontier: Human development and the Anthropocene*. <https://hdr.undp.org/content/human-development-report-2020>
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data mining: Practical machine learning tools and techniques* (4th ed.). Morgan Kaufmann. <https://doi.org/10.1016/C2015-0-02071-8>
- Yavuz, H. (2014). *Mesleki açık öğretim lisesi öğrencilerinin sunulan hizmetlerin etkililiği konusunda görüşleri ve motivasyon düzeyleri* [Student opinions on effectiveness of student services and motivation levels at vocational open high school] (Thesis No: 367965) [Master's thesis, Fatih University]. YÖK Thesis Center. <https://tinyurl.com/3jpe74k4>
- Yukselturk, E., Ozekes, S., & Türel, Y. K. (2014). Predicting dropout student identification: An application of data mining methods in an online education program. *European Journal of Open, Distance and e-Learning*, 17(1), 118–133. <https://doi.org/10.2478/eurodl-2014-0008>
- Yurdakul, S. (2015). *Veri Madenciliği ile lise öğrenci performanslarının değerlendirilmesi* [Assessment of high school students' performance by means of data mining] (Thesis No: 418479) [Master's thesis, Kırıkkale University]. YÖK Thesis Center. <https://tinyurl.com/3t3syc6k>
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2



August – 2026

Book Review: *Mega-Universities and Opening Education by Design*

Editor: Mark Nichols (ICDE, 2025, xxii + 217 pages). ISBNs: 978-82-93172-63-5 (PDF), 978-82-93172-60-4 (paperback), 978-82-93172-61-1 (hardback), 978-82-93172-62-8 (Kindle)

Reviewed by: Junhong Xiao, *Open University of Shantou, China*

Mega-Universities and Opening Education by Design, edited by Mark Nichols, focuses on what current open, flexible, and distance learning (OFDL) can learn from mega-universities' economies of scale. This 2025 publication revisits Sir John Daniel's influential 1996 book, *Mega-Universities and Knowledge Media: Technology Strategies for Higher Education* (Daniel, 1996). In the intervening three decades, the global landscape of OFDL has undergone tremendous changes, following its mainstreaming into the ecosystem of higher education. As a tie-in, Daniel writes the foreword to this 2025 sequel. The chapter authors are internationally renowned OFDL scholars who use plain language to share their praxis in democratizing higher education.

My overview of this book starts with the title using the word *opening* as a verb. Explicitly, this denotes the various strategically deliberate actions by institutions to become more open. Four facets of higher education are explored throughout this book, namely to be (a) available and (b) inclusive (from the students' perspective); and (c) scalable and (d) sustainable (from the institutional perspective). The chapters are presented in three sections. The first is grounded in the design of OFDL, technology, management of mega-universities, and institutional resilience in the era of the COVID-19 pandemic. The second section focuses on what any university can learn about expanding their reach while also increasing their openness. The third section presents two case studies. One analyzes the evolution of Southern New Hampshire University (SNHU) from a residential institution into a mega-university. The other illustrates how the Open University of China, the world's largest university, pursues innovation to maintain its relevance to learners in the trades, higher education, and lifelong learning. Building on the collection of insights from each chapter, the editor has summarized the book in the form of "five strategic considerations for mega-universities and those seeking mega-university status: advocacy, differentiation, adaptability, alignment, and courage" (Nichols, 2025, p. 198).

Several points in the book resonated with me. The first was the specific case study of SNHU. The next three points are relationally intertwined, including technology, capacity building, and cost-effectiveness, which I tie together in my discussion of cost-effectiveness. I conclude by discussing quality, the thread that runs through all chapters of the book.

The case of SNHU, evolving from a small residential institution into a mega-university, is fascinating. At one time, mega-universities were the prerogative of open universities, as only these institutions had the potential to develop into mega-universities. The success of SNHU challenges this assumption and has far-reaching implications for campus-based universities facing strong competition from a variety of sources. This case study shows how campus-based institutions can learn from OFDL institutions and vice versa.

Technology use is indispensable to OFDL and mega-universities. Overall, the book conveys an optimistic stance on leveraging digital technologies, including generative artificial intelligence (GenAI), to expand higher education at scale. The hope is that educational technologies can break the iron triangle of access, quality, and cost. A strength of this book is that it is a collection of competing views. For example, many chapters stress GenAI's future innovations and reimagining of teaching and learning, while others discuss current struggles to deal with GenAI in higher education. However, rather than critique the risks and controversies related to technology in education, the book favours using cutting-edge technologies to support mega-scale access to higher education.

Capacity building is raised only in a few chapters in terms of initiatives for students, staff, and faculty. Chapter 1.5, drawing from a case study of Allama Iqbal Open University that serves over 600,000 students per semester, considers comprehensive capacity building to be among the measures essential to the success of mega-universities. Capacity building figures prominently within a framework for managing technological implementation and change at scale. The importance of capacity building for an institution makes it a critical element in designing mega-universities.

Research into cost effectiveness in higher education has made use of different measures for comparison, making it difficult to synthesize knowledge in this area. The book considers several facets of cost effectiveness within higher education, including as (a) a process constraint, (b) a strategic-competitive differentiator, (c) a transformative priority, (d) an operational business measure, and/or (e) a government requirement. It also presents it in relation to other measures, such as tracking productivity, a measure that combines student success with cost effectiveness (Nichols, 2025, p. 125). Procuring and maintaining technology is a core responsibility of mega-universities and OFDL. The acute problem is that the best pedagogical choice of an educational technology may not be the best financial choice for a given institution. In other non-educational business models, an increased cost of doing business can be passed along to the consumer who is buying a premium option (i.e., an expensive choice) in the sector. Unfortunately, an increased cost of technologies competes with the affordability aims of OFDL and mega-universities. Furthermore, only Chapter 1.3 takes students' perspective into account by introducing specific measures to increase affordability for students at the University of South Africa. One assessment of cost effectiveness of the technological change is to measure the appropriateness of the technology to build capacity.

In the editor's preface, quality compliance is explicitly assumed to exist within the four facets of higher education to be available, inclusive, scalable, and sustainable. While each chapter (except the Introduction) mentions how quality issues affect that chapter's discussion, the question of "how do we both serve more learners, and improve our quality at the same time?" (Nichols, 2025, p. 159) permeates the book. Of the many points raised in Chapter 1.4, there is a substantial exploration of quality assurance. Currently, higher education has a multitude of complex definitions for quality assurance. This chapter discusses a variety of quality assurance statements, guidelines, and best practices applicable to OFDL and mega-universities. It

presents the interplay among the scaling of education in mega-universities, rapid technology changes, and educational quality assurance. This interplay within mega-universities brings a unique nuance to the discussion, as harm to students can be exacerbated by either scaling up education of poor quality or scaling up good quality education poorly. Quality assurance is both critical and significant for all sizes of higher educational institutions.

I am happy to recommend this book to OFDL and mega-university institution leaders and management, practitioners, and researchers, as well as government education policy-makers and campus-based university leaders. I believe all will find it a valuable source of reference and inspiration when making policies, formulating institutional development plans, or designing innovations in learning and teaching. The book is forward-thinking, scholarly, practical, and written in plain language. It inspires an array of processes to increase openness in a higher educational institution.

References

Daniel, J. S. (1996). *Mega-universities and knowledge media: Technology strategies for higher education*. Routledge.

Nichols, M. (Ed.). (2025). *Mega-universities and opening education by design*. ICDE.



August – 2026

Novel Technologies in Nurse Practitioner Education: An Integrative Review

Chelsie Girard, Misha Balciunas, Kelly Robertson, Hannah Ballantyne, Jennifer-Lynn Fournier, and Robyn Gorham
Laurentian University School of Nursing, Sudbury, Ontario, Canada

Abstract

The use of technology in healthcare education is rapidly expanding. Undergraduate nursing programs have integrated advanced technologies such as high-fidelity simulation, immersive virtual reality, augmented reality, and holographic simulations to support learning and improve retention in both in-person and distance learning. Meanwhile, little is known about the use of these clinical learning technologies in nurse practitioner education. This integrative review involved a systematic search of five databases: CINAHL, ERIC (ProQuest), MedLine (OVID), Google Scholar, and PubMed. Eleven studies met the inclusion and exclusion criteria and were analyzed for major themes. Five primary themes emerged, including (a) student uptake, (b) student learning, (c) change management, (d) accessibility, and (e) confidence. Findings suggested that integrating advanced clinical learning technologies into nurse practitioner programs should be prioritized. Going forward, focus should turn to establishing nurse practitioner program best practices for integrating clinical learning technologies, budgetary considerations, ongoing evaluation of learning outcomes, and theory development. Work in these areas may support the enhanced use of clinical learning technologies while providing educational institutions, nurse practitioner programs and public funders with the necessary data to inform future planning and further implementation of new technologies.

Keywords: nurse practitioner education, technology, simulation, virtual reality, augmented reality, high fidelity simulation, change management

Introduction

Technology use within healthcare professional education has advanced quickly with widespread use of applications or apps, clinical calculators, diagnostic tools, and simulation. Clinical learning technologies such as high-fidelity simulation, virtual reality (VR), augmented reality (AR), and holographic simulation support student and professional learning while maintaining active engagement (Carley, 2015). High-fidelity simulation and video medicine simulations replicate real-world clinical scenarios to educate and evaluate learners (Ainslie & Bragdon, 2018). Virtual reality creates an immersive, computer-generated environment, often used for hands-on learning of clinical procedures and management of high-risk scenarios (Kyaw et al., 2019). Augmented reality overlays digital images (objects) in a real-world environment, allowing learners to interact with the object (e.g., simulated patient, fabricated surgical site) and the environment at the same time (Ficarra, 2020). Holographic technologies project interactive holograms into the learner's space. This technology has most often been used for procedures involving complex spatial maneuvers, such as surgery for congenital heart disease and other procedures involving significant dexterity (D'Aiello et al., 2023).

Clinical education technologies such as virtual simulation, holographic imaging for skill development, immersive VR, and assisted learning via artificial intelligence (AI) have proven successful in undergraduate nursing education (Chang et al., 2022; Hoyt & Hauze, 2023; Jallad, 2024; Medvec et al., 2023; Russell, 2023). They have been shown to (a) support clinical reasoning and decision-making skills; (b) optimize hands-on skills; (c) enhance confidence through simulation in safe, controlled environments; and (d) maintain student engagement (Casler et al., 2024). They have also been particularly useful across various models of distributed learning, given their ability to transcend space and time by supporting learners from a distance (Al-Hatem et al., 2018; Liu & Butzlaff, 2021; May, 2019).

Nurse practitioners (NPs), educated at the graduate level to perform advanced, comprehensive patient care, use a significant amount of technology in clinical practice, including cellular applications, camera-based diagnostic tools, handheld bedside ultrasound, lab analyzers, and others (Shelton & Reimer, 2018). However, little is known about the introduction and use of clinical learning technologies in NP education or continuing education programs. Meanwhile, NP education program accreditors have included assessment of technology use in their program reviews, with tours of simulation and learning labs often included in evaluation frameworks (Canadian Association of Schools of Nursing, 2022).

The purpose of this integrative review was to develop an understanding of the current body of empirical work on clinical and educational technologies used within NP education programs to determine what has been implemented and what is currently known. Findings and the identification of gaps may reveal opportunities to advance the use of clinical education technologies to support safe, engaging, high-quality NP education (Whittemore & Knafl, 2005). The impact on student outcomes, faculty uptake, student uptake, and collective experiences were explored (Chang et al., 2022; Wang et al., 2023). This review was undertaken to answer the following research question: What is currently known about applications and outcomes related to the use of clinical education technologies in graduate-level NP programs?

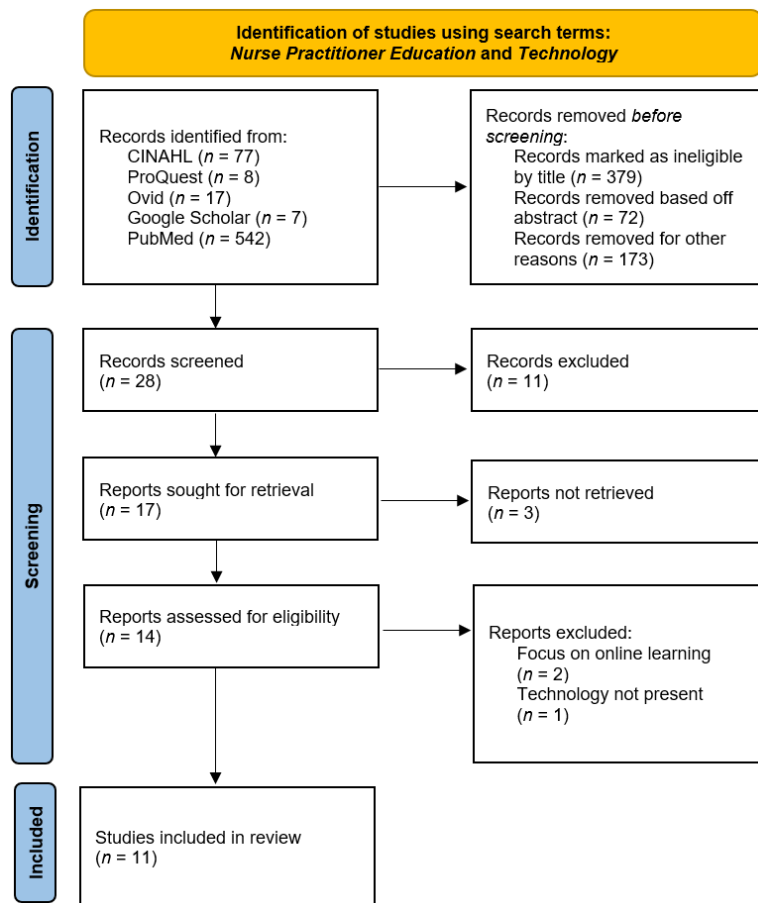
Methodology

Study Design

This integrative review used the framework from Whitemore and Knafl (2005) with its five-step process for review: (a) problem definition, (b) literature search, (c) data evaluation, (d) data analysis, and (e) interpretation and thematic presentation. This framework was selected because it offered a comprehensive and systematic method for synthesizing evidence across a variety of study designs, including both experimental and non-experimental research (Whitemore & Knafl, 2005). This flexibility in incorporating findings from studies with diverse study designs and methodologies was ideal for a topic where the literature base was still developing (Whitemore & Knafl, 2005). Additionally, having a structured approach ensured transparency and replicability, increasing the ability to produce meaningful insights for future practice and research (Cronin & George, 2020). The results of the literature search were reported via the preferred reporting items for systematic reviews and meta-analyses (PRISMA). See Figure 1 for details.

Figure 1

Nurse Practitioner Education and Technology PRISMA



Literature Search

A literature search was conducted via CINAHL, ERIC (ProQuest), Medline (OVID), Google Scholar, and PubMed. To gain an overall understanding of the current integration of clinical education technology in NP programs, authors searched for two terms: (a) nurse practitioner education, and (b) technology. The inclusion criteria consisted of articles published in English since 2010 that pertained to NP education and focused on clinical learning technology. The authors determined that literature from qualitative, quantitative, mixed methods research, and literature reviews were appropriate for inclusion, given the limited field of NP-focused work around the topic of interest. Exclusion criteria included literature published in languages other than English, literature published before 2010, and papers focused on technology integrated within undergraduate nursing programs and non-NP graduate programs.

The first literature search was limited to the previous five years and only included work focused on NP education in Canada, given our interest in current use, available technologies, and experience with these to assess the need for funding initiatives, training, and enhanced uptake within our regional NP education programs. This search drew numerous studies focused on undergraduate learners with very few items pertaining to the NP student population. We decided to broaden the search to include articles published from 2010 to present, and to extend the search to all articles rather than restricting to Canadian sources. The updated search parameters yielded 651 articles. The authors used a tiered process that began with a title and abstract review, allowing for the rapid exclusion of irrelevant studies. Before initiating the detailed review, authors reached consensus to exclude literature focused on online learning platforms, virtual lecture delivery, video conferencing, and telehealth systems. Additionally, articles addressing technology for professional development, patient education, and opinion pieces were also excluded. Upon completion of this review, 11 articles were retained for analysis to determine the current state of knowledge around the use of clinical and educational technologies in NP education.

Data Analysis

Braun and Clark's (2006) framework was used to guide our analysis. This framework provided clear guidance for organizing and coding data, identifying overarching themes, clarifying through revision and refining dominant themes (Braun & Clark, 2006). While designed for qualitative analysis, the authors were able to apply it within this review, given the nature of the task at hand.

Two of the four authors were randomly selected to serve as independent reviewers, and each was assigned review of five or six full articles. Each author independently analyzed the articles for specific themes. Using a dual reviewer process enabled collaboration and verification of relevant themes (Stoll et al., 2019). Each independent review validated that the item met inclusion and exclusion criteria and applied the conceptual framework to perform thematic analysis. Once complete, the authors reconvened to share their findings. When inconsistencies occurred during comparison, a third reviewer was brought in to review and clarify identified inconsistencies. This process allowed the authors to ensure clarity and relevance, and to identify valid themes to help guide future opportunities for integrating technology into NP student curricula.

Table 1 summarizes eleven articles and themes that emerged from the analysis. Code words were recorded as each article was read, and once all authors had completed their screening, the following five themes were identified: student uptake, student learning, change management, accessibility, and confidence.

Table 1

Summary of Articles and Thematic Analysis

Citation	Title	Objective/methodology/ participants	Findings/limitations	Theme
Anderson, Guido-Sanz, Díaz, Lok, Stuart, Akinnola, & Welch (2021)	Augmented reality in nurse practitioner education: Using a triage scenario to pilot technology usability and effectiveness	Mixed methods exploratory study to gain insight into augmented reality usability	AR can be useful and enjoyable, however requires more technological support	Change management Student learning Student uptake Confidence
Casler, Gawlik, & Messenger (2024)	Virtual reality to aid in competency based online nurse practitioner curriculum (VR-NP)	Pilot program for NP health assessment course to use transition to weekly VR simulation to support with increasing competency-based education	Realtime performance feedback, engaging, worthwhile Learning curve for faculty and students regarding software issues	Confidence Student learning Student uptake
Elliott, DeCristofaro, & Carpenter (2012)	Blending technology in teaching advanced health assessment in a family nurse practitioner program: Using personal digital assistants in a simulation laboratory	Studied the use of personal handheld devices and high-fidelity simulation in advanced health assessment course	Blended technology was identified as meeting students' learning needs; difficulty initially with buy in from course coordinators, value was identified once pilot scenarios occurred	Student uptake Change management
Lamarche, Park, Fraser, Rich, & MacKenzie (2016)	In the palm of your hand: Normalizing the use of mobile technology for nurse practitioner education and clinical practice	Use of mobile technology among NP students and the role it plays in NP education	Extremely positive reviews from students, including personal use of technology throughout workday and use for information sharing and in practice Negative view remains surrounding mobile technology	Student uptake Change management
Lee, Bathish, & Nelson (2024)	Transforming nursing education: developing augmented reality procedural training	Cross-sectional feasibility study to test training graduate students in advanced nursing procedures (lumbar puncture) using AR holographic images	Results were positive for feasibility and acceptance of this technology for training; requires financial commitment, curriculum alignment and student/faculty orientation	Student learning Student uptake Confidence Change management Accessibility

Citation	Title	Objective/methodology/ participants	Findings/limitations	Theme
Marrocco, Ginzburg, Feder (2019)	Seeing from the eyes of the beholder: A pilot program for learning through reflection in graduate nursing education	Using wearable technology to allow NP students the opportunity to view their patients' perspective of them throughout a clinical scenario	Students able to refine their interview taking skills and overall saw this as a beneficial and positive experience	Confidence Student uptake Student learning
Nelson, Ridgway, Brown, Anthamatten, Hall, McClure (2025).	Interactive simulation for addressing vaccine hesitancy in nurse practitioner education	Implementation and evaluation of AI virtual simulation to enhance students' communication skills	Overall positive feedback, although technology support required to keep system operating effectively	Student learning Change management
Rogers, Reyes, & Yang (2024).	Development of an interactive 3D visualization tutorial for pathophysiology in graduate nursing education	Teaching 3D visualization of COPD disease process, using AR; working with College of Engineering to create the program	Very positive result with enhanced learning, confidence and engagement	Confidence Student learning Student uptake
Rudolph, Vaughn, Crego, Hueckel, Kuszajewski, Molloy, Brisson, & Shaw (2017)	Integrating telepresence robots into nursing simulation	Use of a telepresence robot among nursing students and NP students clinical simulation, allowing those completing distance education to engage with remaining cohort	Positive experience to help engage those on campus and via distance learning Established that educators require enhanced learning to operate technology in order to provide successful implementation throughout course	Change management Student learning Accessibility
Smith, Jordan, & Li (2022)	Video-based interactive clinical simulation: Preparing nurse practitioner students for clinical	Development and evaluation of a video-based simulation for clinical experience	Resulted in increased confidence and positive feedback from NP students	Confidence
Whited, DeClerk, Berber, & Phelan (2019)	An innovative technique to promote understanding of anatomy for nurse practitioner students	Teaching anatomy to NP students using CT based 3D anatomy table and cadaver lab	Confidence and overall knowledge in all categories of pediatric anatomy increased from pretest to posttest Program is expensive	Change management Student learning Confidence

Results

Five themes emerged from our review of 11 articles, namely student uptake, student learning, change management, accessibility, and confidence. Each theme is discussed below, along with a brief discussion to relate the findings to existing theory. There are gaps in the evidence available to inform implementation and support integration of clinical learning technology in NP education. This underscores a need to address the role of relevant technologies in continuing education to support practicing NPs as lifelong learners.

Student Uptake

Student uptake was a dominant theme in seven articles; all highlighted a positive correlation between effective use of technology and skill development among NP students (Anderson et al., 2021; Casler et al., 2024; Elliott et al., 2012; Lamarche et al., 2016; Lee et al., 2024; Marrocco et al., 2019; Rogers et al., 2024). As students already engage with technology in their daily lives, integrating augmented reality and other clinical learning technologies into simulations enhances the realism of scenarios while maintaining a reasonable learning curve and student comfort levels similar to general technology use (Anderson et al., 2021; Lamarche et al., 2016). Using virtual simulation to replicate the patient's lived experience provided students with a valuable opportunity to exercise assessment and interviewing skills and communication strategies, especially around sensitive topics such as trauma, death and dying, sexual health, forensic care, and critical incidents. This learning will improve patient experience and outcomes in real NP encounters (Lee et al., 2024; Nelson et al., 2025). Meanwhile, several studies reported that NP students appreciated the direct and immediate feedback that AI could provide, reported satisfaction with their learning experience, and indicated they would use AI software for clinical learning in the future (Nelson et al., 2025; Rogers et al., 2024; Smith et al., 2022).

Student Learning

Rogers et al. (2024) found that interactive AR and VR tools support improved learning outcomes and increased motivation for learning compared to a traditional textbook approach to NP education. In Lee et al. (2024), an intervention that allowed NP students to use VR in simulated lumbar puncture also met with positive feedback from the students. The VR served as a safe bridge between the learning of theory and live clinical application. NP students noted that the VR system provided a realistic experience, was easy to use, and created a safe environment to practice motor, procedural, and critical thinking skills without the risk of patient harm (Anderson et al., 2021; Lee et al., 2025; Nelson et al., 2025). In another review of VR simulations, NP students appreciated the opportunity to debrief after difficult scenarios, practice soft skills and communication, and to become better equipped to handle crucial conversations around vaccine hesitancy (Anderson et al., 2021; Nelson et al., 2025; Whited et al., 2019).

Change Management

The theme of change management encompassed a broad range of suggestions for implementation of clinical education technologies into the standard NP curriculum. Adaptation by students and faculty varied and was context dependent. Rogers et al. (2024) highlighted collaboration between nursing and engineering faculty to create a 3D visualization tool for NP students to view an animation of the development of chronic obstructive pulmonary disease in real time. The authors noted that the iterative design generated greater understanding of the technology in both disciplines. Meanwhile, the partnership required developers to

operate the program and to make continual adjustments based on participant feedback (Rogers et al., 2024).

Several studies indicated that when faculty are supported in operating clinical education technology and troubleshooting glitches, program implementation is more efficient, and student satisfaction increases. Supporting faculty in the use of innovative technology in core courses in NP programs will improve competence, enhance the patient experience, and ameliorate better health outcomes (Lee et al., 2024; Rogers et al., 2024; Rudolph et al., 2017).

Accessibility

Technology can transcend borders and space, enable synchronous and asynchronous experiences, and increase accessibility for students. A VR holographic simulation for a clinical encounter or procedure can be conducted in a student's home, where the interaction can be video recorded for faculty review. Use of this technology across locations also ensures NP exposure to important experiences and patient conditions while decreasing logistics challenges related to on-site facilitation in real time and the need to hire and prepare simulated patients. More generally, the use of technology allows for collaboration between students on campus and those through distance education, ensuring each student is participating in the same clinical scenario and is practicing the same skills regardless of their location or ability to attend in person (Rudolph et al., 2017; Smith et al., 2022). This flexibility and facilitation of curricular continuity and collaboration will be particularly relevant within online distributed, on campus, and hybrid NP programs.

Clinical education technology creates opportunities for NP education programs given its contributions to student learning, flexibility, and its applicability to distributed learning models. Efficiencies related to geographic reach and the lack of need for dedicated spaces may facilitate increased NP program enrollment, ultimately supporting the long-term goal of meeting human resource demands in the health sector (Lee et al., 2024). AR and VR can mitigate challenges associated with limited clinical placements by offering students a wide array of controlled, realistic experiences related to various patient health challenges and clinical procedures (Nelson et al., 2025).

Confidence

NP students reported that clinical education technology used in simulation supported the development of a more comprehensive skill set, enhanced competency, and a sense of self efficacy during their initial clinical rotations (Nelson et al., 2025; Smith et al., 2022). Early studies reviewed the use of personal handheld devices and mobile technology in NP education and showed that point-of-care technology improved students' confidence and better prepared them for their NP clinical placements (Elliott et al., 2012; Lamarche et al., 2016). Throughout pretest and posttest surveys with a video-based interactive NP student simulation, average scores increased in all areas of confidence and knowledge acquisition (Smith et al., 2022). Across multiple studies, students reported that using clinical education technology better prepared them for in-person patient interviews, health assessments, and practical skills (Lee et al., 2024; Smith et al., 2022). Despite the long timeframe between the oldest and most recent studies in this review, each consistently showed that NP learners reported increased preparedness for in-person clinical practice when technology was leveraged.

Theoretical Linkages

This review did not uncover works presenting theory to support integration or evaluation of clinical education technologies for NPs. However, our current findings align well with Driscoll and Burners' (2022) constructivist approach to instruction and learning, which emphasized the importance of experience, schemas, collaboration, and complex, realistic and relevant environments. Of note, the inclusion of collaboration was significant as it is an important feature of NP practice within complex systems requiring team-based care. This model also emphasized reflexive practice, a key component in NP student learning and professional practice development.

Discussion and Recommendations

NP students have acknowledged the benefits of clinical learning technologies within programs. The integration of technologies like AR and VR have supported these learners in building self-confidence, provided a safe and non-judgmental learning environment, and enhanced the knowledge and skill required to understand disease processes and communicate effectively with patients (Anderson et al., 2021; Nelson et al., 2025; Rogers et al., 2024; Rudolph et al., 2017; Whited et al., 2021). Meanwhile, research and best practices related to the introduction, use, and uptake of clinical education technologies in NP education will require further development. NP programs must continue to implement, evaluate, and better understand the extent of the effectiveness, benefits, drawbacks, and challenges for both learners and faculty (Casler et al., 2024; Lee et al., 2024; Whited et al., 2019). While studies have supported the ongoing use of clinical learning technologies in NP education overall, studies have also identified challenges in implementation. For example, inconsistent functionality, difficulty navigating remote platforms, reliance on facilitators during simulations, and instances where the technology was perceived as uncomfortable or distracting were all reported in this review (Anderson et al., 2021; Marrocco, et al., 2019; Rudolph et al., 2017).

None of the publications presented the average cost of integrating clinical education technology into an NP education program. Strategies to address the rapid evolution of technology, the need for necessary updates, new platforms, best practices, and interoperability were also not mentioned. Data regarding current levels of use among NP programs were also not reported. Further work in these areas should be considered, as this information could support funding requests and integration across publicly funded NP programs. This work could also inform future planning related to budgetary allocations for upkeep, uptake, training, and technical support. Ongoing evaluation of student and faculty experiences, as well as student outcomes, should be monitored to inform strategic directions around implementation, quality assurance, and best practice initiatives in NP education.

While there has been limited documentation on implementing clinical learning technologies in NP education, substantial work from undergraduate nursing programs has highlighted their effective use. High fidelity simulation, VR, AR, and holographic imaging have been used at the undergraduate level to enhance students' confidence, knowledge acquisition, sense of belonging and inclusion, and overall learning experience (Brown et al., 2023; Chang et al., 2022; Hoyt & Hauze, 2023; Jallad, 2024; Liu & Butzlaff, 2021; Medvec et al., 2023; Russell, 2023; Wang et al., 2023). Promising outcomes in undergraduate education suggest a need to explore the potential benefits of integrating similar technologies within graduate NP programs. Universities and other learning centers will need to source, invest in, and support the technology infrastructure, as well as collaborate with the subject matter experts and faculty leaders required to develop

new content. Consultation with experts in clinical learning technology from undergraduate nursing implementation would leverage existing professional expertise across nursing faculties.

Multiple findings identified adaptation to change, or lack thereof, as a hindrance to technology integration in NP education (Lamarche et al., 2016; Rudolph et al., 2017). Educational institutions, program leaders, and faculty members are encouraged to assess their effectiveness with technology implementation and, given the rapidly evolving nature of clinical education technology, implement a continuous quality improvement approach to ensure ongoing relevance and suitable application within programs. Close attention must be paid to the optimal balance between use of novel technologies and the corresponding in-person patient experiences, so that maximal student outcomes may be realized (Graves et al., 2023).

Limitations

This literature review was subject to some limitations. First, the search was limited to peer-reviewed articles published in English, which may have excluded relevant insights from grey literature or from studies published in other languages. Second, the studies varied significantly in design, outcome measures, the timeline over which data were collected, and the type of technology that was assessed, which may have impacted the ability to synthesize findings or make direct comparisons. Finally, while every effort was made to prevent selection bias, it is important to note that the inclusion and interpretation of studies were influenced by the authors' judgment.

Further Considerations

This review project was undertaken to assess knowledge about the current use of advanced clinical learning technology in graduate-level NP programs. Despite some concept-supporting evidence, the current body of work around clinical education technologies in NP education signals room for expansion and future opportunities for implementation, evaluation, and development. Further work should focus on student, faculty, and funding-specific outcomes, including implementation experiences, change management, student and faculty satisfaction, learning outcomes, and budget.

The current literature has demonstrated that the enhanced use of clinical learning technologies such as VR, AR and holographic applications in NP education would improve learning, increase confidence, enhance patient safety, and contribute to better student and patient outcomes. Given the findings and significant benefits, NP program accrediting bodies, university administrators, and regulators who approve NP education programs should request reporting on the uses of these technologies in their evaluation frameworks.

Meanwhile, very little insight was provided into facilitators and barriers related to the uptake and implementation of relevant technologies among NP program faculty. This highlighted a gap in current knowledge necessary for more widespread application. NP program faculty will be a key linkage in successful implementation of technologies that support clinical learning. There is a need for further research, documentation, publication, and dissemination around learning technology pilots, implementations, challenges, and successes within NP education programs. Student feedback and outcomes related metrics should also be sought. As learning technologies advance, equipping NP programs

and faculty with guidance on their implementation will be essential for optimizing student learning and contributing to safe and effective patient care.

Conclusion

This review of the current field of work focused on the use of clinical learning technologies in NP education revealed key themes, including student uptake, student learning, change management, accessibility, and confidence. Given solid student uptake and the potential of these technologies to transcend geography in distributed learning models, they must be considered for timely integration into NP education programs. Further work should focus on process and outcome evaluation, cost estimation, planning and management, and public funding related to implementation, support, and updating of technologies. Meanwhile, those who fund NP education should consider ongoing investment in existing and emerging technologies that may be applied to enhance NP student learning and professional competence, ultimately improving patient care over the long term. Integration should be supported with adequate funding to cover technology, equipment, and faculty training and support, which is essential for operationalization in the post-secondary environment.

References

- Ainslie, M., & Bragdon, C. (2018). Telemedicine simulation in online family nurse practitioner education: Clinical competency and technology integration. *Journal of the American Association of Nurse Practitioners*, 30(8). <https://doi.org/10.1097/JXX.0000000000000071>
- Al-hatem, A.I., Masood, M., & Al-Samarraie, H. (2018). Fostering student nurses' self-regulated learning with the second life environment: An empirical study. *Journal of Information Technology Education: Research*, 17, 285–307. <https://doi.org/10.28945/4110>
- Anderson, M., Guido-Sanz, F., Díaz, D. A., Lok, B., Stuart, J., Akinola, I., & Welch, G. (2021). Augmented reality in nurse practitioner education: Using a triage scenario to pilot technology usability and effectiveness. *Clinical Simulation in Nursing*, 54, 105–112. <https://doi.org/10.1016/j.ecns.2021.01.006>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Brown, K. M., Perretta, J., Miller, G., Gilbert, G., Anderson, M., Horvath, C., Sullivan, N., Jeffers, J., Swoboda, S., Shilkofski, N., Green, D., & Mudd, S. (2023). Deconstructing healthcare silos: Multiplayer virtual reality for interprofessional & intraprofessional education. *Journal of Interactive Learning Research*, 34(3), 495–517. <https://doi.org/10.70725/193236wcyhb>
- Canadian Association of Schools of Nursing. (2022). *Accreditation nurse practitioner program standards and framework*. [CASN-Accreditation-NP-Standards-and-Framework_EN_FD.pdf](#)
- Carley, A. (2015). Using technology to enhance nurse practitioner student engagement. *The Nurse Practitioner*, 40(7), 47–54. <https://doi.org/10.1097/01.NPR.0000465119.04536.0e>
- Casler, K., Gawlik, K. S., & Messinger, J. (2024). Virtual reality to aid in competency-based online nurse practitioner curriculum (VR-NP). *Journal of Professional Nursing*, 55, 125–132. <https://doi.org/10.1016/j.profnurs.2024.09.009>
- Chang, C.-Y., Sung, H.-Y., Guo, J.-L., Chang, B.-Y., & Kuo, F.-R. (2022). Effects of spherical video-based virtual reality on nursing students' learning performance in childbirth education training. *Interactive Learning Environments*, 30(3), 400–416. <https://doi.org/10.1080/10494820.2019.1661854>
- Cronin, M., & George, E. (2020). The why and how of the integrative review. *Organizational Research Methods*, 26(1). <https://doi.org/10.1177/1094428120935507>
- D'Aiello, A., Cabitza, F., Natali, C., Vigano, S., Ferrero, P., Bognoni, L., Pasqualin, G., Giamberti, A., & Chessa, M. (2018). The effect of holographic heart models and mixed reality for anatomy learning in congenital heart disease: An exploratory study. *Journal of Medical Systems*, 47(64). <https://doi.org/10.1007/s10916-023-01959-8>

- Driscoll, M., & Burner, K. (2022). *Psychology of learning of instruction* (4th ed.). Pearson.
- Elliott, L., DeCristofaro, C., & Carpenter, A. (2012). Blending technology in teaching advanced health assessment in a family nurse practitioner program: Using personal digital assistants in a simulation laboratory. *Journal of the American Academy of Nurse Practitioners*, 24(9), 536–543. <https://doi.org/10.1111/j.1745-7599.2012.00728.x>
- Ficarra, B. (2021). Virtual reality, augmented reality and mixed reality. In W. Carroll (Ed.), *Emerging technologies for nurses: Implications for practice* (pp. 87–120). Springer.
- Graves, L., Dalgarno, N., Hoorn, R. V., Hastings-Truelove, A., Mulder, J., Kolomitro, K., Kirby, F., & van Wylick, R. (2023). Creating change: Kotter's change management model in action. *Canadian Medical Education Journal*, 14(3), 136–139. <https://doi.org/10.36834/cmej.76680>
- Hoyt, H., & Hauze, S. (2023). Nursing education through the lens of immersive simulation: A comparison of student knowledge, skill, and motivation to learn via two-dimensional video versus extended reality. *International Journal on E-Learning*, 22(2), 129–157. <https://doi.org/10.70725/463377yvompo>
- Jallad, S. T. (2024). The effectiveness of immersive virtual reality applications (human anatomy) on self-directed learning competencies among undergraduate nursing students: A cross-sectional study. *Anatomical Sciences Education*, 17(9), 1764–1775. <https://doi.org/10.1002/ase.2534>
- Kyaw, B., Saxena, N., Posadzki, P., Vseteckova, J., Nikolaou, C., George, P., Divakar, U., Masiello, I., Kononowicz, A., Zary, N., & Car, L. (2019). Virtual reality for health professions education: Systematic review and meta-analysis by the digital health education collaboration. *Journal of Medical Internet Research*, 21(1). <https://doi.org/10.2196/12959>
- Lamarche, K., Park, C., Fraser, S., Rich, M., & MacKenzie, S. (2016). In the palm of your hand: Normalizing the use of mobile technology for nurse practitioner education and clinical practice. *Nursing Leadership*, 29(3), 120–132. <https://doi.org/10.12927/cjnl.2016.24898>
- Lee, D., Bathish, M. A., & Nelson, J. (2024). Transforming nursing education: Developing augmented reality procedural training. *CyberPsychology, Behavior & Social Networking*, 27(6), 372–378. <https://doi.org/10.1089/cyber.2023.0403>
- Liu, Y., & Butzlaff, A. (2021). Where's the germs? The effects of using virtual reality on nursing students' hospital infection prevention during the COVID-19 pandemic. *Journal of Computer Assisted Learning*, 37(6), 1622–1628. <https://doi.org/10.1111/jcal.12601>
- May, S. C. (2019). *A comparative analysis of student success and perceptions of engagement between face-to-face and online college courses* (Publication No. 13811251) [Doctoral dissertation, Lindenwood University]. ProQuest Dissertations & Theses Global. <https://www.proquest.com/docview/2202321237?pq-origsite=primo&sourcetype=Dissertations%20&%20Theses>

- Marrocco, G., Ginzburg, E. & Feder, S. (2019). Seeing from the eyes of the beholder: A pilot program for learning through reflection in graduate nursing education. *Nursing Education Perspectives*, 40(2), 113–115. <https://doi.org/10.1097/01.NEP.0000000000000317>
- Medvec, B. R., Knight, S. W., & McLaughlin-Grayson, T. (2023). Cultivating leadership competency using virtual reality: DEI and anti-racism. *Journal of Interactive Learning Research*, 34(3), 475–494. <https://doi.org/10.70725/372356fpywxa>
- Nelson, B., Ridgway, J., Brown, A., Anthamatten, A., Hall, E., & McClure, N. (2025). Interactive simulation for addressing vaccine hesitancy in nurse practitioner education. *Clinical Simulation in Nursing*, 102. <https://doi.org/10.1016/j.ecns.2025.101719>
- Rogers, J. L., Reyes, A., & Yang, X. (2024). Development of an interactive 3D visualization tutorial for pathophysiology in graduate nursing education. *Nurse Educator*, 49(2), E68–E73. <https://doi.org/10.1097/NNE.0000000000001496>
- Rudolph, A., Vaughn, J., Crego, N., Hueckel, R., Kuszajewski, M., Molloy, M., Brisson, R., & Shaw, R. J. (2017). Integrating telepresence robots into nursing simulation. *Nurse Educator*, 42(2), E1–E4. <https://doi.org/10.1097/NNE.0000000000000329>
- Russell, S. L. (2023). *Evaluating a virtual reality simulation-based tool in undergraduate nursing students for impact on accuracy, clinical judgment, fear, and self-confidence when donning, doffing, and disposing of personal protective equipment* (Publication No. 30561029) [Graduate dissertation, West Virginia University]. ProQuest Dissertations & Theses Global. <https://doi.org/10.33915/etd.11960>
- Shelton, A., & Reimer, N. (2018). Telehealth wound applications: Barriers, solutions, and future use by nurse practitioners. *Online Journal of Nursing Informatics*, 22(2). <https://www.himss.org/resources/telehealth-wound-applications-barriers-solutions-and-future-use-nurse-practitioners/>
- Smith, T. S., Jordan, J., & Li, P. (2022). Video-based interactive clinical simulation: Preparing nurse practitioner students for clinical. *Journal for Nurse Practitioners*, 18(9), 995–998. <https://doi.org/10.1016/j.nurpra.2022.07.014>
- Stoll, C. R. T., Izadi, S., Fowler, S., Green, P., Suls, J., & Colditz, G. A. (2019). The value of a second reviewer for study selection in systematic reviews. *Research Synthesis Methods*, 10(4), 539–545. <https://doi.org/10.1002/jrsm.1369>
- Wang, C.-Y., Yin, T., Ma, K.-H., Shyu, J.-F., Cheng, C.-P., Wang, Y.-C., Huang, Y.-L., & Chiang, M.-H. (2023). Enhancing anatomy education through cooperative learning: Harnessing virtual reality for effective gross anatomy learning. *Journal of Microbiology & Biology Education*, 24(3). <https://doi.org/10.1128/jmbe.00100-23>

Whited, T. M., DeClerk, L., Berber, A., & Phelan, K. D. (2019). An innovative technique to promote understanding of anatomy for nurse practitioner students. *Journal of the American Association of Nurse Practitioners*, 33(5), 348–352. <https://doi.org/10.1097/JXX.0000000000000328>

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, 52(5), 546–553. <https://doi.org/10.1111/j.1365-2648.2005.03621.x>



August – 2026

A Systematic Review of Using Social and Epistemic Network Analysis Methods in Community of Inquiry Research

Liu Dong, Victoria L. Lowell, and Chi-Jia Hsieh

Department of Curriculum and Instruction, Purdue University

Abstract

With the increasing application of learning analytics in online learning communities, network analysis methods have emerged as powerful and efficient tools for deepening our understanding of collaborative learning processes, especially within the community of inquiry (CoI) framework, offering insights that extend beyond the examination of interaction patterns. This systematic review synthesizes findings from 43 empirical studies, with a focus on the application of social and epistemic network analysis methods (SNA and ENA) in CoI research in higher education contexts. Key findings highlight the value of SNA and ENA in investigating interaction patterns, refining CoI theory, examining community formation and evolution, analyzing discourse and knowledge construction, tracing the development of CoI presences, and identifying SNA measures relevant to CoI presences. This review also identifies the current challenges and opportunities in using SNA and ENA, as well as the potential for integrating both methods in future CoI research. By revealing both the strengths and limitations of existing studies, this review offers actionable implications for researchers, educators, and instructional designers aiming to enhance the design and analysis of online learning communities.

Keywords: social network analysis, epistemic network analysis, community of inquiry, systematic review

Introduction

The community of inquiry (CoI) framework, rooted in collaborative constructivist theory, was introduced to understand the concept of community in online discussions (Garrison et al., 2000). It comprises three core presences: cognitive, social, and teaching presence. Cognitive presence refers to the ability of learners to construct meaning through sustained reflection and discourse (Garrison et al., 2001), and it can be considered an important indicator of critical thinking. Social presence is defined as “the ability of participants in a community of inquiry to project themselves socially and emotionally, as ‘real’ people (i.e., their full personality), through the medium of communication being used” (Garrison et al., 2000, p. 94). Teaching presence, which encompasses the design, facilitation, and direction of cognitive and social processes, is essential in guiding learners through online educational experience (Anderson et al., 2001; Garrison & Arbaugh, 2007). Together, these presences foster deep and meaningful inquiry, transforming online communication into a medium for collaborative knowledge construction (Kozan & Richardson, 2014; Shea & Bidjerano, 2009). Traditional analyses of CoI in online and blended courses rely on surveys (Arbaugh et al., 2008) and quantitative content analysis (e.g., Satar & Akcan, 2018; Wicks et al., 2015). However, these methods are limited by self-reported data and their inability to capture dynamic interactions or the depth of critical thinking. To address these challenges, researchers increasingly adopt network analysis methods, particularly social network analysis (SNA) and epistemic network analysis (ENA), leveraging learning analytics to explore complex interaction patterns in online discourse.

SNA characterizes networked structures in terms of nodes (people, groups, or organizations) and edges (relationships or interactions) that connect them, producing visual representations and mathematical analysis (Wasserman & Faust, 1994), which can help us understand the network data and provide insights into the various roles and groupings in a network (Cela et al., 2015). SNA commonly employs several measures to describe the structure and characteristics of a network, such as centrality, cohesion, density, and centralization. Based on these measures, SNA focuses on the connections between individuals, identifying influential participants, interaction, social dynamics of knowledge construction and creation, and the community’s overall cohesion (Jan et al., 2019).

ENA is a novel analytical method designed to identify, quantify, and compare connections among elements in coded data by representing these relationships as dynamic networks that model patterns of association (Shaffer et al., 2016). The ENA algorithm comprises three fundamental components: units, conversations, and codes. Units refer to the entities whose network connections are being analyzed, such as individuals, concepts, or groups. Conversations are collections of data segments within which ENA models connections between concepts. Codes represent specific concepts whose patterns of association are being examined. The ENA algorithm uses a moving window to construct a network model for each line in the data, showing how codes in the current line are connected to those that occurred previously (Siebert-Evenstone et al., 2017). Furthermore, ENA is a kind of semantic network analysis which focuses more on the developmental trajectory of students’ online discourses, revealing the development of students’ cognitive thinking (Shaffer et al., 2016).

SNA and ENA both fall under the umbrella of network analysis methods, offering complementary insights in online learning: SNA captures the structural dynamics among learners, while ENA reveals how meaning is constructed through discourse. A literature review synthesizing the use of SNA and ENA in CoI studies is

crucial for understanding current methodological trends, identifying research gaps, and highlighting innovative approaches. More importantly, such a review informs future researchers by exploring the potential for integrating both methods into future CoI studies.

Existing Literature Review

Jan et al.'s (2019) review of nine studies examining CoI in higher education through SNA revealed two key findings: inconsistent results across studies, and a methodological tendency to combine SNA only with content analysis and statistical methods. Similarly, Elmoazen et al. (2022) noted the use of ENA in several CoI-related studies. However, their review lacked detailed analysis or discussion of these applications. Notably, neither review solely concentrated on CoI research, thereby limiting a deeper understanding of the methodological contributions and practical implications of using SNA and ENA within CoI research in distance learning contexts. Given this gap, a comprehensive and focused review is necessary to better understand how network analysis methods contribute to uncovering the relational and cognitive dynamics of online participation within the CoI. Such an understanding can provide researchers and practitioners with contextualized insights and actionable implications for designing more effective online learning environments and facilitating meaningful asynchronous online discussions (AODs). Therefore, this study was conceived to comprehensively review all existing research on the application of SNA and ENA in higher education within open and distance learning environments and examine the following research questions:

RQ1: What are publication trends, popular publication sources, and presences examined in studies that employ SNA and ENA methods?

RQ2: What are the key research purposes and findings in selected studies?

Methodology

A literature review is a respected scholarship method for understanding the scope and content of published literature. In this study, we followed the instructions and guidelines based on preferred reporting items for systematic reviews and meta-analyses (PRISMA; Page et al., 2021) to conduct this review.

Literature Search

To locate relevant articles, we searched the following major databases: Web of Science, Scopus, EBSCO, including Education Full Text (H. W. Wilson), Education Source, ERIC, Social Sciences Full Text (H. W. Wilson), and ScienceDirect. The target studies had to be published between January 2010 and December 2024.

Search terms were developed using a set of predefined keywords (Table 1) and applied using Boolean search operators. Each database was searched independently using combinations of these terms. All retrieved records were then imported into Covidence software (<https://www.covidence.org/>) to facilitate screening and management of the review process.

Table 1

Search Terms for Identifying Community of Inquiry Studies Using SNA and ENA

Term	Search Terms
Community of inquiry	social presence OR teaching presence OR cognitive presence OR learning presence OR communities of inquiry OR community of inquiry OR CoI OR practical inquiry model
Higher education	undergraduate OR graduate OR doctoral OR college* OR university
Network	network* OR interaction*

Eligibility Criteria

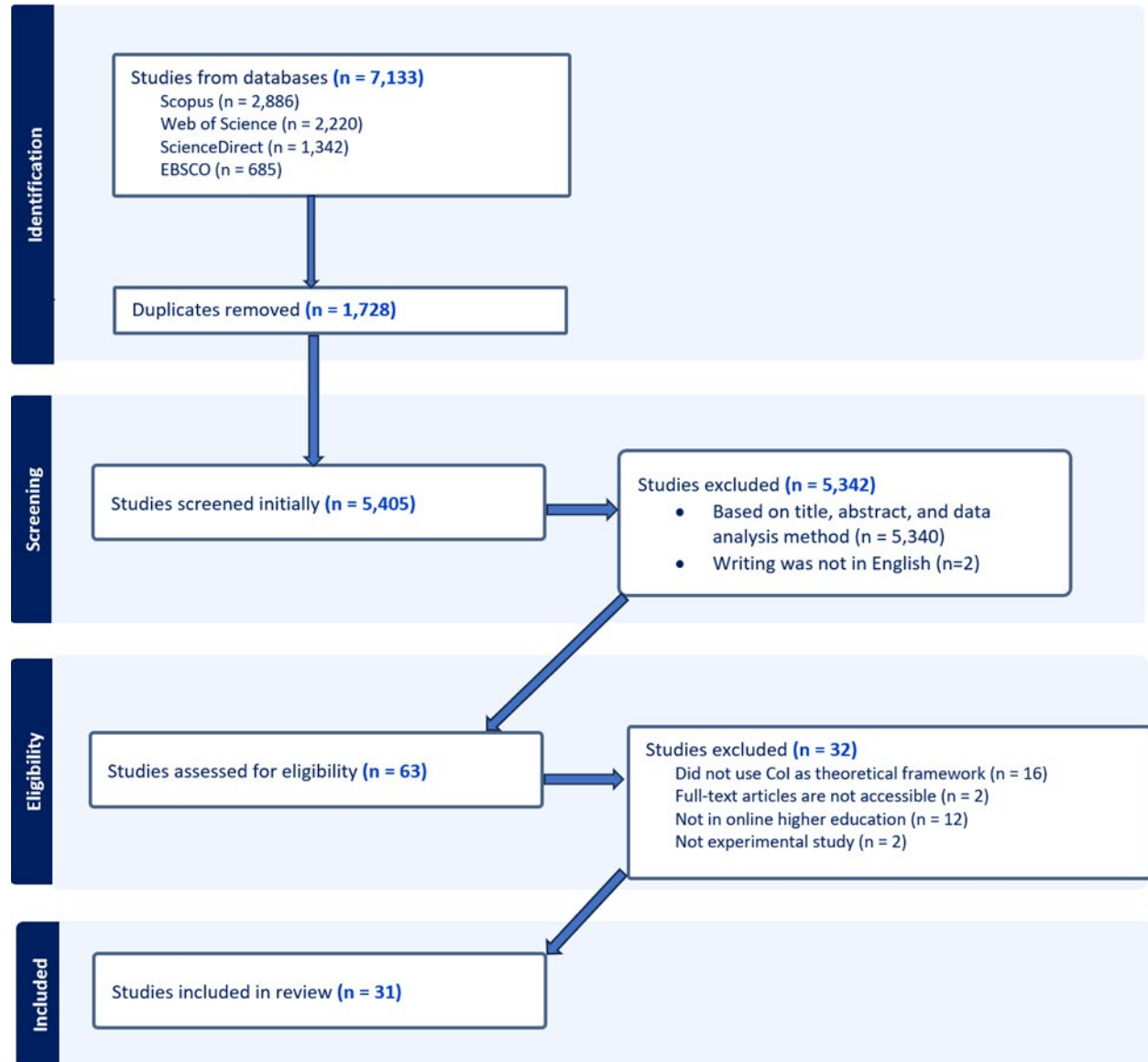
We adopted seven eligibility criteria to guide the screening process of relevant research: (a) studies should have CoI as the foundational framework; (b) studies should be written in English; (c) studies should have full access and be published in peer-reviewed journals or conference proceedings; (d) studies should be conducted in online higher education; (e) the data analyzed in these studies must include AODs, as network analysis methods were employed to capture interaction patterns derived from asynchronous discussion boards; (f) studies should employ SNA or ENA to analyze AODs; and (g) studies should be empirical studies.

Data Collection and Screening Process

For collecting the articles, we conducted a literature search in four stages: identification, screening, eligibility, and inclusion. We began identification by searching Scopus, Web of Science, ScienceDirect, and EBSCO. This search yielded a total of 7,133 studies. After automatically removing duplicates ($n = 1,728$) using Covidence, the remaining 5,405 articles were included in the next screening stage. To screen the literature, we primarily excluded articles based on their title, abstract, and data analysis method. To be noted, the primary reason for the large reduction is that our initial search strategy intentionally employed broad search terms across multiple databases to ensure comprehensive coverage of all potentially relevant studies. We also found that two articles were written in a language other than English. Finally, a total of 63 studies were eligible for full-text assessment. Based on the eligibility criteria, 31 studies were included in this review. The detailed data screening process is presented in Figure 1.

Figure 1

Screening Flowchart for Articles Included in This Review



Note. CoI = community of inquiry.

Additionally, given that the application of learning analytics methods has also been predominantly recognized and disseminated through the International Learning Analytics and Knowledge Conference (LAK) and the International Conference on Quantitative Ethnography (ICQE), the proceedings from these conferences were also included in our search to ensure a comprehensive review. Accordingly, we conducted additional searches on Google Scholar, LAK, and ICQE, and hand-searched relevant studies from the reference lists of the 31 studies identified in the screening. These additional procedures yielded 12 relevant conference papers. In total, 43 studies were included in this review.

Data Coding and Analysis

To code the selected articles, we carefully read, analyzed, and coded each article. They first collected the basic information (publication year, journals and conferences, presences included) from each study to answer RQ1. Then, the coders synthesized the research topics and findings of each study from the abstract, results, discussion, and limitation sections for answering RQ2. To ensure validity and reliability, the coders met each week and discussed the uncertain items until a 100% agreement was reached. The results are introduced and presented in the following section.

Results

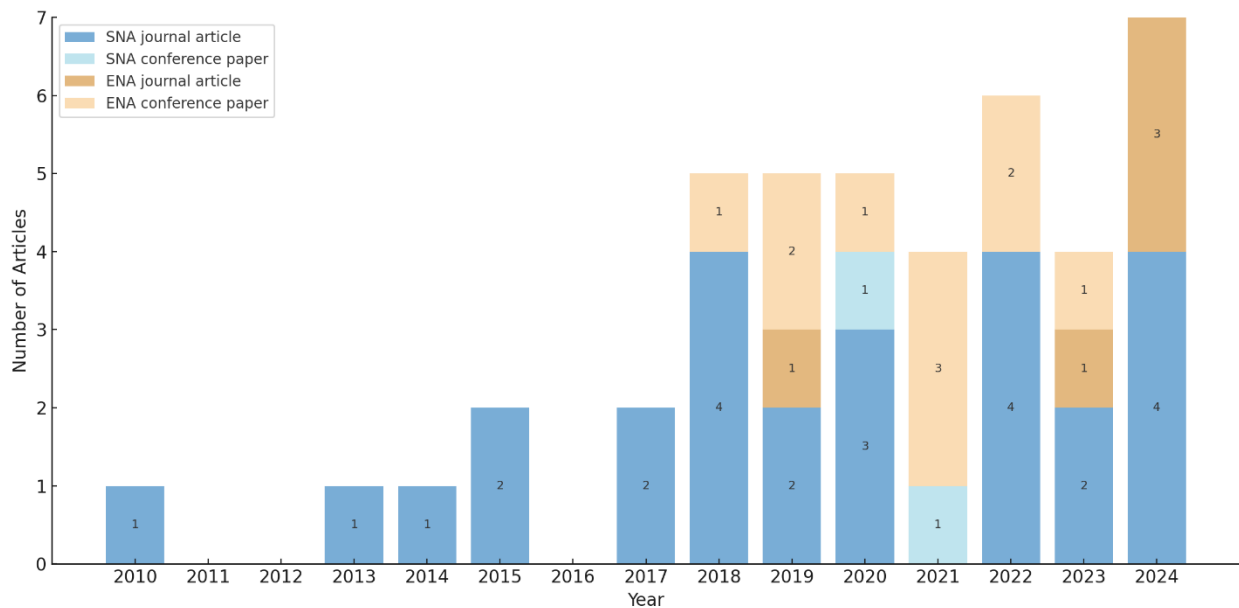
RQ1: What are publication trends, popular publication sources, and presences examined in studies that employ SNA and ENA methods?

Publication Trends

Figure 2 presents the distribution of using SNA and ENA methods in journal articles and conference proceedings. The first article using SNA to assess the CoI framework in higher education appeared in 2010, and the first ENA article was published in 2018. In addition, a noticeable increase in articles employing ENA methods was observed in 2024, suggesting a growing recognition and adoption of the ENA approach among researchers. This reveals that the research on using ENA in CoI research is still in its early stage of development. Overall, the combined number of ENA and SNA publications has gradually increased since 2013.

Figure 2

Number of Published Articles Using ENA or SNA Methods, 2010–2024



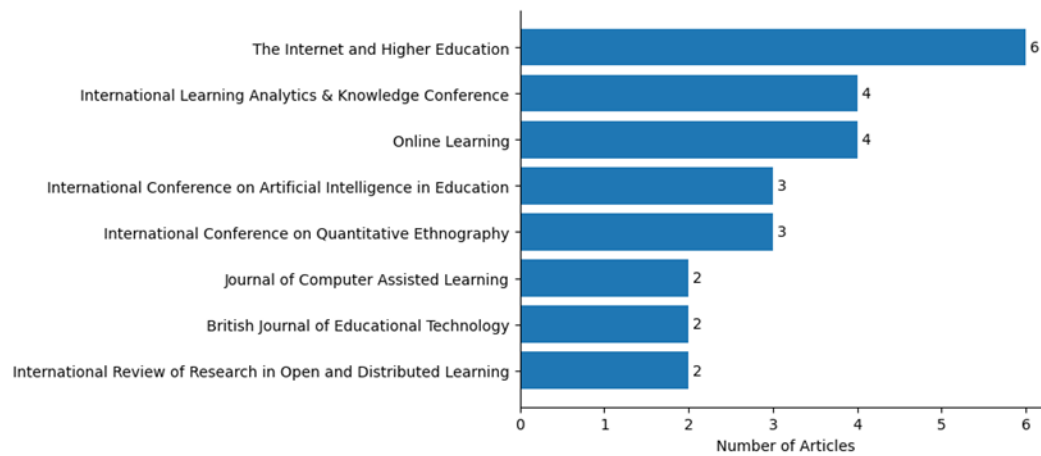
Note. ENA = epistemic network analysis; SNA = social network analysis.

Popular Publication Sources

This review identified 26 publication sources, including 5 conference proceedings and 21 journals, ultimately contributing to the 43 studies. Figure 3 presents the most popular publication sources.

Figure 3

Popular Publication Sources of Articles Identified for This Review



Presences Included

Table 2 provides an overview of the presences investigated using SNA and ENA within all selected studies. Social and cognitive presences were most frequently explored, while teaching presence received less attention. SNA has been more widely used, with studies covering all CoI presences. In contrast, ENA studies have primarily focused on social and cognitive presence, with fewer exploring teaching presence and none addressing learning presence independently. The combination of social and cognitive presence is the most commonly examined across both methods.

Table 2

Distribution of Community of Inquiry Presences and Analysis Methods in Reviewed Articles, 2010–2024

Presence	SNA	ENA
Social	Castellanos-Reyes et al. (2024); Lim (2023); Norz et al. (2023, 2024); Satar & Akcan (2018); Wu et al. (2022); Yen et al. (2022)	Ferreira et al. (2020); Moraes et al. (2021); Rolim, Ferreira Mello, Kovanović, & Gašević (2019); Scianna et al. (2022)
Cognitive	Alwafi (2022); Galikyan & Admiraal (2019); Jo et al. (2017); Lim (2024); Oh et al. (2018); Sadaf & Olesova (2020); Sadaf et al. (2024)	Ba et al. (2023, 2024); Farrow et al. (2021); Ferreira Mello et al. (2018); Iqbal et al. (2022)

Presence	SNA	ENA
Teaching	Wang & Liu (2020)	
Learning	Jimoyiannis & Tsiotakis (2017); Shea et al. (2013)	
Social & cognitive	Tirado-Morueta et al. (2015, 2020); Williams-Dobosz et al. (2021)	Fathali (2024); Ferreira et al. (2021, 2022); Ferreira Mello & Gašević (2019); Rolim, Ferreira Mello, Lins, & Gašević (2019)
Social & cognitive & teaching	Shea et al. (2010)	Elmoazen et al. (2024)
Social & cognitive & teaching & learning	Shea et al. (2014); Wicks et al. (2015)	
Not explicitly referring to any specific presence	Eryilmaz et al. (2022); Jan (2018); Jan & Vlachopoulos (2018, 2019); Papanikolaou et al. (2020)	

Note. SNA = social network analysis; ENA = epistemic network analysis.

RQ2: What are the key research purposes and findings in selected studies?

Participation and Interaction Patterns

Network analysis methods can uncover and map the levels of participation and the evolution of interaction patterns among online learners (Saqr & Alamro, 2019). Drawing on Moore's (1989) framework, the selected studies can be also categorized into: learner-learner interaction (e.g., Sadaf & Olesova et al., 2020; Sadaf et al., 2024), further divided into group and individual level (e.g., Jo et al., 2017; Satar & Akcan, 2018), learner-content interaction (e.g., Alwafi, 2022; Jo et al., 2017; Sadaf & Olesova et al., 2020), and learner-instructor interaction (e.g., Jan, 2018; Lim, 2024; Satar & Akcan et al., 2018; Wang & Liu, 2020). For instance, novice and experienced instructors facilitated discussion board interaction patterns (Lim, 2024), and findings showed that there were no significant differences in students' cognitive presence based on instructors' interaction patterns. Additional interaction types, such as learner-system interaction, have also been noted (Jo et al., 2017). These interactions play a pivotal role in shaping the dynamics and outcomes of AODs. Furthermore, interaction can also be explored through various intervention scaffoldings, including structured prompts, role assignments, and scenario-based tasks, which can foster reflective thinking, collaboration, and deeper cognitive engagement (e.g., Lim, 2024; Oh et al., 2018; Sadaf & Olesova, 2020). Moreover, discussion strategy with learning analytics made learners aware of their level and quality of interaction (Alwafi, 2022).

Refining CoI Theory

Network analysis methods provide valuable insights for refining the CoI framework. Beyond the traditional three presences, learning presence was introduced as a potential fourth construct, emphasizing learners' active engagement in self-regulated learning (Shea & Bidjerano, 2010). It highlights self-efficacy and effort regulation as key components of meaningful participation (Shea & Bidjerano, 2010; Shea et al., 2013).

Additionally, ENA has been widely employed to examine the interrelationships between social and cognitive presence, particularly across the phases of cognitive presence and indicators of social presence (Fathali, 2024; Ferreira et al., 2021, 2022; Ferreira Mello & Gašević, 2019; Rolim, Ferreira Mello, Lins, & Gašević, 2019). Main findings have shown that social presence indicators are more strongly associated with the exploration and integration phases of cognitive presence. Specifically, affective indicators of social presence connect more with advanced cognitive phases, such as integration and resolution, whereas interactive indicators are tied to early phases, such as triggering events and exploration (Fathali, 2024; Rolim, Ferreira Mello, Lins, & Gašević, 2019). Moreover, ENA-based studies provide additional insights by revealing comparisons in the connections between social and cognitive presences across different contexts, such as learners' emerging roles (Elmoazen et al., 2024; Ferreira et al., 2021), performance levels (Fathali, 2024), and experimental conditions, including scaffolding interventions (Rolim, Ferreira Mello, Lins, & Gašević, 2019).

Community Formation and Evolution

The terms network and community are closely related, as a network is a group of entities joined together to form a community (Jan & Vlachopoulos, 2018). The main aim of community detection is to identify different groups or clusters of nodes within the network that share some similar characteristics. Therefore, researchers aimed to use SNA as an analytics tool for exploring the community formation within CoI. For example, Papanikolaou et al. (2020) explored how SNA metrics was related to CoI presences and pointed out that gamification can facilitate CoI development. Two articles focused on the influence of various learning designs on types of community formation and evolution through SNA (Jan & Vlachopoulos, 2018, 2019). Additionally, researchers found that AOD topics and individual reflections influence community formation (Eryilmaz et al., 2022). Furthermore, students with higher social presence are more likely to contribute to community evolution (Yen et al., 2022).

Discourse and Knowledge Construction

Another key theme was exploring how students' collaborative knowledge is constructed through AODs within the CoI framework with the four presences: teaching presence (Shea et al., 2014; Wang & Liu, 2020), learning presence (Shea et al., 2013, 2014), cognitive presence (Shea et al., 2014; Tirado-Morueta et al., 2015) and social presence (Shea et al., 2014; Tirado-Morueta et al., 2015). These studies suggested that the social knowledge construction process can be predicted, measured, and analyzed by students' network positions (Shea et al., 2014; Tirado-Morueta et al., 2015). Notably, the findings emphasized that knowledge construction is closely related to social, cognitive, and learning presences (Tirado-Morueta et al., 2015), while teaching presence appears to have a relatively limited impact on this process (Shea et al., 2014). Thus, SNA emerges as a valuable method for gaining deeper insights into how knowledge is constructed collaboratively within online or blended learning environments.

The Trajectory of Single Presence or Multiple Presences

SNA has proven effective in capturing the evolution of social presence (Castellanos-Reyes et al., 2024; Wu et al., 2022) and mapping the weekly progression of interactions (Jan, 2018). That is, trajectory tracking can be effectively detected and visualized using SNA. For instance, interaction tends to evolve over time, with the middle of a course identified as the most productive and effective period for students to establish social presence and engage in deep-level learning (Wu et al., 2022).

ENA has proven effective in uncovering discourse patterns and developmental trajectories, with studies examining cognitive presence patterns (Ba et al., 2023, 2024; Farrow et al., 2021; Ferreira Mello et al., 2018; Iqbal et al., 2022), the evolution of social presence (Ferreira et al., 2020; Moraes et al., 2021; Rolim, Ferreira Mello, Kovanović, & Gašević, 2019; Scianna et al., 2022), and temporal changes in the relationship between social and cognitive presences (Fathali, 2024; Rolim, Ferreira Mello, Lins, & Gašević, 2019). These findings highlight the temporal progression of cognitive engagement and the dynamic interconnections among cognitive presence indicators. Trajectory analyses also compared variations across roles in AODs (Elmoazen et al., 2024), learning contexts (remote vs. hybrid; Scianna et al., 2022), coding approaches (Ferreira et al., 2020), and instructional scaffolds (Ferreira Mello et al., 2018; Iqbal et al., 2022).

Identifying SNA Measures Relevant to CoI Presences

The efficacy of SNA is to understand the dynamics of online learning within a CoI framework by comparing patterns produced by various SNA measures with those developed using content analyses. Table 3 summarizes the distribution of SNA measures in selected CoI studies.

Table 3

The Distribution of SNA Measures in Community of Inquiry Studies, 2010–2024

SNA measure	Definition	Studies
Node level		
Degree centrality	Number of ties to other nodes in the network.	Castellanos-Reyes et al. (2024); Eryilmaz et al. (2022); Galikyan & Admiraal (2019); Jan (2018); Jimoyiannis & Tsiotakis (2017); Jo et al. (2017); Lim (2023, 2024); Norz et al. (2023, 2024); Oh et al. (2018); Papanikolaou et al. (2020); Sadaf & Olesova (2020); Sadaf et al. (2024); Satar & Akcan (2018); Shea et al. (2013, 2014); Tirado-Morueta et al. (2015, 2020); Wang & Liu (2020); Wicks et al. (2015); Williams-Dobosz et al. (2021); Yen et al. (2022)
Closeness centrality	Average shortest path from a given node to all other nodes in the network.	Eryilmaz et al. (2022); Norz et al. (2023, 2024); Papanikolaou et al. (2020); Sadaf et al. (2024); Yen et al. (2022)

SNA measure	Definition	Studies
Betweenness centrality	Number of shortest paths that have to pass to reach a given node.	Eryilmaz et al. (2022); Jo et al. (2017); Norz et al. (2023, 2024); Sadaf et al. (2024); Yen et al. (2022)
Eigenvector centrality	Quality and level of how individuals strategically connect to other more active or well-connected members.	Sadaf et al. (2024); Williams-Dobosz et al. (2021); Yen et al. (2022)
Dyad level		
Reciprocity	Extent to which two individuals in a network reciprocate each other's connections.	Castellanos-Reyes et al. (2024); Jan (2018); Jan & Vlachopoulos (2018, 2019); Lim (2023, 2024); Norz et al. (2023); Sadaf et al. (2024)
Subgroup and overall network level		
Clique	Small subgroups made of nodes directly connected to one another but loosely connected to the network core.	Jan & Vlachopoulos (2018, 2019); Jimoyiannis & Tsiotakis (2017)
Density	Number of connections formed within a network.	Alwafi (2022); Castellanos-Reyes et al. (2024); Jan (2018); Jo et al. (2017); Lim (2023, 2024); Norz et al. (2023); Oh et al. (2018); Sadaf et al. (2024); Satar & Akcan (2018); Tirado-Morueta et al. (2015, 2020); Wang & Liu (2020); Yen et al. (2022)
Transitivity	Density of loops of the length of three (triangles) in a network.	Jan (2018); Jan & Vlachopoulos (2018, 2019)
Centralization	How tightly interaction within a network is organized around its most focal points.	Jan (2018); Jan & Vlachopoulos (2018, 2019); Jo et al. (2017); Lim (2023, 2024); Sadaf et al. (2024); Satar & Akcan (2018); Tirado-Morueta et al. (2015, 2020)

SNA measure	Definition	Studies
Cohesion	Degree of linkage among participants in the network.	Jimoyiannis & Tsiotakis (2017); Tirado-Morueta et al. (2015, 2020)

Note. Definitions adapted from *Social Network Analysis: Methods and Applications* by S. Wasserman and K. Faust, 1994, Cambridge University Press. CoI = community of inquiry.

In summary, studies using SNA focus on three levels of analysis: node, dyad, and subgroup and overall network. At the node level, degree centrality measures were frequently used, followed by closeness centrality, betweenness centrality, and eigenvector centrality. Degree centrality, including indegree and outdegree centrality, serves as a key indicator of participants' interaction levels and their positional roles within AODs (e.g., Lim, 2023, 2024; Sadaf & Olesova, 2020). High outdegree centrality reflects active participation, while high indegree indicates prestige (e.g., Lim, 2023, 2024). Research shows that students with higher social, cognitive, and learning presence are often central within interaction networks, correlating positively with centrality measures (Shea et al., 2013, 2014; Wicks et al., 2015). For instance, high levels of social presence were observed in contributions of learners who were central to the network (Satar & Akcan, 2018; Shea et al., 2013, 2014). However, SNA measures present contradicting results for teaching presence. Some studies indicate that students with high levels of teaching presence are not always central in the network, suggesting that teaching presence may influence learning differently, possibly through guidance and support rather than direct interaction (Shea et al., 2014). However, an alternative perspective suggests that instructors are key players in the group as central actors (Satar & Akcan, 2018).

At the dyad level, reciprocity measures mutual connections. Novice instructors tend to have higher reciprocity compared to experienced instructors (Lim, 2024), and higher network density is associated with increased reciprocity (Lim, 2023). At the subgroup and overall network level, density and centralization are frequently examined. Density reflects overall connectedness, with higher scores indicating cohesive groups and increased centralization (Lim, 2024; Satar & Akcan, 2018). High centralization suggests interactions revolve around a focal participant (Jo et al., 2017). Research also found that cohesion and density can positively influence social and cognitive presence, with density exerting a stronger effect on social presence (Tirado-Morueta et al., 2020). Overall, these findings highlight the interplay between SNA measures and the CoI framework, demonstrating how network structure influences CoI presences in AODs.

Discussion

The systematic literature review reported in this paper examined 43 empirical studies to deepen our understanding of how SNA and ENA have been applied in CoI research. The results highlight the relevance of CoI research while illustrating how SNA and ENA function as powerful conceptual and methodological resources for capturing the complexities of online learning processes. However, the majority of the selected studies employed relatively small sample sizes and were conducted over short durations, often lacking a longitudinal perspective (e.g., Galikyan & Admiraal, 2019; Lim, 2024). As a result, the generalizability of their findings is limited, as results are context-bound and specific to the participant groups involved. To enhance the validity and applicability of these findings, future research should aim to replicate studies across diverse educational settings and populations. In addition, some studies only used 2-week or 3-week

datasets (e.g., Eryilmaz et al., 2022; Jan & Vlachopoulos, 2018). Therefore, there is also a need for longitudinal studies.

From the theoretical perspective, SNA can be a useful method to validate learning presence as a new construct for the CoI framework (Shea et al., 2013, 2014, Wicks et al., 2015), while ENA effectively examines the interplay between social and cognitive presences across various contexts (e.g., Fathali, 2024; Rolim, Ferreira Mello, Lins, & Gašević, 2019). The resulting insights also invite a refinement of the CoI framework by revealing how cognitive presence develops within specific interaction structures and how social presence influences collective knowledge integration. However, both methods have given limited attention to teaching and learning presences, with ENA yet to explore learning presence. Such findings highlight the potential for reconceptualizing the CoI framework to account for the evolving interplay among its presences and further explore teaching and learning presence in future studies.

From educators' and instructional designers' perspectives, integrating SNA and ENA in analyzing AODs provides valuable qualitative and quantitative insights into students' overall interaction patterns, cognitive engagement and development, and progression of interactions. Such insights enable educators to provide timely feedback and refine course design (e.g., Ferreira et al., 2022). For instance, placing greater emphasis on interactive elements and supporting cognitive development, especially during the exploration and integration phases, can improve learning outcomes (Fathali, 2024). Recognizing these temporal patterns allows instructional designers to align course activities and interventions with high-engagement periods, enhancing the learning experience and self-regulated skills (Galikyan & Admiraal, 2019). There is also a need for instructors to play an active role in facilitating AODs and improving the depth and quality of participation (Wang & Liu, 2020). Therefore, more studies should be conducted to emphasize the importance of learner-to-instructor interaction (Fathali, 2024; Lim, 2024; Satar & Akcan, 2018; Shea et al., 2013; Wang & Liu, 2020). Additionally, network visualizations can demonstrate the impact of discussion prompts, role assignments, and scaffolds: structured prompts and reflective roles lead to more cohesive and interconnected networks, promoting deeper learning (e.g., Lim, 2024; Oh et al., 2018). Researchers could further examine how different instructional interventions shape students' interaction patterns and cognitive engagement. Such inquiry would enable instructors to identify when and how to provide timely, targeted support to learners in online environments.

Methodologically, both ENA and SNA provide visual representations to capture both individual and group-level dynamics that make complex interactions more interpretable. However, most studies using SNA focus on basic measures, such as centrality, cohesion, and centralization. More measures such as multiplexity, clique, and diameter should also be explored and validated to assess their potential applicability. Studies using ENA have demonstrated its validity to capture the complex interplay between social and cognitive presences or solely explore social presence or cognitive presence by comparing different groups. As a novel approach, ENA has the potential to explore more aspects of CoI research. Furthermore, current CoI literature is characterized by a methodological imbalance, as researchers typically apply SNA or ENA independently rather than in combination. It highlights a methodological gap and presents a promising avenue for future inquiry. Specifically, combining SNA and ENA can provide a more holistic understanding of online learning by capturing both the structural patterns of interaction and the epistemic dimensions of discourse.

Implications

The application of SNA and ENA in CoI research addresses many aspects related to the online learning experience. Based on prior results, we point out three research challenges and opportunities essential for future research:

Using SNA in CoI Research: Challenges and Opportunities

SNA has proven to be an effective technique for analyzing CoI research, which enables the mapping and identification of CoI structures, revealing patterns of interaction, group cohesion, and the impact of instructional design on community formation (Jan & Vlachopoulos, 2018, 2019). However, SNA as a quantitative approach may be challenging to enhance the rigor of research in this area. More qualitative data, such as using student and instructor interviews or other instruments, is essential for future researchers to have a more comprehensive and meaningful interpretation of multiple factors and their effects on knowledge building and interaction patterns (Galikyan & Admiraal, 2019; Jo et al., 2017; Lim, 2024). In addition, the selected studies using SNA in CoI research also revealed that the majority of studies focused on the node-level and overall network level of SNA measures. More attention should be given to the dyad level, as examining the relationships between pairs of individuals provides valuable insights into how individual and small-group interactions contribute to the broader learning experience.

Using ENA in CoI Research: Challenges and Opportunities

Although ENA enables researchers to move beyond simple frequency counts, capturing the temporal and structural connections between CoI indicators, such as how social and cognitive presences interact over time (Dong et al., 2025), several challenges remain in current research. For instance, effective use of ENA requires a clear theoretical rationale linking code co-occurrences to meaningful constructs; otherwise, interpretations may lack validity (Elmoazen et al., 2022). Therefore, it is essential to ensure the validity of the coding schemes and critically examine whether these themes accurately capture the core dynamics of CoI and the evolving nature of online discourse (Garrison et al., 2006; Garrison & Arbaugh, 2007). Additionally, most ENA studies rely on manual coding of small datasets, limiting scalability and generalizability. The use of automated coding approaches and larger corpora remains limited, although recent advances in machine learning have demonstrated the feasibility of analyzing extensive textual datasets and automating coding procedures (Ba et al., 2023, 2024; Elmoazen et al., 2024). However, the number of studies in CoI research remains insufficient. Most importantly, the interpretive complexity of ENA visualizations presents an additional challenge, particularly for practitioners who lack prior experience with network analytic techniques or who do not have a well-established theoretical framework to guide interpretation. Those challenges point out the need for methodological refinement and broader adoption.

The Potential of Combining SNA and ENA in CoI Research

A notable methodological limitation within the current body of CoI research is the predominant use of either SNA or ENA in isolation. The limited integration of these analytical approaches highlights a methodological gap and presents a promising avenue for future inquiry. SNA identifies the roles and positions of individual or group actors within a network but lacks the capacity to analyze the content of discussion posts. Conversely, ENA effectively examines collaborative discourse and the connections among key features of collaborative learning. However, it is limited in its ability to analyze the roles actors assume

in collaboration and the social structures that emerge in the learning process (Gašević et al., 2019). The growing use of both methods in CoI-related studies indicates an emerging consensus that single-method approaches may be insufficient to unpack the dynamic nature of online collaborative learning experiences (Swiecki & Shaffer, 2020). Therefore, the integration of both SNA and ENA addresses a critical methodological need by enabling a more holistic understanding of online collaborative learning that captures both the structural properties of learner interaction and the epistemic patterns embedded in discourse (Kaliisa et al., 2022). Moreover, this dual perspective can support the simultaneous examination of social and cognitive connections, offering clearer evidence for how interaction networks shape meaning-making processes (Fan & Dai, 2021; Gašević et al., 2019). Additionally, this combination can inform concrete design decisions, such as identifying collaborative process, optimizing subgroup connectivity, and pinpointing which discourse pathways require instructional scaffolding (Liu et al., 2021). By jointly modeling interaction and discourse, this approach may serve as a more reliable predictor of learning outcomes than either method alone (Swiecki & Shaffer, 2020).

Conclusions and Limitations

This review synthesizes 43 empirical studies, providing comprehensive insights into the use of SNA and ENA for analyzing AODs within the CoI framework in higher education. Overall, by synthesizing the literature and identifying key findings, this study enhances our understanding of the current use of SNA and ENA in CoI studies and offers valuable guidance for researchers and practitioners. However, several limitations should be noted. First, we only included empirical studies published in journals or conference proceedings, potentially excluding some influential reviews or conceptual papers. Future studies should include those kinds of works to gain a broader perspective on the application of SNA and ENA in CoI studies. Second, this review is restricted to the context of online higher education, so it does not explore the practical implications for other educational settings, such as secondary education or MOOCs. Future research should adopt broader inclusion criteria and diverse educational settings.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors used ChatGPT 4o to revise the grammar and proofreading. After using this tool, the authors edited the content as needed and take full responsibility for the publication.

References

References marked with an asterisk indicate studies included in the review.

- *Alwafi, E. M. (2022). Designing an online discussion strategy with learning analytics feedback on the level of cognitive presence and student interaction in an online learning community. *Online Learning*, 26(1), 80–92. <https://doi.org/10.24059/olj.v26i1.3065>
- Anderson, T., Rourke, L., Garrison, D. R., & Archer, W. (2001). Assessing teaching presence in a computer conferencing context. *Journal of Asynchronous Learning Networks*, 5(2), 1–17. <https://doi.org/10.24059/olj.v5i2.1875>
- Arbaugh, J. B., Cleveland-Innes, M., Diaz, S. R., Garrison, D. R., Ice, P., Richardson, J. C., & Swan, K. P. (2008). Developing a community of inquiry instrument: Testing a measure of the community of inquiry framework using a multi-institutional sample. *The Internet and Higher Education*, 11(3–4), 133–136. <https://doi.org/10.1016/j.iheduc.2008.06.003>
- *Ba, S., Hu, X., Stein, D., & Liu, Q. (2023). Assessing cognitive presence in online inquiry-based discussion through text classification and epistemic network analysis. *British Journal of Educational Technology*, 54(1), 247–266. <https://doi.org/10.1111/bjet.13285>
- *Ba, S., Hu, X., Stein, D., & Liu, Q. (2024). Anatomizing online collaborative inquiry using directional epistemic network analysis and trajectory tracking. *British Journal of Educational Technology*, 55(5), 2173–2191. <https://doi.org/10.1111/bjet.13441>
- *Castellanos-Reyes, D., Richardson, J. C., & Maeda, Y. (2024). The evolution of social presence: A longitudinal exploration of the effect of online students' peer-interactions using social network analysis. *The Internet and Higher Education*, 61, Article 100939. <https://doi.org/10.1016/j.iheduc.2024.100939>
- Cela, K. L., Sicilia, M. Á., & Sánchez, S. (2015). Social network analysis in e-learning environments: A preliminary systematic review. *Educational Psychology Review*, 27, 219–246. <https://doi.org/10.1007/s10648-014-9276-0>
- Dong, L., Hsieh, C.-J., Watson, S. L., Yu, J. H., & Lowell, V. L. (2025). Revisiting the interrelationships between social and cognitive presences in asynchronous online courses: Insights from epistemic network analysis. *Online Learning*, 29(3), 417–444. <https://doi.org/10.24059/olj.v29i3.4672>
- *Elmoazen, R., Saqr, M., Hirsto, L., & Tedre, M. (2024). Capturing temporal pathways of collaborative roles: A multilayered analytical approach using community of inquiry. *International Journal of Computer-Supported Collaborative Learning*, 20, 41–77. <https://doi.org/10.1007/s11412-024-09431-6>

- Elmoazen, R., Saqr, M., Tedre, M., & Hirsto, L. (2022). A systematic literature review of empirical research on epistemic network analysis in education. *IEEE Access*, 10, 17330–17348. <https://doi.org/10.1109/ACCESS.2022.3149812>
- *Eryilmaz, E., Thoms, B., & Ahmed, Z. (2022). Cluster analysis in online learning communities: A text mining approach. *Communications of the Association for Information Systems*, 51(1), Article 32. <https://doi.org/10.17705/1CAIS.05132>
- Fan, O., & Dai, X. (2021). Using a three-layered social-cognitive network analysis framework for understanding online collaborative discussions. *Australasian Journal of Educational Technology*, 38(1), 164–181. <https://doi.org/10.14742/ajet.7166>
- *Farrow, E., Moore, J., & Gašević, D. (2021). Ordering effects in a role-based scaffolding intervention for asynchronous online discussions. In I. Roll, D. McNamara, S. Sosnovsky, R. Luckin, & V. Dimitrova (Eds.), *International Conference on Artificial Intelligence in Education. AIED 2021* (pp. 125–136). Springer International Publishing. https://doi.org/10.1007/978-3-030-78292-4_11
- *Fathali, S. (2024). Social and cognitive presence in a community of inquiry: An epistemic network analysis of CALL students' interactions via NowComment. *System*, 121, Article 103233. <https://doi.org/10.1016/j.system.2024.103233>
- *Ferreira, M. A. D., Ferreira Mello, R., Kovanović, V., Nascimento, A., Lins, R. D., & Gašević, D. (2022). NASC: Network analytics to uncover socio-cognitive discourse of student roles. In A. F. Wise, R. Martinez-Maldonado, & I. Hilliger (Chairs), *LAK22: The 12th International Conference on Learning Analytics and Knowledge* (pp. 415–425). ACM. <https://doi.org/10.1145/3506860.3506978>
- *Ferreira, M. A. D., Ferreira Mello, R., Lins, R. D., & Gašević, D. (2021). Analytics of emerging and scripted roles in online discussions: An epistemic network analysis approach. In I. Roll, D. McNamara, S. Sosnovsky, R. Luckin, & V. Dimitrova (Eds.), *International Conference on Artificial Intelligence in Education. AIED 2021* (pp. 156–161). Springer. https://doi.org/10.1007/978-3-030-78270-2_28
- *Ferreira, M. A. D., Rolim, V., Ferreira Mello, R., Lins, R. D., Chen, G., & Gašević, D. (2020). Towards automatic content analysis of social presence in transcripts of online discussions. In M. Scheffel, V. Kovanović, N. Pinkwart, & K. Verbert (Chairs), *LAK20: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 141–150). ACM. <https://doi.org/10.1145/3375462.3375495>
- *Ferreira Mello, R., & Gašević, D. (2019). What is the effect of a dominant code in an epistemic network analysis? In B. Eagan, M. Misfeldt, & A. Siebert-Evenstone (Eds.), *Advances in Quantitative Ethnography: ICQE 2019* (pp. 66–76). Springer. https://doi.org/10.1007/978-3-030-33232-7_6

- *Ferreira Mello, R., Kovanović, V., Gašević, D., & Rolim, V. (2018). Towards combined network and text analytics of student discourse in online discussions. In C. Rosé, R. Martínez-Maldonado, U. Hoppe, R. Luckin, M. Mavrikis, K. Porayska-Pomsta, B. McLaren, & B. du Boulay (Eds.), *International Conference on Artificial Intelligence in Education. AIED 2018* (pp. 111–126). Springer. https://doi.org/10.1007/978-3-319-93843-1_9
- *Galikyan, I., & Admiraal, W. (2019). Students' engagement in asynchronous online discussion: The relationship between cognitive presence, learner prominence, and academic performance. *The Internet and Higher Education*, 43, Article 100692. <https://doi.org/10.1016/j.iheduc.2019.100692>
- 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](https://doi.org/10.1016/S1096-7516(00)00016-6)
- Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education*, 15(1), 7–23. <https://doi.org/10.1080/08923640109527071>
- Garrison, D. R., & Arbaugh, J. B. (2007). Researching the community of inquiry framework: Review, issues, and future directions. *The Internet and Higher Education*, 10(3), 157–172. <https://doi.org/10.1016/j.iheduc.2007.04.001>
- Garrison, D. R., Cleveland-Innes, M., Koole, M., & Kappelman, J. (2006). Revisiting methodological issues in transcript analysis: Negotiated coding and reliability. *The Internet and Higher Education*, 9(1), 1–8. <https://doi.org/10.1016/j.iheduc.2005.11.001>
- Gašević, D., Joksimović, S., Eagan, B. R., & Shaffer, D. W. (2019). SENS: Network analytics to combine social and cognitive perspectives of collaborative learning. *Computers in Human Behavior*, 92, 562–577. <https://doi.org/10.1016/j.chb.2018.07.003>
- *Iqbal, S., Swiecki, Z., Joksimović, S., Ferreira Mello, R., Aljohani, N., Ul Hassan, S., & Gašević, D. (2022). Uncovering associations between cognitive presence and speech acts: A network-based approach. In A. F. Wise, R. Martínez-Maldonado, & I. Hilliger (Chairs), *LAK22: 12th International Learning Analytics & Knowledge Conference* (pp. 315–325). ACM. <https://doi.org/10.1145/3506860.3506908>
- *Jan, S. K. (2018). Identifying online communities of inquiry in higher education using social network analysis. *Research in Learning Technology*, 26, Article 2064. <https://doi.org/10.25304/rlt.v26.2064>
- *Jan, S. K., & Vlachopoulos, P. (2018). Influence of learning design on the formation of online communities of learning. *The International Review of Research in Open and Distributed Learning*, 19(4). <https://doi.org/10.19173/irrodl.v19i4.3620>

- *Jan, S. K., & Vlachopoulos, P. (2019). Social network analysis: A framework for identifying communities in higher education online learning. *Technology, Knowledge and Learning*, 24(4), 621–639. <https://doi.org/10.1007/s10758-018-9375-y>
- Jan, S., Vlachopoulos, P., & Parsell, M. (2019). Social network analysis and online learning communities in higher education: A systematic literature review. *Online Learning*, 23(1). <https://doi.org/10.24059/OLJ.V23I1.1398>
- *Jimoyiannis, A., & Tsiotakis, P. (2017). Beyond students' perceptions: Investigating learning presence in an educational blogging community. *Journal of Applied Research in Higher Education*, 9(1), 129–146. <https://doi.org/10.1108/JARHE-06-2015-0046>
- *Jo, I., Park, Y., & Lee, H. (2017). Three interaction patterns on asynchronous online discussion behaviours: A methodological comparison. *Journal of Computer Assisted Learning*, 33(2), 106–122. <https://doi.org/10.1111/jcal.12168>
- Kaliisa, R., Rienties, B., Mørch, A. I., & Kluge, A. (2022). Social learning analytics in computer-supported collaborative learning environments: A systematic review of empirical studies. *Computers and Education Open*, 3, Article 100073. <https://doi.org/10.1016/j.caeo.2022.100073>
- Kozan, K., & Richardson, J. C. (2014). Interrelationships between and among social, teaching, and cognitive presence. *The Internet and Higher Education*, 21, 68–73. <https://doi.org/10.1016/j.iheduc.2013.10.007>
- *Lim, J. (2023). Exploring the relationships between interaction measures and learning outcomes through social network analysis: The mediating role of social presence. *International Journal of Educational Technology in Higher Education*, 20(1), Article 14. <https://doi.org/10.1186/s41239-023-00384-8>
- *Lim, J. (2024). Students' interaction, satisfaction and cognitive presence in online discussions. *Educational Technology & Society*, 27(1), 35–49. <https://www.jstor.org/stable/48754840>
- Liu, Z., Zhang, N., Peng, X., Liu, S., Yang, Z., Peng, J., Su, Z., & Chen, J. (2021). Exploring the relationship between social interaction, cognitive processing and learning achievements in a MOOC discussion forum. *Journal of Educational Computing Research*, 60(1), 132–169. <https://doi.org/10.1177/07356331211027300>
- *Moraes, M., Folkestad, J., & McKenna, K. (2021). Using conversation analysis and epistemic network analysis to understand social presence in an asynchronous online discussion. In A. R. Ruis & S. B. Lee (Eds.), *Second International Conference on Quantitative Ethnography: Conference proceedings supplement* (p. 15–18). International Society for Quantitative Ethnography. https://www.qesoc.org/images/pdf/ICQE20_Proceedings_Supplement_Final_web.pdf
- Moore, M. G. (1989). Editorial: Three types of interaction. *American Journal of Distance Education*, 3(2), 1–7. <https://doi.org/10.1080/08923648909526659>

- *Norz, L.-M., Dornauer, V., Hackl, W. O., & Ammenwerth, E. (2023). Measuring social presence in online-based learning: An exploratory path analysis using log data and social network analysis. *The Internet and Higher Education*, 56, Article 100894. <https://doi.org/10.1016/j.iheduc.2022.100894>
- *Norz, L.-M., Kaczko, É., Netzer, M., & Ammenwerth, E. (2024). Quantifying social presence in online-based learning: A statistical and pedagogical analysis of indicators from social network analysis. *Online Learning*, 28(4), 414–467. <https://doi.org/10.24059/olj.v28i4.4134>
- *Oh, E. G., Huang, W.-H. D., Hedayati Mehdiabadi, A., & Ju, B. (2018). Facilitating critical thinking in asynchronous online discussion: Comparison between peer-and instructor-redirection. *Journal of Computing in Higher Education*, 30, 489–509. <https://doi.org/10.1007/s12528-018-9180-6>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372(71). <https://doi.org/10.1136/bmj.n71>
- *Papanikolaou, K., Tzelepi, M., Moundridou, M., & Petroulis, I. (2020). Employing social network analysis to enhance community learning. In V. Kumar & C. Troussas (Eds.), *Intelligent tutoring systems. 16th international conference: ITS 2020* (pp. 342–352). Springer. https://doi.org/10.1007/978-3-030-49663-0_41
- *Rolim, V., Ferreira Mello, R., Lins, R. D., & Gašević, D. (2019). A network-based analytic approach to uncovering the relationship between social and cognitive presences in communities of inquiry. *The Internet and Higher Education*, 42, 53–65. <https://doi.org/10.1016/j.iheduc.2019.05.001>
- *Rolim, V., Ferreira Mello, R., Kovanović, V., & Gašević, D. (2019). Analysing social presence in online discussions through network and text analytics. In *IEEE 19th International Conference on Advanced Learning Technologies. ICALT 2019* (pp. 163–167). IEEE. <https://doi.org/10.1109/ICALT.2019.00058>
- *Sadaf, A., & Olesova, L. (2020). Exploring the relationship between interaction and the structure of questions in online discussions using learning analytics. *European Journal of Open, Distance and E-Learning*, 23(1), 46–60. <https://doi.org/10.2478/eurodl-2020-0004>
- *Sadaf, A., Olesova, L., & Choi, H. (2024). Empowering learning networks: Insights from social network analysis in inquiry-based discussions. *Online Learning*, 28(4), 200–227. <https://doi.org/10.24059/olj.v28i4.4635>
- Saqr, M., & Alamro, A. (2019). The role of social network analysis as a learning analytics tool in online problem based learning. *BMC Medical Education*, 19(1), Article 160. <https://doi.org/10.1186/s12909-019-1599-6>

- *Satar, H. M., & Akcan, S. (2018). Pre-service EFL teachers' online participation, interaction, and social presence. *Language Learning & Technology*, 22(1), 157–183. <https://dx.doi.org/10.125/44586>
- *Scianna, J., Woodard, M., Galarza, B., Lee, S., Kaliisa, R., & Quesada, H. V. (2022). Community at a distance: Understanding student interactions in course-based online discussion forums. In C. Damşa & A. Barany (Eds.), *Advances in Quantitative Ethnography. ICQE 2022* (pp. 270–284). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-31726-2_19
- Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics*, 3(3), 9–45. <https://doi.org/10.18608/jla.2016.33.3>
- Shea, P., & Bidjerano, T. (2009). Community of inquiry as a theoretical framework to foster “epistemic engagement” and “cognitive presence” in online education. *Computers & Education*, 52(3), 543–553. <https://doi.org/10.1016/j.compedu.2008.10.007>
- Shea, P., & Bidjerano, T. (2010). Learning presence: Towards a theory of self-efficacy, self-regulation, and the development of a communities of inquiry in online and blended learning environments. *Computers & Education*, 55(4), 1721–1731. <https://doi.org/10.1016/j.compedu.2010.07.017>
- *Shea, P., Hayes, S., Uzuner-Smith, S., Gozza-Cohen, M., Vickers, J., & Bidjerano, T. (2014). Reconceptualizing the community of inquiry framework: An exploratory analysis. *The Internet and Higher Education*, 23, 9–17. <https://doi.org/10.1016/j.iheduc.2014.05.002>
- *Shea, P., Hayes, S., Uzuner Smith, S., Vickers, J., Bidjerano, T., Gozza-Cohen, M., Jian, S.-B., Pickett, A. M., Wilde, J., & Tseng, C.-H. (2013). Online learner self-regulation: Learning presence viewed through quantitative content and social network analysis. *The International Review of Research in Open and Distributed Learning*, 14(3), 427–461. <https://doi.org/10.19173/irrodl.v14i3.1466>
- *Shea, P., Hayes, S., Vickers, J., Gozza-Cohen, M., Uzuner, S., Mehta, R., Valchova, A., & Rangan, P. (2010). A re-examination of the community of inquiry framework: Social network and content analysis. *The Internet and Higher Education*, 13(1–2), 10–21. <https://doi.org/10.1016/j.iheduc.2009.11.002>
- Siebert-Evenstone, A. L., Irgens, G. A., Collier, W., Swiecki, Z., Ruis, A. R., & Shaffer, D. W. (2017). In search of conversational grain size: Modeling semantic structure using moving stanza windows. *Journal of Learning Analytics*, 4(3), 123–139. <https://doi.org/10.18608/jla.2017.43.7>
- Swiecki, Z., & Shaffer, D. W. (2020). iSENS: An integrated approach to combining epistemic and social network analyses. In M. Scheffel, V. Kovanović, N. Pinkwart, & K. Verbert (Chairs), *The Tenth International Conference on Learning Analytics & Knowledge* (pp. 305–313). ACM. <https://doi.org/10.1145/3375462.3375505>

- *Tirado-Morueta, R., Hernando, Á., & Aguaded, J. I. (2015). The effect of centralization and cohesion on the social construction of knowledge in discussion forums. *Interactive Learning Environments*, 23(3), 293–316. <https://doi.org/10.1080/10494820.2012.745437>
- *Tirado-Morueta, R., Maraver-López, P., Pérez-Rodríguez, A., & Hernando-Gómez, Á. (2020). Exploring social network structure patterns suitable to the community of inquiry model moderated by the task. *Journal of Educational Computing Research*, 58(2), 319–342. <https://doi.org/10.1177/0735633119845695>
- *Wang, Y., & Liu, Q. (2020). Effects of online teaching presence on students' interactions and collaborative knowledge construction. *Journal of Computer Assisted Learning*, 36(3), 370–382. <https://doi.org/10.1111/jcal.12408>
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.
- *Wicks, D., Craft, B. B., Lee, D., Lumpe, A., Henrikson, R., Baliram, N., Bian, X., Mehlberg, S., & Wicks, K. (2015). An evaluation of low versus high collaboration in online learning. *Online Learning*, 19(4), 67–87. <https://doi.org/10.24059/olj.v19i4.552>
- *Williams-Dobosz, D., Azevedo, R. F. L., Jeng, A., Thakkar, V., Bhat, S., Bosch, N., & Perry, M. (2021). A social network analysis of online engagement for college students traditionally underrepresented in STEM. In M. Scheffel, N. Dowell, S. Joksimović, & G. Siemens (Chairs), *LAK21: The 11th International Learning Analytics and Knowledge Conference* (pp. 207–215). ACM. <https://doi.org/10.1145/3448139.3448159>
- *Wu, Q., Koszalka, T. A., & Niu, Z. (2022). Exploring levels and patterns of social presence in asynchronous online discussions: A longitudinal study. *Distance Learning*, 19(4), 1–18. <https://doi.org/10.1108/DL-05-2023-0001>
- *Yen, C.-J., Tu, C.-H., Tankari, M., Özkeskin, E. E., Harati, H., & Miller, A. (2022). A predictive study of students' social presence and their interconnectivities in the social network interaction of online discussion board. *Journal of Applied Learning & Teaching*, 5(1), 27–39. <https://doi.org/10.37074/jalt.2022.5.1.13>

