IRRODL

THE INTERNATIONAL REVIEW OF RESEARCH IN OPEN AND DISTRIBUTED LEARNING

Vol 23 No 2

May 2022

creative commons

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

Editorial - Volume 23 Issue 2

Constance Blomgren
Associate Editor, Athabasca University

Welcome *IRRODL* authors, reviewers, and readers to this second issue of 2022.

We have exciting news that the *IRRODL* article by **Aras Bozkurt** and **Olaf Zawacki-Richter** recently has been twice honoured. Trends and Patterns in Distance Education (2014–2019): A Synthesis of Scholarly Publications and a Visualization of the Intellectual Landscape was selected "Fred Mulder Best Open Education Research Paper" by The Global OER Graduate Network (GO-GN) for 2021. As well, this article was one of ten good reads for 2021 recommended by the National Institute for Digital Learning (NIDL). The NIDL additional reading list included two more *IRRODL* articles: **Debra Dell's** Resonance and Current Relevance of IRRODL Highly-cited Articles: An Integrative Retrospective and the article by **Jewoong Moon** and **Yujin Park**, A Scoping Review on Open Educational Resources to Support Interactions of Learners with Disabilities. Congratulations to these authors – we are all richer for your exceptional work.

We have fifteen research articles and two book reviews to offer our readership.

Mapping Network Structure and Diversity of Interdisciplinary Knowledge in Recommended MOOC Offerings by **Jingjing Zhang, Yehong Yang, Elena Barbera, and Yu Lu** offers evidence-based analytics. The authors map disciplinary and interdisciplinary network structures and the implications for future online course design.

Wei He, Li Zhao, and Yu-Sheng Su studied the *Effects of Online Self-Regulated Learning on Learning Ineffectiveness in the Context of COVID-19*. Structural equation modelling data from high school students in Jiangsu Province, China indicate that better performance in the three stages of self-regulated learning decreases the perception of online learning ineffectiveness.

From Physical to Virtual: A New Learning Norm in Music Education for Gifted Students by **Md Jais Ismail, Azu Farhana Anuar, and Fung Chiat Loo** provides research into online music education for youth. The quantitative findings indicate that five domains of motivation for student success are enhanced by online distance learning.

Daniel Villar-Onrubia examined open courseware implementation through a qualitative multi-method approach. The study results indicate a disconnect between the implementation and the opportunities academics encounter in this area of open education. Read the details in *"They Have to Combine the Future of the University and Their Own Future": OpenCourseWare (OCW) Authoring as an Academic Practice in Spain*.

Mete Akcaoglu and Mustafa Ozturk Akcaoglu used a cross-sectional survey study to investigate pre-service teachers and their understanding of key components of online distance learning. With the current interest in online distance learning this research provides a timely contribution. Read the findings in their article, *Understanding the Relationship Among Self-efficacy, Utility Value, and the Community of Inquiry Framework in Preservice Teacher Education*.

The Effects and Implications of Using Open Educational Resources in Secondary Schools by **Paul Harvey and John Bond** contributes to the growing scholarship of OER for primary and secondary education. Twenty-eight Washington State schools using math OER with middle school students provided the context that examines OER curriculum effects, whether time duration of curriculum use influences math results, and the influence of other factors on student achievement when using math OER.

Mohsen Keshavarz, Zohrehsadat Mirmoghtadaie, and Somayyeh Nayyeri through their research, designed and evaluated a tool to measure the effective management of the virtual classroom. Read further about this tool in *Design and Validation of the Virtual Classroom Management Questionnaire A Case Study: Iran*.

Maryna Zhenchenko, Oksana Melnyk, Yaroslava Prykhoda, and Igor Zhenchenko authored *Ukrainian E-Learning Platforms for Schools: Evaluation of Their Functionality*. The findings indicate that Ukrainian e-learning platforms need further support for open access development along with improved collaboration and communication tools.

Anita Samuel and Simone C. O. Conceição contribute to instructional design research with their article, *Using the Critical Incident Questionnaire as a Formative Evaluation Tool to Inform Online Course Design: A Qualitative Study*. Their findings indicate that a formative evaluative tool administered mid-semester influenced real-time online course design and delivery.

Fine-tuned BERT Model for Large Scale and Cognitive Classification of MOOCs by **Hanane Sebbaq and Nour-eddine El Faddouli** adds to our understanding of pedagogy within MOOCs. These researchers used Bloom's taxonomy and automated the pedagogical annotation of MOOCs.

Lintang Matahari Hasani, Harry Budi Santoso, and Kasiyah Junus contribute their study, *Designing Asynchronous Online Discussion Forum Interface and Interaction Based on the Community of Inquiry Framework*. The researchers explored asynchronous online discussion forums and the Community of Inquiry applying a user-centered design method.

Are K–12 Teachers Ready for E-learning? by **Elif Polat, Sinan Hopcan, and Ömer Yahşi** examines both Turkish K–12 teachers' e-learning preparedness, and their readiness to teach online, resulting in a scale to measure this readiness.

Yuanyuan Hu, Claire Donald, and **Nasser Giacaman** take up the Community of Inquiry's cognitive presence in their study, *Cross Validating a Rubric for Automatic Classification of Cognitive Presence in MOOC Discussions*.

The first book review is by **Alexandra Miller.** This review examines *The Hidden Curriculum of Online Learning: Understanding Social Justice through Critical Pedagogy*. **Kelly Hammond** penned our second review and provides comment on *Exploratory Programming in the Arts and Humanities*.

Lots to read in this issue – enjoy!

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

Mapping Network Structure and Diversity of Interdisciplinary Knowledge in Recommended MOOC Offerings

Jingjing Zhang[1], Yehong Yang[1], Elena Barbera[2], and Yu Lu[3*]
[1]School of Educational Technology, Faculty of Education, Beijing Normal University, China;
[2]Faculty of Psychology and Education Sciences. Universitat Oberta de Catalunya, Spain
[3]Advanced Innovation Center for Future Education, Faculty of Education, Beijing Normal University, China
**corresponding author: luyu@bnu.edu.cn*

Abstract

In massive open online courses (MOOCs), recommendation relationships present a collection of associations that imply a new form of integration, such as an interdisciplinary synergy among diverse disciplines. This study took a computer science approach, using the susceptible-infected (SI) model to simulate the process of learners accessing courses within networks of MOOC offerings, and emphasized the potential effects of a network structure. The current low rate of access suggests that a ceiling effect influences learners' access to learning online, given that there are thousands of courses freely available. Interdisciplinary networks were created by adding recommended courses into four disciplinary networks. The diversity of interdisciplinarity was measured by three attributes, namely variety, balance, and disparity. The results attest to interesting changes in how the diversity of interdisciplinary knowledge grows. Particularly remarkable is the degree to which the diversity of interdisciplinarity increased when new recommended courses were first added. However, changing diversity implied that neighbouring disciplines were more likely to come to the forefront to attach to the interdisciplinarity of MOOC offerings, and that the pace of synergy among disparate disciplines slowed as time passed. In the absence of domain experts, expert knowledge is not sufficient to support interdisciplinary curriculum design. More evidence-based analytics studies showing how interdisciplinarity evolves in course offerings could help us to better design online courses that prepare learners with 21st-century skills.

*Keywords: d*istance education and online learning, informal learning, interdisciplinary projects, simulations

Introduction

Today, 21st-century skills demand ways of thinking that go beyond simple categories to include interconnections between disciplinary boundaries. Consequently, contemporary approaches to educating young people call for a new type of learning that incorporates interdisciplinary knowledge across the natural and social sciences, with the goal of solving real-world problems. However, integrating courses across disciplines in a conventional higher education setting is a demanding and challenging pursuit. Researchers (e.g., Holley, 2009; Jarmon et al., 2009; Meyers et al., 2013; Spector, 2015; Zhang, Burgos, & Dawson, 2019) have concluded that interdisciplinary learning to prepare learners with 21st-century skills demands new learning spaces that are no longer simply physical places, but also include an online environment that is supportive of informal learning.

Cormier and Siemens (2010, p. 32) argued that "online open courses allow for innovation in how educators prepare to teach, how learners negotiate knowledge from the information they are encountering, and how courses can have an impact on the broader field of study." Such online spaces are capable of accommodating different types of interaction that are open, flexible, and adaptable to new forms of knowledge integration (e.g., Gillet et al., 2005; Linn et al., 2003; Luo & Chea, 2020; Tucker & Morris, 2011). What has changed is not the knowledge itself in the current practice of online and flexible learning, but rather how the knowledge is delivered and re-constructed by the global body of learners (Maassen et al., 2018). Such learning environments allow learners to sustain interdisciplinary efforts in order to strengthen the relationships between what they have learned and other sources of knowledge and experience (Zhang, Burgos, & Dawson, 2019). Online learning, to some extent, resembles the interconnected world in which one lives and learns, and demonstrates how knowledge intersects and crosses borders (Anderson, 2008). Compared to conventional practices in higher education, the online space offers greater flexibility. It also allows knowledge to be recategorized or divided in much the same way as industry products are reconfigured to meet changing demands from rapidly developing society and digital technology (Adelman, 1999).

Nevertheless, we still do not know how this transformation occurs in the online delivery mode until it is carefully analysed. Nor will we understand how much online courses such as MOOC offerings have diversified and specialized until more courses are offered online, and data-driven research is available. This process of diversification and specialization is barely noticeable at first and then draws more attention from institutional management. Over time, the nature of MOOC offerings worldwide, and what we learn from them through analytic research, may have an impact on the current design of courses for knowledge acquisition in tertiary education. This will, in turn, contribute to preparing learners with 21st-century skills.

Network analysis enables us to make evidence-based decisions about the structural topology of current MOOC offerings. Identifying meaningful network structures among MOOC offerings provides an overall picture of the disciplinary knowledge in MOOCs. In this study, network analysis was used to explore the hidden topological knowledge structure of MOOCs on XuetangX (a widely recognized Chinese MOOC platform). The growing collection of MOOCs can be seen as a network of course nodes with links between the nodes. The linkage between any two courses came from XuetangX's recommendation mechanism. Using metrics developed in network science, this study investigated the network structures of disciplinary knowledge created in the MOOC space, and explored the extent to which such structures allowed for interdisciplinarity.

Problems of Disciplinary Structure in Conventional Higher Education

Our modern disciplinary structure has resulted in the diversification and specialization of labour and knowledge (Kockelmans, 1979). For centuries, disciplines have expanded, integrated, and scaled down; nevertheless, the main structure of higher education has not changed (Holley, 2017). The department is still arguably the dominant unit in which teaching and learning are designed and implemented for students (Trowler et al., 2003). Academic departments follow institute regulations to develop courses that address local issues and contexts (Higham, 2003), and a curriculum committee takes care of reviewing and approving the proposals for programmes and curricula (Massachusetts Institute of Technology [MIT], 2017). In this way, knowledge is, to a large extent, defined and compartmentalized based on institutional decisions. This is widely referred to as a traditional disciplinary approach to curricula (Venville, et al., 2012). This approach rests, however, on the assumption that there is a corpus of disciplinary knowledge (i.e., received wisdom that is beyond criticism; Kelly et al., 2008). "Curricula in higher education are to a large degree 'hidden curricula' . . . they take on certain patterns and relationships, but those patterns and relationships will be hidden from all concerned, except as they are experienced by the students" (Barnett, 2000, p. 260). Disciplinary knowledge is translated into curricula with reference to key criteria, standards, or educational outcomes that might be unintendedly weighted in favour of existing disciplinary capacity (West, 2018).

Given the increasing challenges facing educators in the 21st century, the process of how knowledge is produced today is global and diverse(Kahn & Agnew, 2017). There is an urgent demand for interdisciplinarity, a term commonly used to refer to the integration of knowledge from multiple disciplines (Choi & Pak, 2006). Nevertheless, the ethnocentric nature of disciplinary structures inevitably produces clusters of subject knowledge, thereby leaving gaps between them (Holley, 2017). Campbell (1969) proposed a famous fish-scale model of interdisciplinarity studies to criticize this kind of intuitional structure. Faculty are trained with respect to disciplinary norms, and interdisciplinary training is a complex endeavour for conventional higher education. Students continue to be increasingly trained within disciplinary domains, thus leaving epistemological gaps unexplored. Defining knowledge strictly in disciplinary domains has made it difficult to make potentially rich connections between various epistemological ideas (Holley, 2017). Following the arguments of the crisis of curriculum in the community (e.g., Priestley, 2011; Wheelahan, 2012), a number of researchers have agreed that higher education practices have yet to clearly define the problematic role and meaning of disciplinary knowledge that has existed for centuries (e.g., Graff, 2015; Trowler et al., 2012). Instead, most curricula are "fragmented . . . unconnected [and] rely on students' efforts to make sense of the whole" (Hubbal & Gold, 2007, p. 8).

Large Repository of MOOCs Creates Opportunities for Analytics Studies

Educational researchers face a significant challenge preparing higher education to be proactive regarding these issues. There is no doubt that educational practices have undergone remarkable changes, but many practices remain rooted in 20th-century foundations of learning (Kahn & Agnew, 2017). The disruptive innovation of MOOCs attempt to change current practices while encountering contradictory demands for higher education in modern society. Currently, there are growing concerns about high dropout rates (Ang et al., 2020; Reich & Ruipérez-Valiente, 2019), lack of learner support (Gregori et al., 2018), as well as demand for better instructional design of MOOCs (Zawacki-Richter et al., 2018) and accountability regarding assessment of MOOCs (Suen, 2014). However, what is new in MOOCs is neither the innovative pedagogies nor the transformation of higher education, but rather an unnoticeable and slow process of the increasing specialization of MOOC offerings online (Schuwer et al., 2015). There are now over 163,000 MOOCs available in cyberspace, and trends indicate greater growth and differentiation (Shah, 2020).

As online learning is open and flexible, course offerings are less likely to follow compulsory requirements of university curricula (Brown et al., 2015). For instance, MOOCs are usually offered by experts in various areas from different universities or industries. As long as this is the case, MOOCs will continue to be offered in a bottom-up fashion without necessarily following any curricula or guidelines, which will inevitably create large repositories of courses from diverse disciplines. Traditionally, the course offering decisions that institutions make are governed by conventional disciplinary practices and advisement from programme directors within academic departments (MIT, 2017). However, none of these types of advice are available in the MOOC space.

Although MOOCs offer the opportunity to create something new, we lack the confidence to question why the content of courses bears a considerable resemblance to the face-to-face courses offered by schools and institutions. MOOCs, as an example of online educational practice, predominantly reflect what tertiary education produces (Zhang, Sziegat, et al., 2019). Disciplinary knowledge is commonly regarded as content knowledge. Online higher education preserves the tradition that the majority of such content knowledge is borrowed from conventional higher education (Naidu, 2017). As such, content knowledge has been criticized for decades, since content transfer implies behaviourist-type, old-fashioned approaches to learning (Eynon, 2017). Thus, there is an urgent call for educational researchers and practitioners to think outside the box, and to face the challenges of designing, planning, and implementing strategic changes to knowledge and curricula in the 21st century (Brown et al., 2015). The flexibility necessary for the design of interdisciplinary courses may be difficult to achieve and even harder to preserve in MOOCs. Nevertheless, as Gašević et al. (2014) argued, there is a need for "increased efforts towards enhancing interdisciplinarity" (p. 134) in MOOC research. One way to help interdisciplinary learning occur is by making sure that there is some disciplinary cohesion in the cohorts of MOOC offerings as well.

Since 2012, knowledge delivered online has worked its way up. MOOC offerings have produced a huge amount of data that allowed many analytic studies focused on learning processes, discussion, engagement, and self-regulation. Refer to Mangaroska and Giannakos (2018) and Tsai et al. (2020) for comprehensive reviews. A prominent trend of these studies is the use of an analytic approach to studying learners' behaviours through the exploration of engagement and dropouts, together with a careful examination of prediction and assessment. While such work has arguably illuminated our understanding of what contributes to quality learning and teaching online, the selection of disciplinary knowledge and its organization have often been neglected, as they have been deemed unimportant in this kind of analytic research into MOOCs. Network structures, while widely recognized in network science, have not been used to examine disciplinary knowledge created in the MOOC space. The structure of MOOC offerings reveals conceptual models of the different parts or components as well as course organization or structure. As more online courses are increasingly offered, making connections between different courses from different disciplines in a more synthesized way is an effective mechanism to support strategic institutional decisions on what MOOCs to offer in the future. Identifying the network structures of MOOC offerings could also provide insights into curriculum design for open and flexible learning. Such an approach to evidence-based decision making might discover important insights that would not have been identified through the conventional process of curriculum development (West, 2018). It is also critical for allowing truly interdisciplinary synergy that is not constrained by the unintended bias of a single discipline.

Methodology

This study sought to understand the network structure of MOOC offerings and the diversity of interdisciplinary knowledge offered by MOOCs by simulating the process of learners selecting recommended courses. In this study, we did not examine knowledge, as it has been explored in cognitive science. Instead, knowledge is only used as a result of a course, which is associated with a certain discipline. That is, the fact that a course is offered in a certain discipline is seen as contributing to a certain disciplinary knowledge, which is similar to bibliographic work or science of science research (e.g., Veletsianos & Shepherdson, 2015), in which journal articles are regarded as knowledge produced in certain disciplines. Our representation of disciplinary knowledge in relation to interdisciplinary knowledge was the product of network structures of recommended MOOCs offered on XuetangX.

The Case and Data Collection

XuetangX was selected as the case for this research study. As one of the top five MOOC providers worldwide, XuetangX hosts Chinese MOOCs, as well as courses from a consortium of leading universities worldwide. XuetangX, powered by the open-source platform edX, has expanded a number of features, such as recommendation systems, that are not found in many other MOOC platforms. Recommendations offered on MOOC platforms can play a useful hidden role in the selection of courses. On XuetangX, for each course, three recommended courses are provided on the right-hand side of the course page (as shown in Figure 1). While these current recommendations may not be intelligent enough to suggest that courses are related in any way, there might be an association of some kind that is worth exploring. For example, the fact that course A in engineering is recommended by course B in computer science implies not only that is there a relationship between course A and course B, but also that the knowledge bases of engineering and computer science are related. A collection of these kinds of individual associations offers a new form of integration, such as an interdisciplinary synergy between engineering and computer science. Identifying meaningful network structures of course offerings provides an overall picture of disciplinary knowledge (West, 2018).

Figure 1

Interface of a MOOC on XuetangX

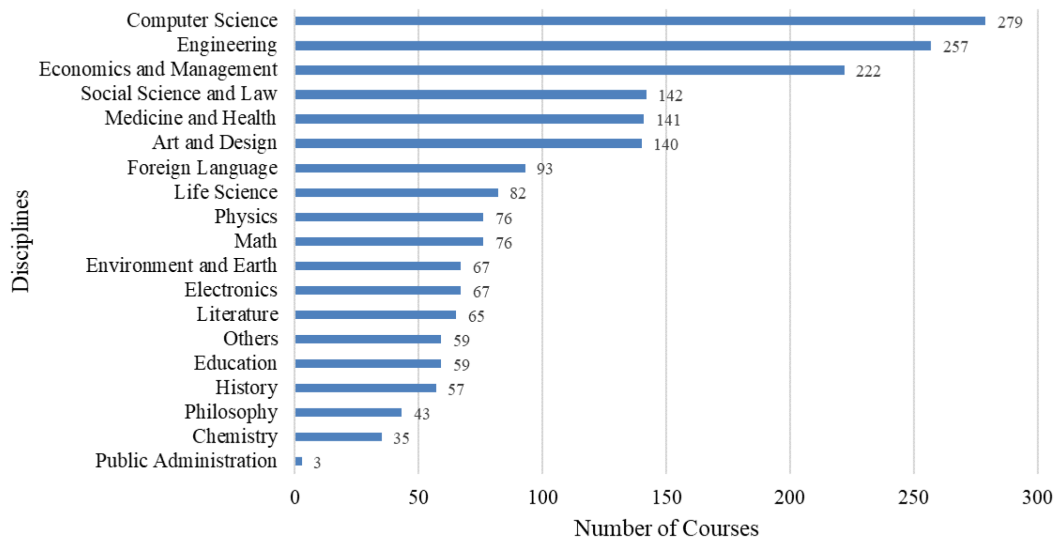
Note. Image captured from XuetangX website (*https://www.xuetangx.com/)*.

In the current study, a crawler programme written in Python was used to collect information about courses and the links recommending these courses from XuetangX. The data were collected until May 2019 from a total of 2,017 courses, among which 526 courses were exported from other MOOC platforms, such as edX. The disciplinary information of 1,990 courses was obtained from the original MOOC platform; 27 courses had no associated disciplinary labels. Each platform adopts different discipline categories to differentiate their courses. For example, 30 different disciplines are provided on edX, while only 21 different disciplines are available on XuetangX. In the present study, we followed XuetangX's discipline categories and merged the disciplines provided on edX into XuetangX's categories. For example, one discipline label—university prerequisite, containing 11 courses—was deleted, and another discipline, entrepreneurship, was merged into economics. By doing so, 19 final categories of disciplines were formed.

As 16 courses had no associated recommended courses (and were thus removed from XuetangX), a total of 1,963 courses were identified within 19 disciplines. As shown in Figure 2, the majority of the courses on XuetangX were within the disciplines of computer science, engineering, economics and management, and social sciences and law. Therefore, these top four disciplines that offer the majority of MOOCs were selected as the cases through which we explored the topological structure of their recommendation networks.

Figure 2

Number of Courses Within Each of the 19 Disciplines on XuetangX



Data Analysis

This study adopted computational and systems modelling, which uses simulation as a model to investigate complex systems, given that social and cultural perspectives of learning are interwoven with multiple levels of interactions (Janežič et al., 2018). As it seems that no researchers have collected learner behaviour data across different courses for a whole MOOC platform, it is impossible to use real learner behaviour data to create a disciplinary structure of MOOC offerings. Thus, we adopted the susceptible-infected (SI) model to illustrate how courses are accessed by learners in the recommended course network. We emphasized the potential effect induced by network structure rather than that from the learners' perspective. The SI model is used to simulate the spread of a disease in a population. It is a simple but common method of modelling the interaction of two populations in a network. In using this model, some assumptions are necessary. First, the size of the population is fixed, and we use N to represent it. Second, there are two classes of individuals in the population. S represents susceptible individuals who do not have the disease but are susceptible to it and I represents infective individuals who have the disease and are infectious; $S + I = N$ is always satisfied. Third, disease spreads through interactions between pairs of individuals—from infective individuals to susceptible individuals. Finally, an infectious rate between 0 and 1 is constant during the whole process (Allen, 1994).

In our research, we assumed that learners played the role of a disease in the SI model, and courses in the network act as individuals. Individuals are connected by recommendation links on Web pages, wherein the disease spreads through recommendation links. During the spreading process, there are only two classes of individuals (i.e., course resources)—S represents course resources not viewed and I represents course resources that have been viewed. The number of course resources in the network is fixed, and is equal to the sum of the unviewed and the viewed courses. Learners acquire neighbour resources moving through the recommendation links within a range of probability, and here, we assumed that the probability was constant.

Analysis of the Diversity of Interdisciplinary Knowledge

In this research, we explored how the degree of diversity changed using SI simulation for four interdisciplinary networks, namely engineering, social science, computer science, and economics and management. Evidence has indicated that the use of multiple metrics can reveal the differences between various bodies of disciplinary knowledge (Porter & Rafols, 2009). We measured the diversity of interdisciplinarity in networks comprising courses from different disciplines by three attributes, namely variety, balance, and disparity (Stirling, 2007). Combining these three metrics enabled us to measure the interdisciplinary knowledge of MOOCs to a level of detail that has been previously unexplored. The diversity of interdisciplinarity in networks comprising courses from different disciplines was measured by three attributes of variety, balance, and disparity/similarity (Stirling, 2007).

The variety attribute measured the number of distinct disciplines in which courses were offered—the greater the variety, the greater the diversity. We calculated variety by the ratio of the number of links that pointed to courses from different disciplines to the total links that pointed to all courses. Balance indicated the even distribution of these disciplines, analogous to statistical variance. The more even the balance, the greater the diversity. Entropy, proposed by Shannon (2001), has been widely used in thermodynamics; we used it as a metric to represent the balance of interdisciplinarity. The greater the entropy, the more even the balance. If an interdisciplinary network contained courses from a wide range of disciplines evenly, then the diversity level of this interdisciplinary network was high. Disparity illustrates the degree to which these disciplines differed. The Rao-Stirling index was used to measure disparities in interdisciplinarity.

Results

Topological Structure of Disciplinary Networks

As shown in Figure 3, the topological structures of the four selected disciplines varied greatly. All four disciplinary networks had significant community structures (Newman & Girvan, 2004), as the modularity values for all four disciplinary networks were greater than 0.3 (see Table 1). Moreover, we found that social sciences and law had the highest modularity value of 0.946. This shows that this disciplinary network had a better community structure than the other three networks.

Table 1

Summary of Statistical Properties of Four Disciplinary Networks

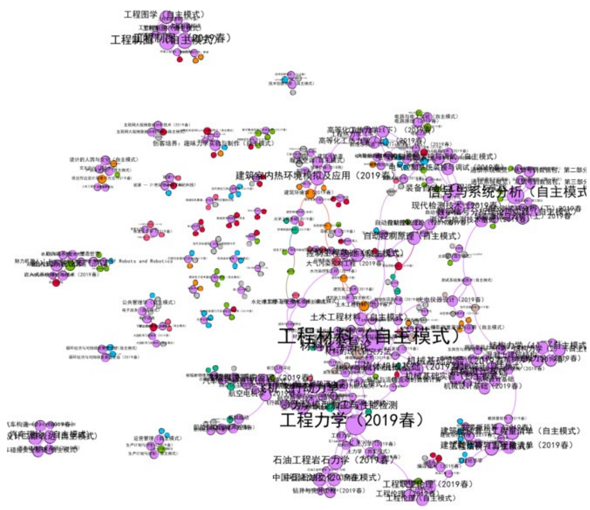
Network metrics	Computer science	Engineering	Economics and management	Social sciences and law	Entire network
Scale	407	451	323	315	1963
Edge	799	757	657	421	5715
Average degree (AD)	1.963	1.678	2.034	1.337	2.911
Diameter (D)	11	11	13	6	47
Average path length (APL)	3.403	2.913	2.842	1.685	10.856
Network density (ND)	0.005	0.004	0.006	0.004	0.001
Number of weakly connected components	19	32	16	45	6
Number of strongly connected components	311	352	246	279	1,076
Number of communities	35	45	29	49	39
Modularity	0.854	0.899	0.859	0.946	0.843

Different communities (represented in different colours in Figure 3) contained courses in different topics. The courses in the same community were very densely connected, which showed that the courses had similar themes. For example, in the subnetwork of computer science, Figure 3 (b), red represents one community. In this community, the courses were *Foundation of C*, *Foundation of C++*, *Advanced Course in C*, and so on. These courses belonged to the same theme: learning programming languages.

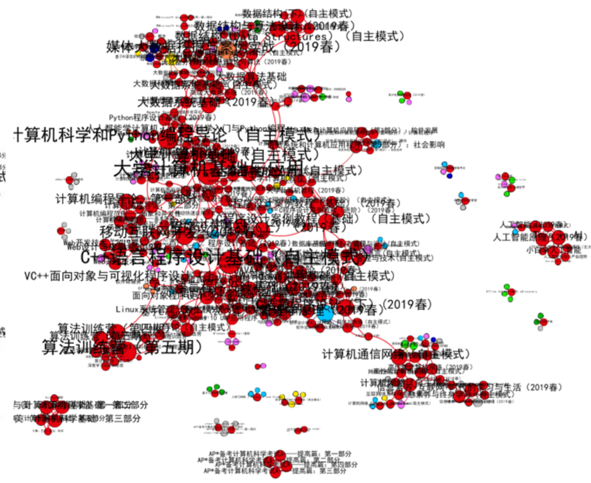
As shown in Figure 3, four disciplinary networks were not complete networks, but rather consisted of multiple independent subgraphs. The network of economics and management had the lowest degree of separation, with 16 subgraphs. The smallest subgraph had only three courses, while most of the remaining courses formed other complete subnetworks. The computer science network also has a similar distribution of subgraphs. The smaller subgraphs in computer science contain only two or three courses. There were some differences between the larger subgraphs and the smaller subgraphs. For example, in the computer science network, the theme of the smallest subgraph was electronics, but the theme of the largest subgraph was programming and engineering.

Figure 3

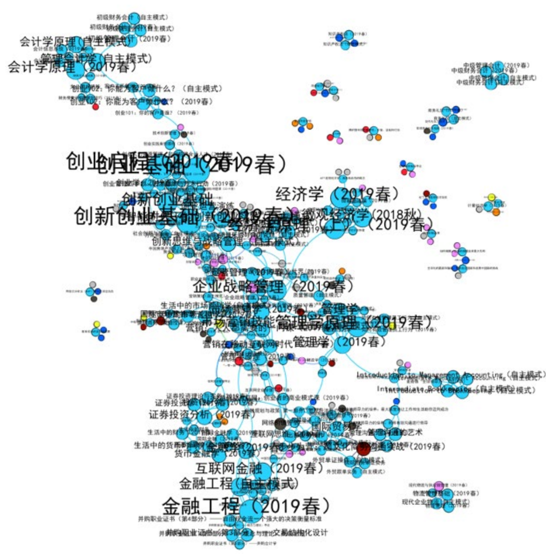
Recommendation Networks of Four Different Disciplines



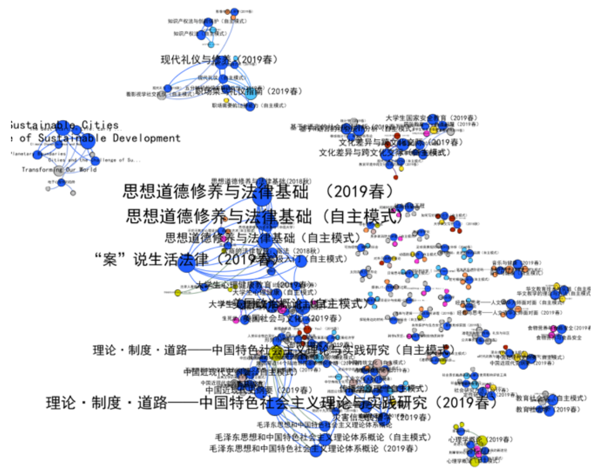
(a) Engineering



(b) Computer Science



(c) Economics and Management



(d) Social Sciences and Law

Note. Colours represent detected communities (purple: engineering; green: EE; light blue: economics and management; red: computer science; dark blue: social sciences and law. The size of a node indicates the number of linkages.

As shown in Figure 3, the courses in economics and management were most densely interconnected. This structure can be measured by the average degree, a global description for all the nodes in the network, and used to measure the average number of neighbours the nodes have. The average degree of the economics and management network was 2.034, which was the largest value among the four disciplinary networks. On the other hand, the average degree of the social sciences and law network was only 1.337, the lowest value among all four disciplinary networks.

The information dissemination capability of a network can be measured by the network diameter (i.e., the maximum distance across the network), and the average path (i.e., the average of any distance between two nodes in the network; Lin, 2009). In the present study, a path length refers to the number of intermediate nodes that need to be communicated or exchanged between any two given courses. The network diameter and average path length for the social sciences and law subnetwork (D = 6, APL = 1.685) were shorter than those of the other three sub-networks (i.e., engineering: D = 11, APL = 2.913; computer science: D = 11, APL = 3.403; and economics and management: D = 13, APL = 2.842). These results indicated that learners in social sciences and law course networks took fewer steps to find other related courses in the same discipline. In this way, the knowledge structure of this discipline seems to be more conducive to learners' access.

In a directed graph, there are two kinds of connected components, namely strongly connected and weakly connected, depending on how we treat the directionality of links. In a strongly connected component, every node is reachable from every other node, while in a weakly connected component, every node is reachable from every other node, ignoring directions (Tabassum et al., 2018). Regarding the computer science discipline, the whole network contained 19 weakly connected components. Eight out of ten (79.85%) of the courses were bounded together in the largest connected component, which contained almost all of the total recommendation links (n = 694, 86.86%). It consisted predominantly of courses in the topics of programming and data analysis, which are the basic and fundamental courses in computer science. In this network, programming courses such as *C++ Programming* and *Computer Fundamentals and Applications* occupied central positions in the computer science recommendation network, in which we used the network metrics - in-degree - to measure the importance of a course.

Forming Interdisciplinary Networks of MOOCs

Interdisciplinary networks (i.e., extended disciplinary networks) were created by adding neighbour courses into the four disciplinary networks (as shown in Figure 4). We continued this process until no further course in the previous steps was added. Four interdisciplinary networks, one for each discipline, were created. For example, in the engineering discipline, in the first step, 194 courses that were recommended by the courses in engineering disciplines were added, of which most were courses in electronics, computer science, and economics. In the second step, using a similar mechanism, 158 courses were added, and this process was continued until step 12. After step 12, there were no courses available that had not been added in the previous steps. For three of the disciplinary networks, it took 12 to 13 steps to converge; however, economics and management required 22 steps to form an extended disciplinary network. For the extended computer science network, most of the newly added courses were from computer science, economics, and engineering. For the extended economics network, engineering and computer science made up most of the recommended courses. The social sciences and law network favoured courses from the disciplines of engineering, computer science, and economics.

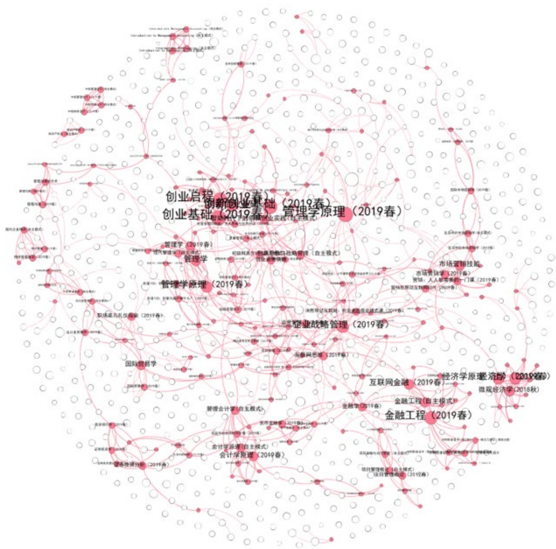
At the end of this process, for each of these connected interdisciplinary networks, there were approximately 857 to 917 courses. There was a large overlap among these four networks, with a total of 561 courses common to all four interdisciplinary networks. Between each pair of the two networks, at least 66% of the courses were the same, and approximately 563 courses belonged exclusively to only one network.

Interestingly, the strongest connected components across the four extended disciplinary networks were in the same network and included the same 30 courses. Among these courses, 15 were related to physics, 10 were related to engineering, and one was related to each of the environment and earth, computer science, history, and math. One course in the set was categorized as other.

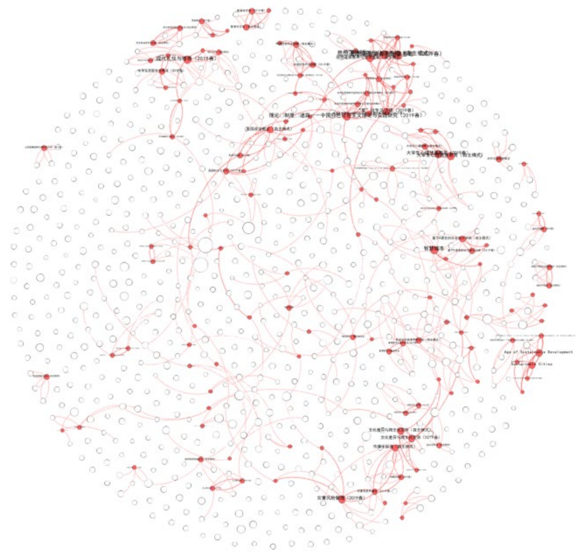
Figure 4

The Four Interdisciplinary Networks

(a) Engineering (b) Computer Science



(c) Economics and Management



(d) Social Sciences and Law

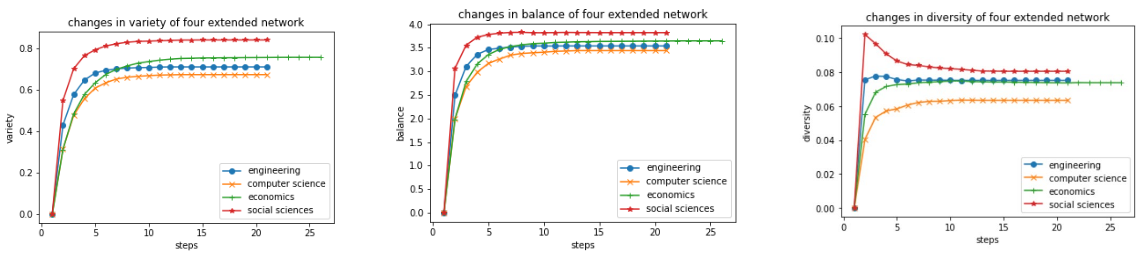
Note. Red nodes represent the courses within the same disciplines, and white nodes represent the courses from other disciplines.

The Diversity of The Interdisciplinary Networks

As shown in Figure 5 (left), while forming interdisciplinary networks by simulation, the variety of interdisciplinarity increased sharply first and then tended to stay steady. The variety of the extended social sciences and law network was relatively higher than that of the other three networks. In the first step, the variety of the extended social sciences and law network increased to 54.92% and then up to 84.19%. That is, learners who took courses in the social sciences and law were more likely to access courses from other disciplines. In the second simulation step, the disciplines that were added to the interdisciplinary networks increased sharply; courses recommended by the first cohort were likely to be those from different disciplines. This further implied that the first cohort of recommended courses to extend beyond the original disciplines were interdisciplinary and were likely to recommend further courses from different areas. As shown in Figure 5 (middle), the changes in balance in each interdisciplinary network followed the same pattern of variety. As shown in Figure 5 (right), the value of the Rao-Stirling index increased sharply while simulating the process of forming interdisciplinary networks. That is, the disparity of the interdisciplinary networks increased. This implied that when learners seek courses that extend beyond their own discipline, they were likely to access courses that are distinct from those they have previously studied.

Figure 5

Forming Interdisciplinary Networks Through Simulation



Note. Left-hand panel depicts variety, middle panel depicts balance, and right panel depicts disparity.

Figure 5 illustrates the use of variety, balance, and disparity to measure the diversity of interdisciplinarity networks. In our study, social sciences and law had the highest level of diversity, followed by engineering, economics, and computer science. The interdisciplinary network for social sciences and law evidenced more balance than that of computer science, which was dominated by its own courses. Additionally, computer science tended to take more steps to converge to a upper limit than did the other disciplines. That is, the recommended courses tended to stick to their own disciplines or neighbours.

While adding new courses to disciplinary networks, the variety, balance, and disparity all increased sharply at first and then tended to converge to a upper limit. It is worth noting that, for social sciences and law, disparity increased at first and then declined. Disparity measured how different disciplines were integrated when new courses were added step by step. During the first step of adding new recommended courses, the degree of difference between disciplines was very high. In other words, the recommended social sciences and law courses belonged to very different disciplines.

Information Diffusion Within Four Interdisciplinary networks

From Table 2, we see that although there were many courses available for the learners in different disciplinary networks, learners were limited to accessing all the recommended courses in the same discipline. For example, in the computer science network, students had access to an average of only 2.33% of the recommended courses, while the average increased to 5.42% to 7.51% for interdisciplinary network. This low information diffusion rate was because some courses did not recommend other courses even though they themselves had been recommended by them.

Table 2

Results of Information Diffusion of Eight Interdisciplinary Networks

Disciplinary network			Discipline	Extended disciplinary (interdisciplinary) network		
Minimum (%)	Maximum* (%)	Average (%)		Minimum (%)	Maximum* (%)	Average (%)
0.22 1 course	12.86 58 courses	1.42 6 courses	Engineering	0.11 1 course	52.53 468 courses	7.21 64 courses
0.25 1 course	12.78 52 courses	2.33 9 courses	Computer science	0.12 1 course	52.39 449 courses	7.51 64 courses
0.31 1 course	14.86 48 courses	2.61 8 courses	Economics and management	0.11 1 course	55.40 51 courses	5.42 50 courses
0.32 1 course	5.08 16 courses	0.93 3 courses	Social sciences and law	0.11 1 course	52.12 47 courses	6.65 60 courses

* The strongest connected component in the network reached the maximum capacity.

Using the SI model, the strongest connected component reached a high level of convergence. Nevertheless, the convergence only reached approximately 5.08% to 14.86% for any course, and belonged to the strongest connected component in the recommendation networks of four major disciplines. Learners needed to navigate approximately 62 to 122 times to access 5.08% to 14.86% of the courses on this particular recommendation network. In contrast, approximately 52.12% to 55.40% of courses could be accessed by navigating from any course belonging to this strongest connected component of the interdisciplinary networks. The remaining strongest connected components with a smaller number of courses were likely to reach low levels of convergence, with few exceptions. That is, the topological structures of these interdisciplinary networks tended to follow a large strongly connected component conducive to global information diffusion, while the remaining smaller strongly connected components supported local information diffusion.

Discussion

Network Structures of MOOCs Affect Access to Disciplinary Knowledge

In this research, MOOC recommendation relationships were used to create networks of MOOC offerings, and the topological structures of the recommended MOOC networks presented earlier were carefully examined to illustrate how the network structures of four disciplinary knowledge differed.

The four disciplinary networks had significant community structures, but the topological structures varied greatly. The disciplinary networks were not complete networks but they all demonstrated one large connected component and multiple smaller, independent subgraphs. In computer science, for example, eight out of ten of the courses were bounded together in the largest connected component, consisting predominantly of courses in the topics of programming and data analysis. The courses in economics and management were most densely interconnected. In contrast, the social sciences and law network was loosely connected, as indicated by the average degree. Nevertheless, social sciences and law had the shortest network diameter, which implied that these courses have a greater capability for information dissemination. These findings are important as they extended past research using bibliographic data to examine disciplinary knowledge.

The topological structures of interdisciplinary networks represent the intrinsic and potential constraints on the flow of information, which alters the online context in which interdisciplinary learning might occur. To understand how these structures influence information diffusion, the SI model was adopted to simulate how courses were accessed in four different disciplinary networks. This simulation helped us see how efficiently courses (representing disciplinary knowledge) were accessed, and how many courses were eventually acquired in forming respective networks of interdisciplinary knowledge. The process of disease (access to courses) spread started from each course in the network; we then calculated the average value of the proportion of accessed (infective) courses over time. For example, in the computer science network, students had access an average of only 2.33% of the recommended courses. One could argue that learners might not follow the recommended courses in regard to their learning, but as we argued earlier, recommendations for a course of study were examined as a mechanism for providing courses to learners, similar to courses suggested by tutors or institutions in conventional education. Such recommendations also implied a certain association between different disciplines or topics. This low rate of accessibility suggested that there was a ceiling effecting what learners could access or learn online, given that there are approximately 2,700 courses freely available. The concept of more is better might seem desirable in many circumstances that enable access to education, but these strategies, frequently used by industry, must be challenged by educational researchers (Sheail, 2018). Given the rate at which ever-increasing offerings of MOOCs are developed, the rate at which learners move around and select courses from different disciplines is actually very slow. The results of this study showed that the network structure of recommended courses served as a structure of disciplinary knowledge, and it affected how far and how quickly learners could approach all these courses. The results reported herein are consistent with studies that have emphasized how the network structure affects information diffusion in the knowledge management area (Arnaboldi et al., 2016; Lambiotte & Panzarasa, 2009; Reagans & McEvily, 2003).

Diversity of Interdisciplinary Knowledge in the Process of Knowledge Integration

In this study, interdisciplinary networks (i.e., extended disciplinary networks) were created by adding neighbour courses into the current four disciplinary networks. The findings shed light on the role that disciplinary knowledge plays in forming a new interdisciplinary network. Strong connections (i.e., the same 30 courses in the strong connect component shared by all four interdisciplinary networks) made up an important part of each interdisciplinary network, as such connections were conducive to learners accessing more courses in the whole network.

The disparity metric we used not only considered the number of disciplines added into the interdisciplinary network but also measured how distant the knowledge sources were, in other words, the disparity of disciplines (Porter & Rafols, 2009). To control the number of disciplines added to the network, it was very important to measure distance. Interdisciplinarity is often interpreted as inherent in a MOOC designed by experts from two or more disciplines, though this provides no basis for exploring the kinds of disciplines that should be integrated to create interdisciplinary knowledge.

The results of simulating the process of adding new recommended courses to the original networks attested to the interesting changes in how the diversity of interdisciplinary knowledge grew. Particularly remarkable is the degree to which the diversity of interdisciplinarity increased when first adding recommended new courses. All three metrics—variety, balance, and disparity—increased sharply. However, the changes in diversity implied that neighbouring disciplines were more likely to come to the forefront to attach to interdisciplinarity in MOOC offerings, and that the synergy between disparate disciplines proceeded at a much slower pace. This is mainly because when adding more recommended courses to the network, the added courses tended to be in a discipline that was not distant from or was the same as the previous ones; thus, they did not add as much interdisciplinarity. Moreover, for disciplines such as the social sciences and law, in which courses were loosely connected, the disparity level increased sharply when new recommended courses were first added. This implied that newly added courses belonged to disciplines that were distant from the previous ones.

It seems that the measures of interdisciplinarity in the studied networks indicated the considerable interchange of subject knowledge when learners considered moving out of their comfortable zone (i.e., their own discipline). By no means did the modern knowledge gap between different disciplines restrict learners from taking the initiative to select courses in another discipline, although the difficulty attached to doing so was beyond that which many of us anticipated.

In summary, knowledge integration, as evidenced by simulating the process of learners following recommendation links to access courses, drew mainly on neighbouring disciplines. The knowledge offered in such a MOOC space is arguably becoming more interdisciplinary but in a modest manner. Only a slow increase in the small segments of knowledge from more distant disciplines was observed, which is consistent with the findings from studies using journal articles to map the changes in interdisciplinary knowledge (Porter & Rafols, 2009).

Limitations and Future Work

Focusing on the question of how to prepare students to meet the 21st-century demand for skills led us to explore how delineating the structures of MOOC offerings allowed for knowledge integration using a simulation model. Similar to arguments by Fernández-Díaz et al. (2017) regarding the pedagogic architecture of a MOOC, we believe that the topological network structure of MOOC offering and the diversity of such networks influence learners' access to courses in the MOOC space. However, we are very conscious that an examination of the typological structure and diversity of interdisciplinary knowledge can perhaps only make an indirect contribution to the debates surrounding what online courses to offer in order to prepare students regarding 21st-century skills, and how these courses relate to each other. Nevertheless, this study offers an alternative approach to mapping interdisciplinary knowledge using network analysis, and it urges that analytics studies come to the forefront regarding using available course information to provide evidence in support of the design of online education. As we lack domain experts, expert knowledge is not sufficient to support interdisciplinary curriculum design, as argued earlier. More evidence-based analytics studies could help us by providing more evidence on how to increase access by a greater number of learners to respectively form interdisciplinary networks of knowledge (Rohs & Ganz, 2015).

Acknowledgements

This work was supported by Beijing Municipal Educational Science Foundation [Project No. AHDB19030].

References

Adelman, C. (1999). *The new college course map and transcript files: Changes in course-taking and achievement, 1972-1993*. [Based on the postsecondary records from two national longitudinal studies]. US Government Printing Office.

Allen, L. J. (1994). Some discrete-time SI, SIR, and SIS epidemic models. *Mathematical Biosciences, 124*(1), 83–105. https://doi.org/10.1016/0025-5564(94)90025-6

Anderson, T. (Ed.). (2008). *The theory and practice of online learning*. Athabasca University Press.

Ang, K. L. M., Ge, F. L., & Seng, K. P. (2020). Big educational data analytics: Survey, architecture and challenges. *IEEE Access, 8*, 116392–116414. https://doi.org/10.1109/ACCESS.2020.2994561

Arnaboldi, V., Conti, M., La Gala, M., Passarella, A., & Pezzoni, F. (2016). Ego network structure in online social networks and its impact on information diffusion. *Computer Communications, 76*, 26–41. https://doi.org/10.1016/j.comcom.2015.09.028

Barnett, R. (2000). Supercomplexity and the curriculum. *Studies in Higher Education, 25*(3), 255–265. https://doi.org/10.1080/713696156

Brown, M., Costello, E., Donlon, E., & Giolla-Mhichil, M. N. (2015). A strategic response to MOOCs: How one European university is approaching the challenge. *The International Review of Research in Open and Distributed Learning*, *16*(6). https://doi.org/10.19173/irrodl.v16i6.2151

Campbell, D. T. (1969). Ethnocentrism of disciplines and the fish-scale model of omniscience. In M. Sherif & C. W. Sherif (Eds.), *Interdisciplinary relationships in the social sciences* (pp. 328–348). Routledge. https://doi.org/10.4324/9780203788363-19

Choi, B., & Pak, A. (2006). Multidisciplinarity, interdisciplinarity and transdisciplinarity in health research, services, education and policy: Definitions, objectives, and evidence of effectiveness. *Clinical and Investigative Medicine, 29*, 351–364.

Cormier, D., & Siemens, G. (2010). The open course through the open door: Open courses as research, learning, and engagement. *Educause Review*, (July/August), 31–39. https://www.islandscholar.ca/islandora/object/ir%3A3675/datastream/PDF/view

Eynon, R. (2017). Crowds, learning and knowledge construction: questions of power and responsibility for the academy. *Learning Media, and Technology*, 42(3), 257-258. http://doi.org/10.1080/17439884.2017.1366920

Fernández-Díaz, E., Rodríguez-Hoyos, C., & Calvo Salvador, A. (2017). The pedagogic architecture of MOOC: A research project on educational courses in Spanish. *International Review of Research in Open and Distributed Learning, 18*(6), 18–35. https://doi.org/10.19173/irrodl.v18i6.2964

Gašević, D., Kovanović, V., Joksimović, S., & Siemens, G. (2014). Where is research on massive open online courses headed? A data analysis of the MOOC Research Initiative. *International Review of Research in Open and Distributed Learning, 15*(5), 134–176. https://doi.org/10.19173/irrodl.v15i5.1954

Gillet, D., Ngoc, A. V. N., & Rekik, Y. (2005). Collaborative web-based experimentation in flexible engineering education. *IEEE Transactions on Education, 48*(4), 696–704. https://doi.org/10.1109/te.2005.852592

Graff, H. J. (2015). *Undisciplining knowledge: Interdisciplinarity in the twentieth century*. JHU Press.

Barbera, E.., Zhang, J., Galván-Fernàndez, C., & Fernández-Navarro, F. de A. (2018). Learner support in MOOCs: Identifying variables linked to completion. *Computers & Education, 122*. https://doi.org/10.1016/j.compedu.2018.03.014

Higham, J. (2003). Curriculum change: A study of the implementation of general national vocational qualifications. *The Curriculum Journal, 14*(3), 327–350. https://doi.org/10.1080/0958517032000137649

Holley, K. A. (2009). Interdisciplinary strategies as transformative change in higher education. *Innovative Higher Education, 34*(5), 331–344. https://doi.org/10.1007/s10755-009-9121-4

Holley, K. (2017). Interdisciplinary curriculum and learning in higher education. In G. Noblit (Ed.), *Oxford Research Encyclopedia of Education*. https://doi.org/10.1093/acrefore/9780190264093.013.138

Hubball, H., & Gold, N. (2007). The scholarship of curriculum practice and undergraduate program reform: Integrating theory into practice. *New Directions for Teaching and Learning*, *2007*(112), 5–14. https://doi.org/10.1002/tl.293

Janežič, M., Dimovski, V., & Hodoščček, M. (2018). Modeling a learning organization using a molecular network framework. *Computers & Education, 118*, 56–69. https://doi.org/10.1016/j.compedu.2017.11.008

Jarmon, L., Traphagan, T., Mayrath, M., & Trivedi, A. (2009). Virtual world teaching, experiential learning, and assessment: An interdisciplinary communication course in Second Life. *Computers & Education, 53*(1), 169–182. https://doi.org/10.1016/j.compedu.2009.01.010

Kahn, H. E., & Agnew, M. (2017). Global learning through difference: Considerations for teaching, learning, and the internationalization of higher education. *Journal of Studies in International Education, 21*(1), 52–64. https://doi.org/10.1177/1028315315622022

Kelly, G. J., Luke, A., & Green, J. (2008). What counts as knowledge in educational settings: Disciplinary knowledge, assessment, and curriculum. *Review of Research in Education, 32*(1), vii–x. https://doi.org/10.3102/0091732x07311063

Kockelmans, J. J. (1979). *Why interdisciplinarity. Interdisciplinarity and higher education.* Penn State Press.

Lambiotte, R., & Panzarasa, P. (2009). Communities, knowledge creation, and information diffusion. *Journal of Informetrics, 3*(3), 180–190. https://doi.org/10.1016/j.joi.2009.03.007

Lin, S. (2009, September). Resource transferring mechanism of entrepreneurial network: Based on establishment of the scale-free network. *Proceedings of 2009 International Conference on Management Science and Engineering* (pp. 1638–1643).

Linn, M. C., Clark, D., & Slotta, J. D. (2003). WISE design for knowledge integration. *Science Education, 87*(4), 517–538. https://doi.org/10.1002/sce.10086

Luo, M. M., & Chea, S. (2020). Wiki use for knowledge integration and learning: A three tier conceptualization. *Computers & Education, 154*, 103920. https://doi.org/10.1016/j.compedu.2020.103920

Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E, 69*(2), 026113.

Maassen, P., Nerland, M., & Yates, L. (Eds.). (2018). *Reconfiguring knowledge in higher education* (Vol. 50). Springer.

Mangaroska, K., & Giannakos, M. (2018). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies, 12*(4), 516–534. https://doi.org/10.1109/TLT.2018.2868673

Meyers, E. M., Erickson, I., & Small, R. V. (2013). Digital literacy and informal learning environments: an introduction. *Learning, Media and Technology, 38*(4), 355–367. https://doi.org/10.1080/17439884.2013.783597

Massachusetts Institute of Technology. (2017). *Classes, grades & evaluations*. Retrieved December 12, 2020 from http://web.mit.edu/registrar/subjects/cmtes/coc/

Naidu, S. (2017). *Open educational practice: Caveat emptor*. Globethics.net. http://hdl.handle.net/20.500.12424/164357

Porter, A., & Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics, 81*(3), 719–745. https://doi.org/10.1007/s11192-008-2197-2

Priestley, M. (2011). Whatever happened to curriculum theory? Critical realism and curriculum change. *Pedagogy, Culture & Society, 19*(2), 221–237. https://doi.org/10.1080/14681366.2011.582258

Reagans, R., & McEvily, B. (2003). Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly, 48*(2), 240–267. https://doi.org/10.2307/3556658

Reich, J., & Ruipérez-Valiente, J. A. (2019). The MOOC pivot. *Science, 363*(6423), 130–131. https://doi.org/10.1126/science.aav7958

Rohs, M. & Ganz, M. (2015). MOOCs and the claim of education for all: A disillusion by empirical data. *International Review of Research in Open and Distributed Learning, 16*(6). https://doi.org/10.19173/irrodl.v16i6.2033

Schuwer, R., Gil Jaurena, I., Aydin, C. H., Costello, E., Dalsgaard, C., Brown, M., Jansen, D., & Teixeira, A. (2015). Opportunities and threats of the MOOC movement for higher education: The European perspective. *International Review of Research in Open and Distributed Learning*, *16*(6), 20–38. https://doi.org/10.19173/irrodl.v16i6.2153

Shah, D. (2020, November 30). By the numbers: MOOCs in 2020. *The Report*. Retrieved January 18, 2021 from https://www.classcentral.com/report/mooc-stats-2020/

Shannon, C. E. (2001). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, *5*(1), 3–55. https://doi.org/10.1145/584091.584093

Sheail, P. (2018). Temporal flexibility in the digital university: Full-time, part-time, flexitime. *Distance Education, 39*, 462–479. https://doi.org/10.1080/01587919.2018.1520039

Spector, J. M. (2015). *Foundations of educational technology: Integrative approaches and interdisciplinary perspectives*. Routledge.

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface, 4*(15), 707–719. https://doi.org/10.1098/rsif.2007.0213

Suen, H. K. (2014). Peer assessment for massive open online courses (MOOCs). *International Review of Research in Open and Distributed Learning*, *15*(3), 312–327. https://doi.org/10.19173/irrodl.v15i3.1680

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), e1256. https://doi.org/10.1002/widm.1256

Trowler, P., Knight, P., & Saunders, M. (2003). *Change thinking change practices*. Learning and Teaching Support Network. https://www.academia.edu/10275967/Change_Thinking_Change_Practices

Trowler, P., Saunders, M., & Bamber, V. (Eds.). (2012). *Tribes and territories in the 21st century: Rethinking the significance of disciplines in higher education*. Routledge.

Tsai, Y. S., Rates, D., Moreno-Marcos, P. M., Munoz-Merino, P. J., Jivet, I., Scheffel, M., Drachslet, H., Kloos, C. D., & Gašević, D. (2020). Learning analytics in European higher education—Trends and barriers. *Computers & Education, 155*, 103933. https://doi.org/10.1016/j.compedu.2020.103933

Tucker, R., & Morris, G. (2011). Anytime, anywhere, anyplace: Articulating the meaning of flexible delivery in built environment education. *British Journal of Educational Technology, 42*(6), 904–915. https://doi.org/10.1111/j.1467-8535.2010.01138.x

Veletsianos, G., & Shepherdson, P. (2015). Who studies MOOCs? Interdisciplinarity in MOOC research and its changes over time. *International Review of Research in Open and Distributed Learning*, *16*(3), 1–17. https://doi.org/10.19173/irrodl.v16i3.2202

Venville G., Rennie L. J., & Wallace J. (2012) Curriculum integration: Challenging the assumption of school science as powerful knowledge. In Fraser B., Tobin K., & McRobbie C. (Eds.), *Second International Handbook of Science Education* (Vol. 24, pp. 737–749). Springer. https://doi.org/10.1007/978-1-4020-9041-7_49

West, J. (2018). Teaching data science: an objective approach to curriculum validation. *Computer Science Education, 28*(2), 136–157. https://doi.org/10.1080/08993408.2018.1486120

Wheelahan, L. (2012). *Why knowledge matters in curriculum: A social realist argument*. Routledge.

Zawacki-Richter, O., Bozkurt, A., Alturki, U., & Aldraiweesh, A. (2018). What research says about MOOCs–An explorative content analysis. *The International Review of Research in Open and Distributed Learning*, *19*(1). https://doi.org/10.19173/irrodl.v19i1.3356

Zhang, J., Burgos, D., & Dawson, S. (2019). Advancing open, flexible and distance learning through learning analytics. *Distance Education, 40*(3), 303–308. https://doi.org/10.1080/01587919.2019.1656151

Zhang, J., Lou, X., Zhang, H., & Zhang, J. (2019). Modeling collective attention in online and flexible learning environments. *Distance Education, 40*(2), 278–301. https://doi.org/10.1080/01587919.2019.1600368

Zhang, J., Sziegat, H., Perris, K., & Zhou, C. (2019). More than access: MOOCs and changes in Chinese higher education. *Learning, Media and Technology*, *44*(2), 108–123. https://doi.org/10.1080/17439884.2019.1602541

Athabasca University

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

Effects of Online Self-Regulated Learning on Learning Ineffectiveness in the Context of COVID-19

Wei He[1], Li Zhao[1], and Yu-Sheng Su[2*]
[1]School of Education Science, Nanjing Normal University, China, [2]Department of Computer Science and Engineering, National Taiwan Ocean University, Keelung, Taiwan
[*]The corresponding author is Yu-Sheng Su.

Abstract

Within the COVID-19 pandemic and the new normal period, online learning has become one of the main options for learning. Previous studies on self-regulated learning have shown that it was a better predictor of online learning effectiveness. However, this discussion has not been extended to the situation of the COVID-19 pandemic. To address this gap, this study aims to explore the relationship between the three stages of self-regulated learning (SRL) and learning ineffectiveness (LI). Data of 370 high school students were collected during the period of COVID-19. Structural equation modeling was used to perform confirmatory factor analysis on the data. Findings show that the preparatory stage was positively related to the stages of performance and appraisal, and the performance stage was positively related to the appraisal stage; on the other hand, the stages of performance and appraisal were negatively related to learning ineffectiveness. In addition, the preparatory stage had no direct relation to learning ineffectiveness, but the preparatory stage was correlated with learning ineffectiveness, mediated by the stages of performance and appraisal. These results suggest that better performance in the three stages of self-regulated learning decrease learners' perceived online learning ineffectiveness. This understanding can have implications for global education.

Keywords: online learning, self-regulated learning, learning ineffectiveness, COVID-19

Introduction

COVID-19 has had a destructive impact on the field of society, culture, religion, economy and education all over the world (Mustajab et al., 2020). Offline teaching activities in schools have been suspended and replaced with online education (Zhang et al., 2020). Compared with traditional school-based education, online learning is based on open and distributed learning, without the limitations of place, time, and physical materials. Open and distributed learning gives learners more autonomy in their online self-regulated learning (SRL). Samruayruen et al. (2013) have shown that in an open and distributed education environment, learners' SRL was more successful. SRL is a process that is initiated by learners to control their learning (Tuti et al., 2021). However, online learners seldom interact with or receive guidance and supervision from instructors (Broadbent & Poon, 2015; Su & Wu, 2021), which might result in learners struggling to regulate their learning processes (Jansen et al., 2019). It is therefore important to study learners' online SRL during the period of COVID-19 (Zhu et al., 2020).

Hong et al. (2021) has divided SRL into 6 sub-constructs: task strategy, mood adjustment, self-evaluation, environmental structure, time management, and help-seeking. There are several models of SRL with similar components and processes (Chen & Bonner, 2020). Adam et al. (2017), in their review, conclude that previous SRL models comprised the three stages: preparatory, performance, and appraisal. Many researchers have discussed the effects of multiple components of SRL or a single stage on other factors. For instance, the relationships between learning environments, students' beliefs, and multiple dimensions of SRL were explored by Maison and Syamsurizal (2019). Cosnefroy et al. (2018) analyzed the correlation between the forethought stage of SRL and self-regulation failure. Nevertheless, Zeidner and Stoeger (2019) indicate that few studies have considered all stages of SRL simultaneously. However, Liu et al. (2021) discussed the gender difference in each of the three stages of online SRL. The results found that in each of the three stages of SRL female students performed better than male students. Hong et al. (2021) examined the impact of academic procrastination on each of the six sub-constructs of SRL, and each of the six sub-constructs of SRL on learning ineffectiveness (LI). Thus, this study aims to explore the impact of the three stages of SRL: preparatory, performance, and appraisal, not the six sub-constructs of SRL, on LI. According to the effects of the three stages of SRL on learning effectiveness or ineffectiveness, instructors can provide targeted and efficient support for students.

Benefiting from the openness and distribution of online learning, students' online learning effectiveness has been improved accordingly. For example, Zhao, Liu, and Su (2021) have shown students to demonstrate better learning performance in open and distributed education than in face-to-face learning. Students with better ability to self-regulate their online learning were found to have significantly higher levels of perceived effectiveness than those with less ability in this area (Charo et al., 2020). When engaged in online learning, if students lack SRL skills, they may not be able to complete the learning tasks they are assigned in their online courses (Barnard et al., 2009). The abovementioned studies show that students' SRL reduces their learning ineffectiveness (LI) level. However, in the context of COVID-19, all the offline learning suddenly changed to 100% online learning. This study is to explore whether students' SRL was effective and how the different SRL stages affected students' learning effectiveness during this transformation. When students study online, it is necessary and significant to grasp their perceptions of online learning effectiveness or ineffectiveness (Hong et al., 2021). In this study, LI was adopted for high school students to self-rate their perceived learning performance. Therefore, this study explores how the three stages of SRL were related to

LI while high school students were learning online. The findings of the relationships between online SRL and online LI can provide new insights into distance education and provide relevant references for coping with the future online learning research on the normalization of the epidemic.

Literature Review

Online Self-Regulated Learning

Before COVID-19, most students studied face to face in classrooms and did not experience 100% online learning. Prior studies on SRL were conducted in online or offline learning contexts, but little research has been looked at during COVID-19. Students were suddenly faced with the extremely difficult task of self-regulating their learning activities at home amid the influence of the COVID-19 pandemic (Zhang et al., 2021). The shift from offline to online learning during COVID-19 has caused students to lack instructors' guidance, requiring them to have a greater ability to regulate themselves in their learning (Lee et al., 2020). During the specific time of COVID-19, many factors might have multiple negative effects on learners' SRL processes (Cai et al., 2020). SRL is an important capability to actively participate in constructing and interpreting knowledge in a student-centered learning environment (Alsancak Sirakaya & Ozdemir, 2018). "SRL is an active and constructive process in which learners set their own learning goals and then attempt to regulate, plan and control their motivation, cognition, and behavior" (Pintrich, 2000, p.453). During this process, they are both guided and limited by their goals and the environmental background characteristics (Pintrich, 2000.p.453). Learning tasks must have clear beginnings, middles, and ends (Cleary et al., 2012). In online learning courses, there is a clear learning process for before, during, and after lessons.

Several SRL models presenting different stages and subprocesses have been proposed. For example, based on social cognitive theory, Zimmerman (2000) described an SRL model as comprising the forethought, performance, and self-reflection stages. Hadwin et al. (2018) has developed a self-regulation model and divided it into the three components of negotiation and awareness of the task, strategic task engagement, and adaptation. Adam et al. (2017) proposed SRL comprises the stages of preparatory, performance, and appraisal. Thus, the present study selected the three stages which correspond to behaviors of before, during, and after online lessons, as the cyclic processes during the COVID-19 pandemic. During each stage, students use different strategies to monitor and control their learning (Zimmerman, 2000).

In the preparatory stage, the learning environment (e.g., stable Internet connection) and individual characteristics (e.g., mood) have been highlighted as essential components by Hong et al. (2021). Thus, this study specified the preparatory behaviors before engaging in online lessons focusing on mood adjustment and structuring environments. In addition, Adam et al. (2017) stated that the performance stage of SRL is when the actual task is accomplished while monitoring and controlling the progress of performance. In the performance stage, learners use cognitive and certain strategies (e.g., task strategies) and meta-cognitive monitoring processes (e.g., time management) to accomplish tasks (Ridgley et al., 2020; Zhang et al., 2021). Thus, this study specified the performance behaviors during online lessons, from two aspects of time management and the task strategy, When learners completed the learning tasks, they enter the appraisal stage, during which they monitor their learning progress, design help-seeking plan (Zimmerman, 2000), evaluate learning effectiveness (Cleary et al., 2012; Zimmerman, 1990). Considering this, this study specified the appraisal behaviors after engaging in online lessons, from two aspects of help seeking and self-evaluation.

Many studies of online learning have shown a relationship between learning achievement and the subscales of SRL, such as help seeking (Won et al., 2021), and learning environments (Maison & Syamsurizal, 2019). However, most studies did not use all stages of the SRL model (Zeidner & Stoeger, 2019). To improve learners' learning and SRL skills, differential effects of SRL in models and theories should be applied by scholars and teachers (Ernesto, 2017). Therefore, this study focuses on three stages of SRL during the COVID-19 pandemic.

Learning Ineffectiveness (LI) in the Context of COVID-19

With the help of distance education and technology support, students have easy and convenient access to online learning. Online learning effectiveness can be reflected by learners' evaluation of the performance in the cognitive, affective, and psychomotor domains of online learning (Zhao, He, & Su, 2021). The online learning carried out during the COVID-19 pandemic has been effective (Bahasoan et al., 2020). Although learners can learn and benefit from online learning, the learning effectiveness of online learning compared with traditional learning is still considered a debatable issue. For example, Zhao, Liu, and Su (2021) show that compared with traditional learning, flipping classroom learning supported by MOOCs results in better learning achievement; however, Carrol and Burke (2010) support that traditional learning was better than online learning. To ensure effective online learning, teachers and course designers must understand learners' perceptions of online learning effectiveness or ineffectiveness (Hong et al., 2021).

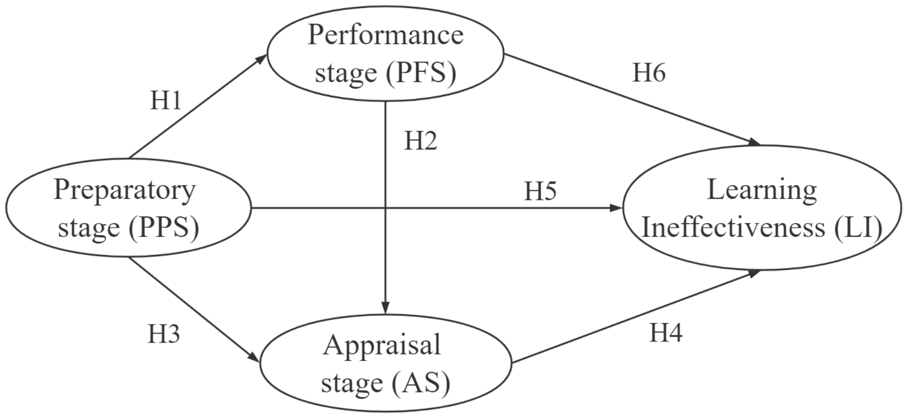
Adolescents are likely to be biased in their tendency to react (van Herk et al., 2004). For example, facing difficulties, learners may feel dissatisfied with engaging in online courses (Rabin et al., 2020). Ruhland and Brewer (2001) argue that students' perceived ineffectiveness of online learning is also a factor that should be captured as part of learning outcomes. Therefore, a good measurement of learning effectiveness requires a considerable understanding of how to best link the course to online learning and how to make online learning meaningful for students' needs and experiences. Hong et al. (2021) originally proposed LI, defined as learners self-evaluating how they feel about their online learning performance. However, limited research exists on LI related to learners' online learning. Thus, this study aimed to investigate students' perceptions of LI during the COVID-19 pandemic.

Research Hypotheses

Extensive research has been done on the impact of SRL on learners' learning and academic achievement (Jansen et al., 2019). Previous studies have found that SRL was a good predictor of academic achievement (e.g., Moghadari-Koosha et al., 2020). Six SRL sub-constructs influenced perceived LI (Hong et al., 2021). In online learning courses, there is a clear learning process of before, during, and after lessons. These three steps in the learning process correspond to the three stages of the SRL process model (Adam et al., 2017). This study furthers the previous study to explore the relationship between the stages of SRL and LI in the context of online learning. Therefore, based on Adam et al.'s (2017) proposed SRL model, we developed the conceptual model shown in Figure 1 for this study. This model reveals the relationship between learners' SRL behaviors at various stages of SRL and their respective influences on their LI.

Figure 1

Research Model



According to Zimmerman (2015), SRL is a cyclical process whereby learners are engaged in three distinct stages. Boom et al. (2004) revealed the self-regulated learning competence map, showing that the learning process includes beginning, performing, and finishing. The beginning stage is directed to the performing stage, and the performing stage is directed to the finishing stage. The emotion regulation strategy is associated with adaptive strategies, such as reappraisal (Aldao et al., 2010). Mood adjustment is considered as an essential component of the preparatory stage. Thus, the following hypotheses were proposed:

> H1: The preparatory stage (PPS) is positively correlated with the performance stage (PFS) in online learning.
>
> H2: PFS is positively correlated with the appraisal stage (AS) in online learning.
>
> H3: PPS is positively correlated with AS in online learning.

Academic achievement has significant relationships with the behaviors of the preparatory (e.g., Lehmann et al., 2014), performance (e.g., Alghamdi et al., 2020), and appraisal stages (e.g., Colthorpe et al., 2019). According to Adam et al.'s (2017) SRL process model, this study divided the six constructs of Hong et al.'s (2021) SRL into the preparatory, performance, and appraisal stages. Hong et al. (2021) found all of the six sub-constructs of SRL were negatively correlated with LI. Thus, the interaction effects between the three stages of SRL and LI were hypothesized as follows:

> H4: AS is negatively correlated with students' LI in online learning.
>
> H5: PPS is negatively correlated with students' LI in online learning.
>
> H6: PFS is negatively correlated with students' LI in online learning.

Methodology

Participants and Procedure

High school students have to face the college entrance examination, which is a concern for students all around the world, and especially in China. Compared with other levels of education, high school is more intense, and the online learning of high school students has received substantial attention from society during the COVID-19 pandemic. Therefore, we selected high school students at various grade levels in Jiangsu Province, China, as participants for this study. Adapting purposive sampling, teachers who gave online courses were invited to distribute the questionnaire to their students between April 10 and April 20, 2020. All participants were informed that the online questionnaire would be used only for this study and that their privacy would be protected. A total of 395 students from the high schools voluntarily and anonymously completed the online survey. If the questionnaires have missing values needed for the data analysis, they would be removed, leaving 370 samples for analysis.

The participants from grades 1 to 3 (M = 2.14, SD = 1.140) included 75 males (20.3%) and 295 females (79.7%), whose ages were from 15 to 21 years (M = 16.85, SD = 1.156). In addition, all participants had taken part in online lessons. Of all participants, the average study hours per day was 2.40 (SD = 0.847), and that semester's online courses number was between two and nine (M = 4.72, SD = 1.190). Of all participants, 96% studied online courses for 50% of the time during that semester.

Instruments

The questionnaire items were adapted from prior studies and were translated into Chinese by experts. Three high school students were invited to check the whole questionnaire and give comments to all the items to ensure the readability of the measurement items. Each of the items was scored by a 5-point Likert scale, from 1 for strongly disagree to 5 for strongly agree, with 3 representing neutral. Finally, the reliability of the constructs was subsequently tested.

Online SRL of Measurement

According to the SRL instrument of Hong et al. (2021), we designed the scale of the instrument with 22 items consisting of six sub-constructs with good reliability and validity, covering the three stages. The preparatory stage includes mood management, and environment structuring; the performance stage includes adapting time management and task strategies; and the appraisal stage includes help seeking and self-evaluation. The preparatory stage contains 8 items, such as, "Before I study online, I am used to finishing the coursework to avoid distractions in the online class." The performance stage contains 7 items, such as "During learning online, I will adjust my learning style according to the actual learning." The appraisal stage contains 7 items, for example, "After learning online, I test and summarize what I have learned."

Learning Ineffectiveness of Online Learning Measurement

A good learning effectiveness measurement must capture changes in learners' cognitive and affective development as a result of their learning experiences. Previous studies took ineffectiveness instead of effectiveness to assess learners' online learning performance. For example, the scale of learning ineffectiveness in the online learning context was developed to measure college students' LI in the context of COVID-19 (Hong et al., 2021). Therefore, eight items were designed in this study to measure the online learning ineffectiveness of high school students, for example, "Since learning online, my learning confidence has decreased."

Data Analysis

According to Thompson's (2000) recommendation, the number of samples should be between 10:1 and 15:1 for the number of observed variables. This ratio of sample size (N = 370) to observed variables (30 items) is reasonable. IBM SPSS Statistics was used to analyze data from all 370 high school students. Next, descriptive statistics of population information and correlation analysis were obtained using SPSS 24. Then confirmatory factor analysis (CFA) was performed to further test whether the questionnaire satisfied the reliability and validity via Amos (version 22.0). Finally, we conducted structural equation modeling (SEM) to evaluate the hypothetical structural model.

Results

Reliability and Validity Analysis

First, items with a value of factor loadings lower than .50 were deleted in each construct (Hair et al., 2011). During this process, three items in PPS, one item in PFS, and two in AS with factor loadings lower than .50 were deleted. After conducting CFA, items with the highest residual value in each construct were deleted (Hair et al., 2019). During this process, two items in PPS and two in PFS with the highest residual values were deleted. To meet the criteria, some items in each construct needed to be removed: one item in AS and three in LI were removed. The measurement model finally exhibited a good fit, with chi-square divided by the degrees of freedom ($\chi 2/df$) = 2.482, goodness of fit index (GFI) = .925, Bentler–Bonett normed fit index (NFI) = .959, comparative fit index (CFI) = .975, and root mean square error of approximation (RMSEA) = .063. The remaining 16 items—which contained three PPS items, four PFS items, four AS items, and five LI items—were reserved for further analysis.

Second, composite reliability (CR) and Cronbach's alpha (α) were considered together to assess the internal model's consistency. Hair et al. (2019) suggest that the CR should exceed .70. DeVellis (2012) recommends that an acceptable α value should be above .70. Thus, a construct is considered to have achieved internal consistency when both the CR and α exceed .70. Table 1 shows that the CR of all constructs ranged from .862 to .939, and α ranged from .721 to .900. Therefore, the results suggest that each construct measurement variable in the questionnaire had acceptable reliability and internal consistency.

Third, we calculated the construct's average variance extracted (AVE) and the variable measurement condition factor to ascertain the convergent validity. When the convergent effectiveness of construct is sufficient, the value of AVE should exceed .50 (Fornell & Larcker, 1981). Additionally, the convergent validity requirement is satisfied if the variable's measurement factor is greater than .50 (Hair et al., 2019). Table 1 indicates that the AVE of all constructs exceeded .50 (ranging from .677 to .756), and each item's standardized factor loading also exceeded .50 (ranging from .696 to .937). Therefore, the questionnaire had acceptable convergent validity.

Table 1

Reliability and Validity Analysis

Latent variable	Measure item	Standardized factor loading	CR	AVE	Cronbach's α
Preparatory stage (PPS)	PPS1	.748	.862	.677	.721
	PPS2	.888			
	PPS3	.826			
Performance stage (PFS)	PFS1	.897	.921	.747	.819
	PFS2	.921			
	PFS3	.922			
	PFS4	.696			
Appraisal stage (AS)	AS1	.781	.913	.725	.839
	AS2	.803			
	AS3	.907			
	AS4	.906			
Learning ineffectiveness (LI)	LI1	.743	.939	.756	.900
	LI2	.903			
	LI3	.921			
	LI4	.937			
	LI5	.828			

Note. CR = composite reliability; AVE = average variance extracted.

Model Fit Analysis

The model fit and statistical significance of the hypothesized paths between the four potential variables were examined to test the structural model. Kline (2011) suggests that GFI, NFI, and CFI values exceeding .90, a $\chi 2/df$ value less than 3, and an RMSEA value less than .08 can generally be regarded as representing acceptable goodness of fit. The results show that the data ($\chi 2/df$ = 2.466, GFI = .902, NFI = .904, CFI = .940, RMSEA = .077) had an acceptable fit of the hypothesized model. It indicated that the hypothesis model proposed in this study has good fitness.

Path Analysis

The standardized path coefficients (β) of the model of the study are represented in Figure 2 and Table 2. The results indicate that hypotheses 1, 2, 3, 4, and 6 were supported. PPS was positively related to PFS and AS ($\beta = .808$, $t = 15.695$; and $\beta = .325$, $t = 5.357$, respectively). PFS was positively related to AS ($\beta = .636$, $t = 9.936$). Moreover, PFS and AS were negatively related to LI ($\beta = -.453$, $t = -4.865$; and $\beta = -.365$, $t = -3.495$, respectively). However, PPS was not significantly related to LI ($\beta = -.077$, $t = -1.062$). These results indicate that H5 was not supported.

Table 2

Coefficients of the Hypothesized Model

Hypothesis	Path	β	*SE*	*t*	Supported
H1	PPS→PFS	.808	.060	15.695*	Yes
H2	PFS→AS	.636	.062	9.936*	Yes
H3	PPS→AS	.325	.068	5.357*	Yes
H4	AS→LI	−.365	.096	−3.495*	Yes
H5	PPS→LI	−.077	.075	-1.062	No
H6	PFS→LI	−.453	.082	−4.865*	Yes

Note. β = standardized coefficient; H = hypothesis; PPS = preparatory stage; PFS = performance stage; AS = appraisal stage; LI = learning ineffectiveness.

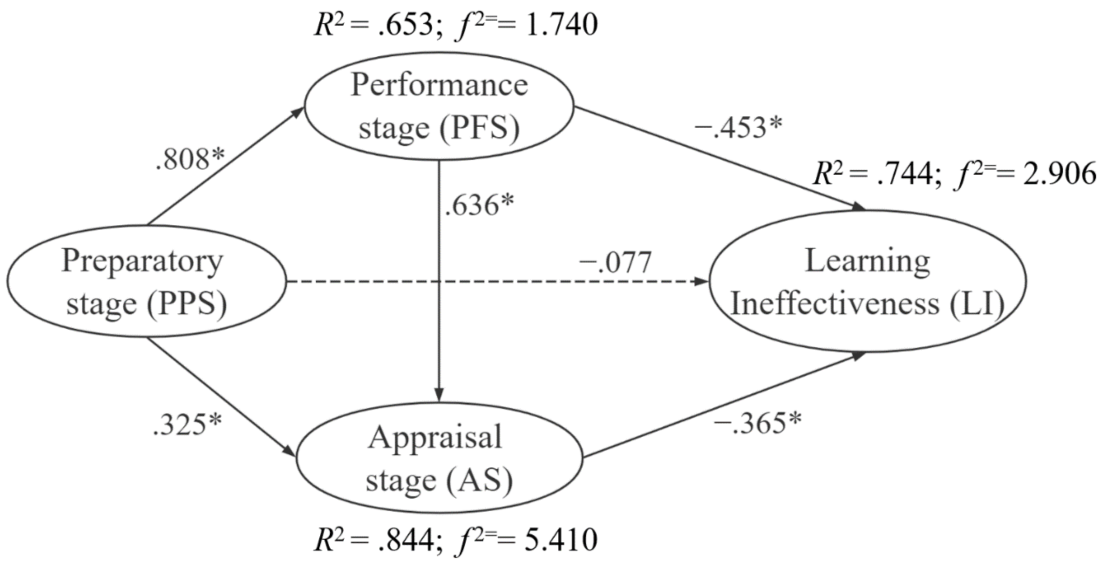
* $p < .001$.

The coefficient of determination (R^2) represents the predictive ability of the model (Fornell & Larcker, 1981), and R^2 values higher than .6 are considered to indicate a high impact effect (Sanchez, 2013). The explanatory power of PPS to PFS was 65.3%, the explanatory power of PPS and PFS to AS was 84.4%, and the explanatory power of PFS and AS to LI was 74.4%. Therefore, all variables had good predictive capacity (Hair et al., 2012).

In addition, f^2 values greater than .8 are considered to have a high effect size, between .2 and .8 are considered medium, and less than .2, small (Cohen, 1988). As shown in Figure 2, the f^2 ranged from 1.740 to 5.410, indicating that the effect size was good. Therefore, the paths between the variables in this study were well verified (Hair et al., 2019).

Figure 2

The Structural Model with Standardized Coefficients



Note. * p < .001.

Finally, 5,000 resample bootstrappings were performed to provide additional evidence related to the significance of the indirect effects. The bootstrapping 95% confidence interval (CI) of the lower and upper bounds of indirect effects did not include zero, indicating that the paths were significant (Preacher et al., 2007). It was significant for the mediated effect of the study model (β = −.672) with 95% CI from −.798 to −.565, indicating that PFS and AS of SRL did have a full mediating effect on the negative correlation between PPS and online LI.

Discussion

The COVID-19 outbreak has caused a growing number of students to adopt online learning, but the effectiveness of online learning is a controversial issue. Research on SRL has been largely from a macro perspective, for example, learning behavior, learning ability, and academic performance. However, there is still a lack of micro perspectives on the interaction mechanism among the various stages of SRL, and it is not clear which stage of behavior has the strongest effect on learning effectiveness. Adopting a micro perspective, this study focused on exploring how stages of learners' SRL behavior affected their perceptions of learning ineffectiveness. The results show that SRL behaviors in the stage of preparatory had a positive effect on the stages of performance and appraisal, and the performance stage had a positive influence on the appraisal stage. We also found that the preparatory stage of SRL affects learning ineffectiveness by mediating the effect of the performance and appraisal stages.

Previous studies indicate that each process in the preparatory stage initiates actions that the learner engages in when performing the task (Ridgley et al., 2020). For example, mood actively activates pre-reflection in

SRL (Lehmann et al., 2014), which initiates actions of the performance stage. In this study, the behaviors in the preparatory stage had a direct positive impact on the performance stage, indicating that if the preparation during the preparatory stage is sufficient, the performance stage process will be easier. Thus, H1 was positively supported.

The use of strategy and meta-cognitive monitoring in the performance stage subsequently influence the appraisal stage, reflect on and evaluate their progress and goal attainment (Ridgley et al., 2020). For example, time management is related to evaluation, reflection, and reaction (Wolters & Brady, 2020). The results show that behaviors in the performance stage had a positive effect on the appraisal stage, indicating that learners would perform better in the appraisal stage according to the adopted task strategies and actively monitor the length of time in the performance stage, positively supporting H2.

Learners need to manage environmental factors such as computer access at home before studying (Cai et al., 2020). Emotion regulation strategies were correlated with reappraisal (Aldao et al., 2010). Pekrun et al. (2011) have proposed that positive emotions may be beneficial in most cases. This could indicate that learners will perform better in the appraisal stage according to their mood adjustment and environment structuring during the preparatory stage. The results of the present research verify that the preparatory stage is positively related to the appraisal stage, supporting H3.

Tzeng and Nieh (2015) state that in the appraisal stage, self-evaluations and self-reactions led learners to feel that their learning was effective and motivated them to continue to work diligently because they believed they could make further progress. Moreover, Zhu et al. (2011) found that learners who were developing help-seeking schemes, such as searching for help on the Internet, were more likely to have good academic performance. By investing more energy in self-evaluation and help seeking after online courses, learners increase their learning effectiveness. The results of the present research verify that the appraisal stage can negatively predict perceived learning ineffectiveness, negatively supporting H4.

Cosnefroy et al. (2018), constructing a self-regulated learning failure model, shows that forethought processes affect academic performance by affecting the performance stage. The actual situation (e.g., noise) and individual characteristics (e.g., mood) influence learning outcomes (Lehmann et al., 2014). Based on the studies mentioned, this research considered that higher SRL when regulating mood and preparing the environment for distance learning can promote learners' behavior within the performance and appraisal stages and reduce learning ineffectiveness. Although the preparatory stage did not show a direct influence on learning ineffectiveness in this study, the preparatory stage was correlated with learners' learning ineffectiveness by mediating the effect of the other two stages. Therefore, H5 was not supported.

The results show that the performance stage of SRL has a high indirect effect on learning ineffectiveness. If they consider a variety of factors such as task strategy and time management during online courses, learners can reduce their learning ineffectiveness; some previous studies (e.g., Alghamdi et al., 2020; Wolters & Brady, 2020) report similar results. Their research shows that task strategy and time management can have positive impacts on academic performance and achievement, respectively. Thus, H6 was negatively supported.

Conclusions

During the outbreak of COVID-19, online learning was comprehensively applied in education. Ways to promote online learning effectiveness in the context of COVID-19 is an important issue. The online learning environment demands learner-centeredness and self-regulation. Self-regulated learning plays a crucial role in online learning. This study divided SRL into three stages and explored the relationship of high school students' SRL from the three stages and learning ineffectiveness. Results indicate that SRL has a predictive effect on learning effectiveness, and high SRL levels can reduce the ineffectiveness of online learning.

Implications

The COVID-19 pandemic has led to the closure of schools around the world, and offline learning has been replaced with distance learning. This study has some implications for online learning in distance education. Epidemic prevention and treatment are moving the world toward normalization. Learning in the post-pandemic era must integrate online and offline learning and maximize students' learning (Mei, 2020), highlighting the importance of online learning and SRL. The exploration of students' online SRL is conducive to understanding the current situation of students' online SRL and points to further improving it. This study has certain reference value for coping with future online learning to deal with such emergencies, which may occur anywhere in the world.

The theoretical significance of the present research is to clarify the impacts of SRL on learning ineffectiveness during COVID-19. This study is also to provide a practical contribution, which is the results show that the preparatory stage of SRL through the performance and appraisal stages affects learners' online learning ineffectiveness. SRL interventions effectively improved learners' SRL, performance, and academic achievement (e.g., Jansen et al., 2019). The results of this study can be applied by high school teachers to enhance students' adaptability in SRL situations by implementing different interventions before, during, and after lessons.

Limitations and Future Study

Several limitations should be acknowledged in this study. First, the sample size was small, and all participants were from the Jiangsu Province, so it is hard to make sure that the sample represents high school education institutions at all levels. Thus, the sample cannot represent all Chinese high school students. Future studies need to collect more and larger representative samples to enhance the conclusions of the study.

Second, the population of the present study was almost 80% female, which may have led to the distribution bias of the results. Gender difference is a potentially important factor affecting SRL and learning performance (Bezzina, 2010). Future studies may explore the role of gender, the three stages of SRL, and learning ineffectiveness.

Moreover, it is increasingly important to explore predictors of online learning success as online courses are becoming more flexible and accessible (Broadbent & Poon, 2015; Su, Ding, & Chen, 2021). Other factors not examined in this study, such as self-efficacy, self-direction, learning motivation, and learning satisfaction, may also affect students' perceived ineffectiveness of online learning. Researchers might consider including other factors that may affect the perceived ineffectiveness of online learning in future studies.

Acknowledgements

This study was supported in part by a grant from National Social Science Foundation (No. BCA200093) and the Priority Academic Program Development of Jiangsu Higher Education Institutions in China. Moreover, this study was supported by the Ministry of Science and Technology, Taiwan, under grant MOST 109-2511-H-019-004-MY2 and MOST 109-2511-H-019-001.

References

Adam, N. L., Alzahri, F. B., Soh, S. C., Bakar, N. A., & Kamal, N. (2017). Self-regulated learning and online learning: A systematic review. In H. Badioze Zaman, P. Robinson, A. F. Smeaton, T. K. Shih, S. Velastin, T. Terutoshi, A. Jaafar, & N. Mohamad Ali (Eds.), *Advances in visual informatics: 5th International Visual Informatics Conference, IVIC 2017, Bangi, Malaysia, November 28–30, 2017, Proceedings* (pp. 143–154). https://doi.org/10.1007/978-3-319-70010-6_14

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review-science direct. *Clinical Psychology Review*, *30*(2), 217–237. https://doi.org/10.1016/j.cpr.2009.11.004

Alghamdi, A., Karpinski, A. C., Lepp, A., & Barkley, J. (2020). Online and face-to-face classroom multitasking and academic performance: Moderated mediation with self-efficacy for self-regulated learning and gender. *Computers in Human Behavior*, *102*, 214–222. https://doi.org/10.1016/j.chb.2019.08.018

Alsancak Sirakaya, D., & Ozdemir, S. (2018). The effect of a flipped classroom model on academic achievement, self-directed learning readiness, motivation and retention. *Malaysian Online Journal of Educational Technology*, *6*(1), 76–91. https://files.eric.ed.gov/fulltext/EJ1165484.pdf

Bahasoan, A. N., Ayuandiani, W., Mukhram, M., & Rahmat, A. (2020). Effectiveness of online learning in pandemic COVID-19. *International Journal of Science, Technology & Management*, *1*(2), 100–106. https://doi.org/10.46729/ijstm.v1i2.30

Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., & Lai, S. L. (2009). Measuring self-regulation in online and blended learning environments. *The Internet & Higher Education*, *12*(1), 1–6. https://doi.org/10.1016/j.iheduc.2008.10.005

Bezzina, F. H. (2010). Investigating gender differences in mathematics performance and in self-regulated learning: An empirical study from Malta. *Equality, Diversity and Inclusion*, *29*(7), 669–693. https://doi.org/10.1108/02610151011074407

Boom, G., Paas, F., Merrieenboer, J. J. G. V., & Gog, T. V. (2004). Reflection prompts and tutor feedback in a Web-based learning environment: Effects on students' self-regulated learning competence. *Computers in Human Behavior*, *20*(4), 551–567. https://doi.org/10.1016/j.chb.2003.10.001

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, *27*, 1–13. http://doi.org/10.1016/j.iheduc.2015.04.007

Cai, R., Wang, Q., Xu, J. & Zhou, L. (2020). Effectiveness of students' self-regulated learning during the COVID-19 pandemic. *Social Science Electronic Publishing*, *34*(1), 175–182. https://ssrn.com/abstract=3622569

Carrol, N., & Burke, M. (2010). Learning effectiveness using different teaching modalities. *Journal of Business & Economics Research*, *3*(12), 65–72. http://doi.org/10.19030/ajbe.v3i12.966

Charo, R., Maite, A. S., & Guillermo, M. (2020). Self-regulation of learning and MOOC retention. *Computers in Human Behavior*, *111*, Article 106423. https://doi.org/10.1016/j.chb.2020.106423

Chen, P., & Bonner, S. (2020). A framework for classroom assessment, learning, and self-regulation. *Assessment in Education Principles Policy and Practice*, *27*(4), 373–393. https://doi.org/10.1080/0969594X.2019.1619515

Cleary, T. J., Callan, G. L., & Zimmerman, B. J. (2012). Assessing self-regulation as a cyclical, context-specific phenomenon: An overview and analysis of SRL microanalysis protocols. *Education Research International*, *2012*, 1–19. https://doi.org/10.1155/2012/428639

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. http://www.utstat.toronto.edu/~brunner/oldclass/378f16/readings/CohenPower.pdf

Colthorpe, K., Ogiji, J., Ainscough, L., Zimbardi, K., & Anderson, S. (2019). Effect of metacognitive prompts on undergraduate pharmacy students' self-regulated learning behavior. *American Journal of Pharmaceutical Education*, *83*, Article 6646. https://doi.org/10.5688/ajpe6646

Cosnefroy, L., Fe Nouillet, F., Corinne Mazé, & Bonnefoy, B. (2018). On the relationship between the forethought phase of self-regulated learning and self-regulation failure. *Issues in Educational Research*, *28*(2), 329–348. http://www.iier.org.au/iier28/cosnefroy.pdf

DeVellis, R. F., (2012). *Scale development: Theory and applications* (3rd ed.). Sage Publications.

Ernesto, P. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*, Article 422. https://doi.org/10.3389/fpsyg.2017.00422

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/10.1177/002224378101800104

Hadwin, A. F., Järvelä, S., and Miller, M. (2018). Self-regulation, co-regulation and shared regulation in collaborative learning environments. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 83–106). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315697048-6/self-regulation-co-regulation-shared-regulation-collaborative-learning-environments-allyson-hadwin-sanna-j%C3%A4rvel%C3%A4-mariel-miller

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage. https://www.cengage.co.uk/books/9781473756540/

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–152. https://doi.org/10.2753/MTP1069-6679190202

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science*, *40*(3), 414–433. https://doi.org/10.1007/s11747-011-0261-6

Hong, J. C., Lee, Y. F., & Ye, J. H. (2021). Procrastination predicts online self-regulated learning and online learning ineffectiveness during the coronavirus lockdown. *Personality and Individual Differences*, *174*(7), Article 110673. https://doi.org/10.1016/j.paid.2021.110673

Jansen, R. S., van Leeuwen, A., Janssen, J., Conijn, R., & Kester, L. (2019). Supporting learners' self-regulated learning in massive open online courses. *Computers & Education*, *146*, Article 103771. https://doi.org/10.1016/j.compedu.2019.103771

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. The Guilford Press. https://www.guilford.com/books/Principles-and-Practice-of-Structural-Equation-Modeling/Rex-Kline/9781462523344

Lee, D., Watson, S. L., & Watson, W. R. (2020). The relationships between self-efficacy, task value, and self-regulated learning strategies in massive open online courses. *International Review of Research in Open and Distance Learning*, *21*(1), 23–39. https://doi.org/10.19173/irrodl.v20i5.4389

Lehmann, T., Haehnlein, I., & Ifenthaler, D. (2014). Cognitive, metacognitive and motivational perspectives on preflection in self-regulated online learning. *Computers in Human Behavior*, *32*, 313–323. https://doi.org/10.1016/j.chb.2013.07.051

Liu, X., He, W., Zhao, L., & Hong, J. C. (2021). Gender differences in self-regulated online learning during the COVID-19 lockdown. *Frontiers in Psychology*, *12*, Article 752131. https://doi.org/10.3389/fpsyg.2021.752131

Maison, S., & Syamsurizal, T. (2019). Learning environment, students' beliefs, and self-regulation in learning physics: Structural equation modeling. *Journal of Baltic Science Education*, *18*(3), 389–403. https://doi.org/10.33225/jbse/19.18.389

Mei, D. (2020, January). Research on multi-integrated online and offline teaching model in the post-epidemic era: Taking numerical analysis course as an example. In *Proceedings of the 2020 5th International Conference on Modern Management and Education Technology (MMET 2020)* (pp. 356–359). Atlantis Press. https://doi.org/10.2991/assehr.k.201023.072

Moghadari-Koosha, M., Moghadasi-Amiri, M., Cheraghi, F., Mozafari, H., & Zandieh, M. (2020). Self-efficacy, self-regulated learning, and motivation as factors influencing academic achievement among paramedical students: A correlation study. *Journal of Allied Health*, *49*(3), e145–e152. https://pubmed.ncbi.nlm.nih.gov/32877483/

Mustajab, M., Baharun, H. & Fawa'Iedah, Z. (2020). Adapting to teaching and learning during COVID-19: A case of Islamic school's initiative of self-regulated learning. *Nadwa*, *14*(2), 241–264. https://doi.org/10.21580/nw.2020.14.2.6515

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The achievement emotions questionnaire (AEQ). *Contemporary Educational Psychology*, *36*(1), 36–48. https://doi.org/10.1016/j.cedpsych.2010.10.002

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation: Theory, research, and applications* (pp. 451–502). Academic Press. https://doi.org/10.1016/B978-012109890-2/50043-3

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, *42*(1), 185–227. https://doi.org/10.1080/00273170701341316

Rabin, E., Henderikx, M., Kalman, Y. M., & Kalz, M. (2020). What are the barriers to learners' satisfaction in MOOCs and what predicts them? The role of age, intention, self-regulation, self-efficacy and motivation. *Australasian Journal of Educational Technology*, *36*(3), 119–131. https://doi.org/10.14742/ajet.5919

Ridgley, L. M., Rubenstein, D., & Callan, G. L. (2020). Gifted underachievement within a self-regulated learning framework: Proposing a task-dependent model to guide early identification and intervention. *Psychology in the Schools*, *57*(3), 1365–1384. https://doi.org/10.1002/pits.22408

Ruhland, S. K., & Brewer, J. A. (2001). Implementing an assessment plan to document student learning in a two-year technical college. *Journal of Vocational Education Research*, *26*, 141–171. https://doi.org/10.5328/JVER26.2.141

Samruayruen, B., Enriquez, J., Natakuatoong, O., & Samruayruen, K. (2013). Self-regulated learning: A key of a successful learner in online learning environments in Thailand. *Journal of Educational Computing Research*, *48*(1), 45–69. https://doi.org/10.2190/EC.48.1.c

Sanchez, G.. (2013). *PLS path modeling with R*. Trowchez Editions. https://www.gastonsanchez.com/PLS_Path_Modeling_with_R.pdf

Su, Y. S., Ding, T. J., & Chen, M. Y. (2021). Deep learning methods in internet of medical things for valvular heart disease screening system. *IEEE Internet of Things Journal, 8*(23), 16921-16932. https://doi.org/ 10.1109/JIOT.2021.3053420

Su, Y. S., & Wu, S. Y. (2021). Applying data mining techniques to explore users behaviors and viewing video patterns in converged it environments. *Journal of Ambient Intelligence and Humanized Computing*. https://doi.org/10.1007/s12652-020-02712-6

Thompson, B. (2000). Ten commandments of structural equation modeling. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 261–283). American Psychological Association.

Tuti, T., Paton, C., & Winters, N. (2021). The counterintuitive self-regulated learning behaviours of healthcare providers from low-income settings. *Computers & Education*, *166*(1), Article 104136. https://doi.org/10.1016/j.compedu.2021.104136

Tzeng, S., & Nieh, H. (2015, September). How self-concept, self-efficacy and self-evaluation relate to relate to achievement outcomes: New technology-based learning models for science and technology universities students. In *Proceedings of 2015 International Conference on Interactive Collaborative Learning (ICL), 20–24 September 2015, Firenze, Italy* (pp. 863–870). https://doi.org/10.1109/ICL.2015.7318141

van Herk, H., Poortinga, Y. H., & Verhallen, T. M. M. (2004). Response styles in rating scales: Evidence of method bias in data from six EU countries. *Journal of Cross-Cultural Psychology*, *35*, 346–360. https://doi.org/10.1177/0022022104264126

Wolters, C. A., & Brady, A. C. (2020). College students' time management: A self-regulated learning perspective. *Educational Psychology Review*. Advance online publication. https://doi.org/10.1007/s10648-020-09519-z

Won, S., Hensley, L. C., & Wolters, C. A. (2021). Brief research report: Sense of belonging and academic help-seeking as self-regulated learning. *The Journal of Experimental Education*, *89*(1), 1–13. https://doi.org/10.1080/00220973.2019.1703095

Zeidner, M., & Stoeger, H. (2019). Self-regulated learning (SRL): A guide for the perplexed. *High Ability Studies*, *30*(1–2), 9–51. https://doi.org/10.1080/13598139.2019.1589369

Zhang, T., Taub, M., & Chen, Z. (2021, April). Measuring the impact of COVID-19 induced campus closure on student self-regulated learning in physics online learning modules. In *LAK21: 11th International Learning Analytics and Knowledge Conference* (pp. 110–120). Association for Computing Machinery. https://doi.org/10.1145/3448139.3448150

Zhang, W., Wang, Y., Yang, L., & Wang, C. (2020). Suspending classes without stopping learning: China's education emergency management policy in the COVID-19 outbreak. *Journal of Risk and Financial Management*, *13*(3), Article 55. https://doi.org/10.3390/jrfm13030055

Zhao, L., He, W., & Su, Y. S. (2021). Innovative pedagogy and design-based research on flipped learning in higher education. *Frontiers in Psychology*, *12*, Article 230. https://doi.org/10.3389/fpsyg.2021.577002.

Zhao, L., Liu, X., & Su, Y. S. (2021). The differentiate effect of self-efficacy, motivation, and satisfaction on pre-service teacher students' learning achievement in a flipped classroom: A case of a modern educational technology course. *Sustainability*, *13*, Article 2888. https://doi.org/10.3390/su13052888

Zhu, Y. Q., Chen, L. Y., Chen, H. G., Chern, C. C. (2011). How does Internet information seeking help academic performance? The moderating and mediating roles of academic self-efficacy. *Computers & Education*, *57*(4), 2476–2484. https://doi.org/10.1016/j.compedu.2011.07.006

Zhu, Y., Zhang, J. H., Au, W., & Yates, G. (2020). University students' online learning attitudes and continuous intention to undertake online courses: A self-regulated learning perspective. *Educational Technology Research and Development*, *68*, 1485–1519. https://doi.org/10.1007/s11423-020-09753-w

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, *25*(1), 3–17. https://doi.org/10.1207/s15326985ep2501_2

Zimmerman, B. J. (2000). Attaining self-regulation: a social cognitive perspective. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press. https://doi.org/10.1016/B978-012109890-2/50031-7

Zimmerman, B. J. (2015). Self-regulated learning: Theories, measures, and outcomes. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed, pp. 541–546). https://doi.org/10.1016/B978-0-08-097086-8.26060-1

Athabasca University

CC BY

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

From Physical to Virtual: A New Learning Norm in Music Education for Gifted Students

Md Jais Ismail[1], Azu Farhana Anuar[2], and Fung Chiat Loo [3]
[1]Faculty of Music, Universiti Teknologi MARA (UiTM), Shah Alam, Selangor, Malaysia; [2]Student Development, UniKL MICET, Universiti Kuala Lumpur, Alor Gajah Campus, Melaka, Malaysia; [3]Department of Music, Faculty of Human Ecology, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

Abstract

Music education is a subject that is generally thought to have much physical activity involved. However, virtual learning has been mandatary applied to most schools worldwide due to the COVID-19 pandemic. The landscape of music learning has had to be switched to online distance learning (ODL), where students learn music virtually using technological tools. Gifted students are among those affected by the implementation of music ODL throughout 2020. Thus, the purpose of this study is to identify the effectiveness of music ODL on gifted students' motivation. The researchers framed this quantitative study by involving 81 secondary gifted students, aged 13 years, from 13 states in Malaysia. The sample was selected through random sampling, and a preexperimental design was applied to conduct the study. Respondents had been exposed to the music ODL intervention for a month. Data were collected through an adapted questionnaire, namely, the MUSIC Inventory, with a five-point scale. Data were further analysed by descriptive and inferential statistics, integrating two-way MANOVA, using SPSS Statistics version 23. Results reveal that an ODL approach to music classes is significantly effective to enhance gifted students' motivation domains of empowerment, usefulness, success, interest, and caring. Yet, no significant difference was found in gifted students' genders and locations on the four domains. Different approaches in music teaching could be further explored for music ODL to gifted students in future studies.

Keywords: music education, distance learning, COVID-19, preexperimental, two-way MANOVA, online learning

Introduction

COVID-19 has caused tragic consequences to human life. Governments all over the world have taken drastic steps by implementing lockdowns for some sectors, particularly education, as learning tends to involve crowds and physical interaction. Schools have been closed, affecting billion of students and teachers who have had to face teaching and learning at home. Online distance learning (ODL), which has been implemented during this critical situation, has been found to be an effective approach to education. Yet, ODL's effectiveness remains obscure, as schools are rushing to implement it without thorough planning. The subject of music, which has always been taught face-to-face (physically), also has had no alternative but to be implemented online. Teachers are struggling to deliver their teaching contents with the best methods to ensure students understand the music concepts and skills through a virtual platform.

This situation has similarly affected gifted education schools. Regardless of different standards and cognitive levels, students with high cognitive ability in this category have had to adjust to learning through the new norm. As music is found to be significant on these students particularly to cater emotion issues, as a motivation booster and creates awareness, as mentioned by Md Jais and Azu Farhana (2020), they need an effective way to explore music, with teachers' guidance. In the normal situation, music is taught in a music room, in which all gifted students learn music instruments, history, theory, dance, and composition. They may complete music tasks through groups, in pairs, or individually. Teachers may monitor students and offer specific interventions to students personally, as well as giving them advice and encouragement. It is rather demanding, nonetheless, to implement this method via ODL. Without a proper approach to music ODL, gifted students may face and feel stress to learn music. ODL may discourage them, causing them to withdraw from musical activities. Consequently, their music achievements and talents may be worsened and depleted.

ODL has become a new teaching norm in music learning around the world, from preschool to tertiary education institutions (Almusharraf & Khahro, 2020; Atabey, 2021; Wen & Kim Hua, 2020). Schools were closed for a quite a long time, and we have considered some challenges while designing this ODL method. We aim to determine the effectiveness of music ODL from the perspective of gifted students in term of four domains: empowerment, usefulness, caring, and success. We hope the findings from this study help music practitioners to improve their ODL pedagogy, especially during critical situations where students have to learn from home. Previous studies (Dori et al., 2018; Hernández-Torrano, 2018; Kerr & Huffman, 2018) have proven that gifted students' achievement is interrelated with gender and location; as such, we involved these variables in this study in a Malaysian music education context.

Transition from Physical Class to Music Digitisation

Music is a subject that emphasizes practical skills, such as playing music instruments, singing, and dancing, rather than theory. The evolution of the approach to teaching music began a few decades ago, when students were taught outdoors, under trees with a radio and phonograph. Thus far, teachers have continued to teach students in physical, face-to-face, personal music classes, as well as integrating blended learning in their teaching process. Teachers are trained to apply Kodaly, Dalcroze, Orff, and Suzuki methods.

In gifted education, music is needed to enrich talents among gifted students. Gifted individuals are defined as those who possess the combination of three traits as discussed in the Three Rings Model (Renzulli, 2016; Ismail et al., 2021). This model states that gifted students require challenging learning, above and beyond the classroom's four walls. Teachers are urged to implement differentiated learning when instructing gifted students. Hence, music is normally taught by applying differentiated instructions in which students are taught to prioritize their abilities. Students are allowed to learn by differentiated learning processes, contents, products, and their preferred environment. This is aligned with the research of Hymer and Michel (2013), who promoted a wide, balanced and appropriate curriculum, the differentiated education in the classroom by enriching and expanding the curriculum, and their dedication to the entire gifted personal, social and intellectual growth. Hymer and Michel accept that each child has the right to be given a high-quality education, that the primary function of a school is to provide all children with opportunities to achieve educational goals, and that deep learning takes place collaboratively, not competitively.

Gifted and talented young people have special needs that call for differentiated strategies and educational methods (VanTassel-Baska, 1994). They need interactions with intellectual peers as well as independent study experiences (Feldhusen, 2005). Feldhusen offers a list of methods for improving talented young people's learning experiences. These include appraisal, individualisation, high expectations, challenges, intellectual uncertainty, mentors, generative learning or constructivism, and meta-cognition. Teachers must understand how people learn, develop awareness, and apply teaching practice that is adequate for talented learners in order to teach them effectively. This validates the educational contexts of high schools and opportunities because they are structured to explicitly appeal to gifted and talented students.

Atterbury (1990) has examined the link between music education and the education of academically gifted and talented students:

> Not all gifted and talented students are performers. Music educators must find ways of meeting the needs of those students whose cognitive processes are substantially different from those of their peers. … Goals and objectives for these learners should be constructed to include more complex cognitive processes; that is, analysis, synthesis, and evaluation should be emphasized rather than an extra accumulation of facts. (Atterbury, 1990, p. 49)

This statement demonstrates the idea that gifted and talented students need music learning experiences to fulfil their particular education needs. Thus, music teachers in high schools that are in charge of academic courses and classroom resources are responsible for providing their students with cognitive and rigorous music programmes.

The issue of the generalist teacher's musical competence is important when it comes to bringing music to the elementary school classroom and particularly to gifted classes. The skill efficiency of generalist teachers who teach music, and the contribution of previous experience and education to their work, is discussed by Bartel and Cameron (2004). These authors look at the ties between self-effectiveness and ability and conclude that non-musician systems are expecting that they can't do anything. This concern is related to academically gifted students' need for complex, intellectual, and high-level cognitive learning. Teachers must have a sound understanding of musical principles to teach them efficiently.

Previous studies have shown a significant impact of online music learning on students' achievement. Edward et al. (2019) reveal that applying a blended learning strategy in an Oriental music class more greatly improved students' music academic performances compared with the traditional "chalk-and-talk" method. Keast (2009) strengthened this point by conducting a study on implementing distance learning with his music history class. The results indicate that distance learning, with the integration of advanced technology, heightened the constructivism method to a higher level. Ruokonen and Ruismäki (2016) have also proven that online music learning provides more opportunities for independent and constructive learning. By integrating technology in music composition, this study proved the enhancement of students' behaviour and further improves students' musical skills. This is related to Ruthmann and Hebert's (2012) study, which emphasizes that music can indeed be taught online, whether as virtual or blended learning, to diversify music education. This process enhances interaction between musicians, teachers, and students via a digital platform. Some studies of gifted education (Abakumova et al., 2019; Wallace, 2005) have proven the effectiveness of ODL but do not particularly focus on the music subject. Hence, as music is evinced, by Md Jais and Azu Farhana (2020), to be significant to gifted students, this study provides a special exploration on music to be taught through digital platform.

Objectives

The objectives of this study were the following:

- to identify the effectiveness of level of music ODL among gifted students in terms of empowerment, usefulness, success, interest, and caring;
- to identify the effectiveness of music ODL between male and female gifted students in terms of empowerment, usefulness, success, interest, and caring; and
- to identify the effectiveness of music ODL between urban and rural gifted students in terms of empowerment, usefulness, success, interest, and caring.

Hence, we developed two hypotheses to be tested based on the research objectives. The hypotheses are as follows:

> H_1: There is a significant difference on the effectiveness of music ODL between male and female gifted students in term of empowerment, usefulness, success, interest, and caring.
>
> H_2: There is a significant difference on the effectiveness of music ODL between urban and rural gifted students in term of empowerment, usefulness, success, interest, and caring.

Methodology

This study applied a preexperimental approach (Rogers & Revesz, 2020) by using a questionnaire to obtain data from students. The questionnaire was adapted from Jones's (2017) instrument after obtaining permission from the owner. It was then followed by a pilot test on 36 Malaysian gifted school students. The internal reliability of the questionnaire with Cronbach's alpha was 0.91, indicating a reasonable internal consistency. Data were collected from 81 gifted students from various locations (urban and rural) in 13 states in Malaysia. The number of respondents is aligned with Roscoe's (1975) recommendation that an experimental study is best conducted with at least 30 respondents. The respondents were 13-year-old students who took Malaysian music classes and possessed similar mental age characteristics (IQ test: > 130). This is aligned with research by Urruzola and Bernaras (2020) and Md Jais and Azu Farhana (2020), who previously conducted music research involving gifted students 12 and 13 years of age. Respondents' profiles (gender and locations) are shown in Tables 1 and 2.

Table 1

Respondents' Gender

Gender	Frequency	%
Male	38	47
Female	43	53
Total	81	100

Table 2

Respondents' Location

Location	Frequency	%
Urban	47	58
Rural	34	42
Total	81	100

Random sampling was used to select respondents. The questionnaire, named the MUSIC Inventory, was distributed to respondents and marked through Google Forms. Before answering the questionnaire, all respondents were briefed regarding the study and completed ethical consent forms. The study time frame was one month, between April 1 and April 30, 2020, in which music class was conducted once a week (four session in a month), one hour per session.

Instrument

We believed it would not be sufficient to use achievement scores to measure music ODL's effectiveness. We decided to measure the effectiveness of music ODL in a more holistic way, in which the instrument measures gifted students' intrinsic character. The instrument employed in this study was the MUSIC Inventory, adapted from Jones (2017). This consisted of 18 items under five domains, as shown in Table 3. The instrument was distributed to respondents as a five-point scale (from 1, strongly disagree, to 5, strongly agree), as recommended by Doshi et al. (2020). Respondents accessed the questionnaire on Google Forms via a smartphone, tablet, laptop computer, or desktop computer. Respondents were allowed to answer the questions once at a time.

Table 3

MUSIC Inventory

Items	No. of items	Domain
I have the freedom to complete my music class work on my own way.	5	Empowerment

I have choices on what I am allowed to do in music class.	12	
I have control over how I learn the content in music class.	16	
I have options on how to achieve the goals in music class.	18	
The knowledge I gain in music class is important for my future.	1	Usefulness
In general, music class work is useful for me.	9	
I find music class work to be relevant to my future.	13	
I am confident that I can succeed in music class work.	2	Success
I am capable of getting a high grade in music class.	4	
I feel that I can be successful in meeting the academic challenges in music class.	7	
During music class, I feel that I can be successful on the class work.	10	
The music class work is interesting to me.	6	Interest
I enjoy completing music class work.	8	
The music class work holds my attention.	17	
My music teacher cares about how well I do in music class.	3	Caring
My music teacher is friendly.	11	
My music teacher is willing to assist me if I need help in music class.	14	
My music teacher is respectful of me.	15	

Note. Adapted from *User Guide for assessing the Components of the MUSIC® Model of Motivation*, by B. D. Jones, p. 24, December 2017 (https://www.themusicmodel.com/wp-content/uploads/2019/06/User-Guide-to-Assessing-the-MUSIC-Model-Components-December-2017-2.pdf). Copyright 2021 by Brett D. Jones.

Data collected from the questionnaires were tabled and analysed using SPSS Statistics version 23. Descriptive and inferential statistics were reported, including percentage, mean, and one-way multivariate analysis of variance (MANOVA). We determined the level of the mean score by following Hassan et al. (2009) mean interpretation as in Table 4.

Table 4

Mean Score Interpretation

Mean score	Interpretation
1.00–1.99	Weak
2.00–2.99	Low
3.00–3.99	Moderate
4.00–5.00	High

Note. Adapted from "Kajian persepsi pelajar terhadap tahap profesionalisme guru Pendidikan Islam MRSM [A study on students' perception towards teachers' profesionalism in MRSM school]" by S.N.S. Hassan , 2009, Journal of Islamic and Arabic Education, 1(2), p. 38. Copyright 2009 by the Journal of Islamic and Arabic Education.

Procedure

As shown in Table 5, a four-week music lesson comprising various music fields, such as movement, history, and instruments, was planned. The distance learning lesson was executed in a way in which respondents attended the class virtually via either a desktop computer, smartphone, tablet, or laptop computer. Participants required a strong Internet connection to access the class.

Table 5

Lesson Plan

Week/time	Activity
Week 1 (Tuesday) 9.00 am–10.00 am	I. Students were taught through video about music creative movement, namely *Inang*, as shown:  The video was uploaded in a group on Telegram messenger. II. Students imitated and improvised the dance steps of the video based on *Inang* rhythm:  III. Students recorded their dance steps and sent to the respective teacher through Telegram. The teacher exhibited the best dance video in the Telegram group.
Week 2 (Tuesday) 9.00 am–10.00 am	I. Students watched a video about Malaysian traditional music uploaded in the Telegram group. II. Students gave their opinion through discussion with the teacher in the Telegram group. III. Students were given a slide presentation in the group. Students read the slide and answered a quiz.
Week 3 (Tuesday) 9.00 am–10.00 am	I. Students learned music entitled *Muzik Istana* through Google Meet. II. The teacher presented slides and videos from YouTube (https://youtu.be/J9Skrt_rxI8). Students prepared a digital folio and submitted to the teacher.
Week 4 (Tuesday) 9.00 am– 10.00 am	I. Video, notes, and a quiz about *Muzik Gamelan* were sent through Google Classroom. The Gamelan music movement was as below.

II. Students accessed Google Classroom, listened, watched the video, and comprehended the notes and music piece.
III. The students answered simple essay questions through the Google Form and practised the music piece using a recorder or any pitched instruments.
IV. Students were required to enclose their video presentation in the Google Form.

The best presentations were uploaded in the Telegram group.

Results

The results of this study were analysed and reported in descriptive and inferential statistics. As shown in Table 6, we found that respondents scored high in almost all domains tested, with a mean score above 4.00. Hasan et al. (2009) interpreted a mean score above 4.00 to be considered high. Specifically, in the empowerment domain, female students scored slightly higher than male students (*M* = 4.10, *SD* = 0.49; *M* = 4.00, *SD* = 0.68, respectively). In the usefulness domain, female students (*M* = 4.23, *SD* = 0.58) scored higher than male students (*M* = 4.02, *SD* = 0.64), while female students (*M* = 3.78, *SD* = 0.70) scored slightly higher than male students (*M* = 3.72, *SD* = 0.77) in the success domain. Additionally, female students (*M* = 4.06, *SD* = 0.62) scored higher than male students (*M* = 3.75, *SD* = 0.94) in the interest domain and in the caring domain, male students (*M* = 4.41, *SD* = 0.48) scored slightly higher than female students (*M* = 4.40, *SD* = 0.52). We conclude that a higher mean was recorded for female students in most domains, with the compared mean value between 0.06 and 0.20.

The Google Forms results showed slightly different mean scores with reference to location. Students who lived in rural areas (*M* = 4.10, *SD* = 0.54) scored higher those living in urban areas (*M* = 4.02, *SD* = 0.63) in the empowerment domain. In the useful domain, urban students (*M* = 4.15, *SD* = 0.62) scored slightly higher than rural students (*M* = 4.10, *SD* = 0.62). While in the success domain, urban students (*M* = 3.76, *SD* = 0.69) scored slightly higher than rural students (*M* = 3.73, *SD* = 0.79), urban students (*M* = 3.92, *SD* = 0.81) were shown to score slightly higher than rural students (*M* = 3.89, *SD* = 0.79) in the interest domain. Additionally, in the caring domain, urban students (*M* = 4.41, *SD* = 0.52) scored slightly higher than rural students (*M* = 4.40, *SD* = 0.47). We conclude that urban students scored higher in most domains with the compared mean value within an extremely small range, between 0.01 and 0.05.

Table 6

Descriptive Statistics

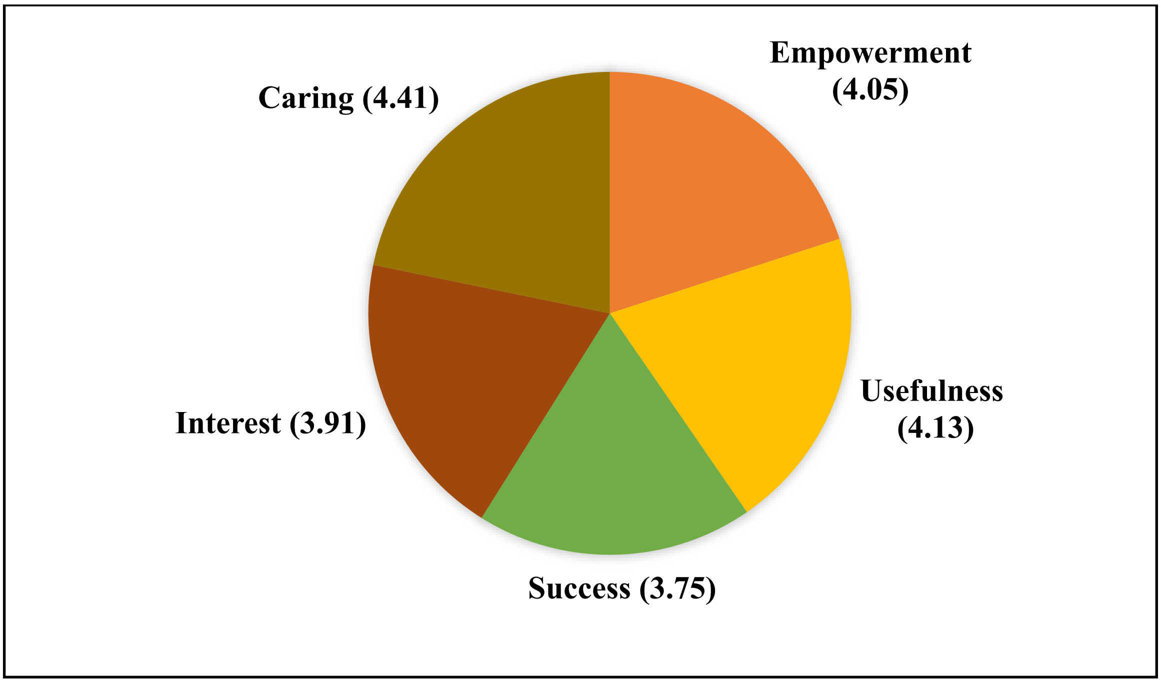
Domain	Gender	Location	*N*	Mean	*SD*
Empowerment	Male	Urban	19	3.94	0.78
		Rural	20	4.06	0.59
		Total	39	4.00	0.68
	Female	Urban	28	4.08	0.50
		Rural	14	4.16	0.49
		Total	42	4.10	0.49
	Total	Urban	47	4.02	0.63
		Rural	34	4.10	0.54
		Total	81	4.05	0.59
Usefulness	Male	Urban	19	4.05	0.67
		Rural	20	4.00	0.63
		Total	39	4.02	0.64
	Female	Urban	28	4.22	0.58
		Rural	14	4.26	0.60
		Total	42	4.23	0.58
	Total	Urban	47	4.15	0.62
		Rural	34	4.10	0.62
		Total	81	4.13	0.62
Success	Male	Urban	19	3.80	0.70
		Rural	20	3.65	0.85
		Total	39	3.72	0.77
	Female	Urban	28	3.74	0.69
		Rural	14	3.85	0.72
		Total	42	3.78	0.70
	Total	Urban	47	3.76	0.69
		Rural	34	3.73	0.79
		Total	81	3.75	0.73
Interest	Male	Urban	19	3.76	1.07
		Rural	20	3.74	0.82
		Total	39	3.75	0.94
	Female	Urban	28	4.03	0.57
		Rural	14	4.11	0.73
		Total	42	4.06	0.62
	Total	Urban	47	3.92	0.81

		Rural	34	3.89	0.79
		Total	81	3.91	0.80
Caring	Male	Urban	19	4.55	0.45
		Rural	20	4.29	0.49
		Total	39	4.41	0.48
	Female	Urban	28	4.32	0.56
		Rural	14	4.57	0.39
		Total	42	4.40	0.52
	Total	Urban	47	4.41	0.52
		Rural	34	4.40	0.47
		Total	81	4.41	0.50

Table 6 shows that gifted students scored highest in the caring domain, with a total mean value of 4.41, followed by usefulness (*M* = 4.13), empowerment (*M* = 4.05), interest (*M* = 3.91), and success (*M* = 3.75). Figure 1 shows the total mean for each domain of the effectiveness music ODL.

Figure 1

Total Means of the Motivation Domains



MANOVA analysis was used to test whether there was a significant difference on the effectiveness of ODL for the two hypotheses. In conducting MANOVA, we ensured that the analysis met the MANOVA assumptions as recommended by Tabachnick and Fidell (1989). Box's *M* test is a prerequisite that must be performed before performing the MANOVA test to determine the homogeneity of variance–

covariance among independent variables (Pallant, 2011). The results of the analysis showed that no variance–covariance difference existed between the dependent variable and the independent variable ($F = 1.21$, $p = 0.16$) ($p > 0.05$). This allowed the MANOVA test to proceed since the number of samples was also robust. Pallant (2011) further minimized the probability of type I errors. Table 7 shows the results of Box's *M* test.

Table 7

Box's M Test Result

Box's *M*	*F*	df_1	df_2	*p*
62.37	1.21	45	9744.502	0.16

The results of the Levene's test can be seen in Table 8. With this test, we found that the significance values for the four variables, namely, empowerment, usefulness, success, and caring, were greater than 0.05. This means that these variables have identical (homogeneous) variance, which allows the MANOVA test to proceed. In contrast, Levene's test findings for interest variables were less than 0.05. This indicates differences in variants. Even so, Pallant (2011) asserts that differences in variants are not an obstacle to conducting MANOVA tests if the sample size is robust.

Table 8

Levene's Test of Equality of Error Variances

Dependent variable	*F*	df_1	df_2	*p*
Empowerment	1.44	3	77	0.23
Usefulness	0.02	3	77	0.99
Success	0.47	3	77	0.69
Interest	3.11	3	77	0.03
Caring	1.01	3	77	0.39

Next, Wilks's lambda (λ) statistic shows a comparison of mean scores in terms of music ODL implementation based on gender in gifted students. Based on Table 9, the value of Wilks's $\lambda = 0.93$, $p = 0.35$ ($p > 0.05$). The eta squared value of 7.2% was found, which means that the effect of the difference between skills tests is very small. Therefore, no sufficient evidence exists to accept H_1. This means that there is no significant difference between male and female gifted students using music ODL in terms of empowerment, usefulness, success, interest, and caring.

Table 9

Multivariate Tests Result

Effect	Wilks's λ	*F*	*p*	Partial eta squared
Gender	0.93	1.139	0.35	0.072

The next λ statistic shows the comparison of mean scores in the aspect of implementing music ODL based on the location of gifted students. Based on Table 10, the value of $\lambda = 0.98$, $p = 0.91$ ($p > 0.05$). The eta squared value of 2% was found, meaning that the effect of difference between skill tests is very small. Therefore, there is no sufficient evidence to accept hypothesis H_2. Our second hypothesis was rejected as no significant difference was found between urban and rural gifted students using music ODL in terms of empowerment, usefulness, success, interest, and caring.

Table 10

Multivariate Tests Result

Effect	Wilks's λ	*F*	*p*	Partial eta squared
Location	0.98	0.303	0.91	0.02

Since the independent variables in this MANOVA test are divided into two categories, namely, gender and location, a two-way MANOVA analysis should be performed (Pallant, 2011). Table 11 shows the test of between-subjects effects to demonstrate the differences in the dependent variables. Findings show that the difference in effect for empowerment is 0% with value $F = 0.02$, $p = 0.88$ ($p > 0.05$); usefulness is 0.1% with value $F = 0.09$, $p = 0.75$ ($p > 0.05$); success is 0.8% with value $F = 0.59$, $p = 0.44$ ($p > 0.05$); interest is 0.1% with value of $F = 0.81$, $p = 0.77$ ($p > 0.05$); and caring is 6.2% with value of $F = 5.05$, $p = 0.02$ ($p < 0.05$). This offers the impression that no significant difference exists in the aspects of empowerment, usefulness, success, and interest between the variables of gender and location during music ODL implementation. However, there are significant differences in the caring aspect between the gender and location variables in the same instance.

Table 11

Tests of Between-Subjects Effects

Dependent variable	Sum of squares	*df*	Mean square	*F*	*p*	Eta squared
Empowerment	0.007	1	0.007	0.02	0.88	0.000
Usefulness	0.037	1	0.037	0.09	0.75	0.001
Success	0.329	1	0.329	0.59	0.44	0.008
Interest	0.052	1	0.052	0.08	0.77	0.001
Caring	1.245	1	1.245	5.05	0.02	0.062

Discussion

We found that music ODL has impacted gifted students' learning in certain domains of empowerment, usefulness, success, interest, and caring. Our results indicate that music ODL is effective, according to gifted students, since all the respondents scored high in most domains. Students in our study believed that music ODL empowered them to learn music topics and further enabled them to complete all tasks given by teachers. This is aligned with the findings of Edward et al. (2019), who found that through online learning, students may feel it is easier to complete music tasks rather than through the traditional method. This is due to the flexibility and uniqueness of distance learning that allows students to commit to and focus on the topic, as mentioned by Spencer (2020).

Additionally, gifted students agreed that music ODL is useful. They believed that the knowledge that they gained during ODL could help them to achieve their goals and further attain a bright future. Skills included those related to computers and music, which students could master during ODL sessions. This is supported by Keast (2009): his study shows that students gradually become more proficient in using high-tech materials in music classes. Song et al. (2004) adds that distance learning experiences enable students to find styles that best match their learning preferences. This allows them to achieve a bright future.

All gifted students also were found to believe that music ODL would help them to achieve success. It would help them to increase their comprehension of music subjects and, further, score high marks. They agreed that the tasks given, as well as the activities conducted, on the digital platform could improve their musical skills. This is in line with Schmidt's (2005) findings that students involved in music activities have a high motivation to attain success. Music intrinsically motivates students, which relates to their academic achievement and grade level. Montacute and Cullinane (2018) state that parents should support their children's learning by providing facilities, encouragement, and a proper learning environment. With parents' support, students may experience quality online classes and may score higher marks than those who do not receive parents' support.

Gifted students also felt that music ODL held their interest, made their learning joyful, and created an interesting environment to complete tasks. They believed that virtual classes and online tasks heightened their interest in music. Both approaches could be a solution for some situations where the students lose their interest in learning, as mentioned by Anderhag et al. (2016). Teaching music remotely can be an intervention to attract students' attention and make an online class active. Thus, it is strongly suggested that while teaching students virtually, we may also replicate the method recommended in the present study to catch students' interest in a subject. It is essential to capture gifted students' interest as they may easily lose intentions to study music once they find that a class is boring and not challenging. This study proves that the topics suggested in this study that consist of creative movement, traditional music, *Muzik Istana* and *Muzik Gamelan* are significant and interesting to gifted students.

In terms of caring, gifted students believed that teachers were concerned about their needs. This means that teachers played their roles and achieved the students' expectation fulfilled what the gifted students required. Particularly during the COVID-19 pandemic, students need support from teachers to understand their feelings and guide them to fix technical issues. This goes beyond the actual practice of teaching, where teachers not only help students understand a topic but also need to be friendly and always account for students' needs. In this study, students also felt that teachers helped them, noted their level of achievement, and respected their ideas. Zhang et al. (2019) similarly found that teachers' caring behaviour affected students' cognitive reappraisal and suppression of expression. Noddings (2012) also found that teachers' concerned behaviour can be described as listening, advising, critical thinking, reflecting, and establishing thoughtful connection between disciplines and life itself.

From the results, we found that female students scored higher than male students in empowerment, usefulness, success, and interest, while male students scored higher in caring. Numerous studies show that females score better than males in music studies (Armstrong, 2013; Comber et al., 1993; Kuntsche et al., 2016), and the present study also shows evidence that female students score higher than males in music. Our findings indicate that females are likely to be more motivated than males, as they scored higher in most of the motivation domains. Males recorded a higher score in the caring domain, probably because the teacher was male, which may have caused male students to feel more comfortable as they were learning with a teacher of the same gender. This is related to Mills's (2000) findings suggesting that males should work with other males as they can understand each other, and thus, male students feel more comfortable talking to and learning from male teachers.

In addition, students who lived in urban areas believed that music ODL was useful, helped them to succeed, heightened their interest, and created a more caring teacher–student atmosphere. Gifted students living in rural areas more strongly believed that that music ODL empowered them to learn music. These results indicate that urban students feel more comfortable with music ODL than rural students. We assume that they possessed proper and better facilities to partake in distance learning than rural students. In fact, urban students used online tools in their daily lives—for example, chatting through Skype, online gaming, using social media (e.g., Facebook, Instagram, Twitter), and so forth. These daily activities help urban students feel more comfortable with distance learning and may influence their beliefs that ODL is useful to them and could lead them to success.

Researchers feel that rural students still need more sophisticated learning facilities to enable music ODL to be carried out and enable them to master a subject. Although there are differences between genders and locations, the results indicate that the effects between domains were extremely minor. Due to a minor value that each domain contributed to genders and locations, the findings revealed no significant difference on the effectiveness of music ODL between male and female gifted students in terms of empowerment, usefulness, success, interest, and caring. We also found no significant difference on the effectiveness of music ODL between urban and rural gifted students in term of empowerment, usefulness, caring, and success. However, from gifted students' perspectives, ODL is quite effective, with moderate to high ratings recorded for the five domains in the questionnaire (Table 6). Further research could be conducted with larger sample sizes involving students from more urban and rural areas. Other tests looking at how ODL affects the different academic and practical results of music students could also be conducted. Nevertheless, this approach also encourages further dissemination of the many perspectives of music learning, not only on instrumental classes but also on institutionalization of traditional music (Ismail et al., 2020, 2021).

Conclusion

The novelty of this research demonstrates that music ODL has changed the music education landscape from physical to virtual instruction. Data from the descriptive results show that gifted students tend to be motivated with regard to the implementation of music ODL. Although schools' implementation of distance learning has been rushed, the present study provides an overview of music ODL from gifted students' perspective that might help schools to consider implementing proper music ODL amidst during the critical situation of the COVID-19 pandemic. We believe that it is important for gifted students to be involved in evaluating the effectiveness of music ODL as they are unique students with extra sensitive learning needs. Without a proper education strategy, gifted students are more prone to feeling bored and tend to withdraw from music class. This may greatly impact their emotions, making them vulnerable to depression, burnout, and even suicidal thoughts.

Findings from the present study have practical implications in deploying a music distance learning system for gifted students. Schools could deploy online applications as outlined in this study for music activities for gifted students. In developing music distance learning, key activities such as virtual instruction, online tasks, and virtual discussions could be implemented to ensure learning effectiveness. We believe that music ODL is more flexible, is cost-effective, and increases students' motivation. Flexibility in ODL as highlighted in this study resonates with past mobile learning studies that have explored the possibilities of simultaneous learning of embedded secondary learning material, such as cultural heritage (Loo & Loo, 2021; Loo et al., 2016), and we see potential for further studies on gifted students. Thus, a proper intervention could be planned based on the evidence in the present study to fulfil gifted students' needs. It is hoped that the transition from physical to virtual music classes can be conducted smoothly by providing motivating material to gifted students. We recommend a true experimental research design related to music ODL among gifted students to be conducted in the future.

Acknowledgements

This study has been supported by Universiti Teknologi MARA, Malaysia. There are no conflicts of interests.

References

Abakumova, I., Bakaeva, I., Grishina, A., & Dyakova, E. (2019). Active learning technologies in distance education of gifted students. *International Journal of Cognitive Research in Science, Engineering and Education*, *7*(1), 85–94. https://doi.org/10.5937/IJCRSEE1901085A

Almusharraf, N., & Khahro, S. (2020). Students satisfaction with online learning experiences during the COVID-19 pandemic. *International Journal of Emerging Technologies in Learning (iJET)*, *15*(21), 246–267. https://online-journals.org/index.php/i-jet/article/view/15647/0

Anderhag, P., Wickman, P. O., Bergqvist, K., Jakobson, B., Hamza, K. M., & Säljö, R. (2016). Why do secondary school students lose their interest in science? Or does it never emerge? A possible and overlooked explanation. *Science Education*, *100*(5), 791–813. https://doi.org/10.1002/sce.21231

Armstrong, V. (2013). *Technology and the gendering of music education*. Ashgate Publishing, Ltd.

Atabey, D. (2021). COVID-19 from the perspective of preschool prospective teachers: What can we do for children? *International Journal on Social and Education Sciences*, *3*(1), 82–94. https://doi.org/10.46328/ijonses.76

Atterbury, B. W. (1990). Speaking the "gifted and talented" language: The key to program success. *Music Educators Journal*, *76*(7), 46–49. https://doi.org/10.2307/3401037

Bartel, L., & Cameron, Linda. (2004). From dilemmas to experience: Shaping the conditions of learning. In L. Bartel (Ed). *Questioning the music education paradigm* (Vol. II, pp. 39-61). Canadian Music Educators Association. http://hdl.handle.net/1807/24437

Comber, C., Hargreaves, D. J., & Colley, A. (1993). Girls, boys and technology in music education. *British Journal of Music Education*, *10*(2), 123–134. https://doi.org/10.1017/S0265051700001583

Davies, R. B. (1971). Hypothesis testing when a nuisance parameter is present only under the alternatives. *Biometrika*, *74*(1), 33–43. https://doi.org/10.2307/2336019

Dori, Y. J., Zohar, A., Fischer-Shachor, D., Kohan-Mass, J., & Carmi, M. (2018). Gender-fair assessment of young gifted students' scientific thinking skills. *International Journal of Science Education*, *40*(6), 595–620. https://doi.org/10.1080/09500693.2018.1431419

Doshi, D., Karunakar, P., Sukhabogi, J. R., Prasanna, J. S., & Mahajan, S. V. (2020). Assessing coronavirus fear in Indian population using the Fear of COVID-19 Scale. *International Journal of Mental Health and Addiction*. Advance online publication. https://doi.org/10.1007/s11469-020-00332-x

Edward, C. N., Asirvatham, D., & Johar, G. (2019). The impact of teaching Oriental music using blended learning approach: An experimental study. *Malaysian Journal of Learning and Instruction*, *16*(1), 81–103. https://eric.ed.gov/?id=EJ1219911

Feldhusen, J. F. (2005). Giftedness, talent, expertise, and creative achievement. *Conceptions of giftedness*, *2*, 64–79. https://doi.org/10.1017/CBO9780511610455.006

Hassan, S. N. S., Tamuri, A. H., Othaman, I., & Mamat, M. S. (2009). Kajian persepsi pelajar terhadap tahap profesionalisme guru Pendidikan Islam MRSM. *JIAE: Journal of Islamic and Arabic Education*, *1*(2), 31–50. http://journalarticle.ukm.my/768/

Hernández-Torrano, D. (2018). Urban–rural excellence gaps: Features, factors, and implications. *Roeper Review*, *40*(1), 36–45. https://doi.org/10.1080/02783193.2018.1393610

Hymer, B., & Michel, D. (2013). Gifted and talented learners: Creating a policy for inclusion. Routledge.

Ismail, M. J., Hamuzan, H. A., & Maarof, N. H. (2021). Meneroka tingkah laku unik pelajar pintar cerdas berbakat akademik [Exploring Unique Behavior of Gifted Students with Academic Talented]. *Malaysian Journal of Learning and Instruction*, 18(2), 301-328. https://doi.org/10.32890/mjli2021.18.2.11

Ismail, M. J., Loo, F. C., & Anuar, A. F. (2021). Learning music through rhythmic movements in Malaysia. *Malaysian Journal of Learning and Instruction*, *18*(1), 241–263. https://doi.org/10.32890/mjli2021.18.1.10

Ismail, M. J., Loo, F. C., Anuar, A. F, Yusof, R. (2020). Institutionalising the Kompang for primary school students in Malaysia. *International Journal of Innovation, Creativity and Change*, *13*(5), 275–292. https://www.ijicc.net/index.php/volume-13-2020/188-vol-13-iss-5

Jones, B. D. (2017, December). *User guide for assessing the components of the MUSIC® Model of Motivation*. https://www.themusicmodel.com/wp-content/uploads/2019/06/User-Guide-to-Assessing-the-MUSIC-Model-Components-December-2017-2.pdf

Keast, D. A. (2009). A constructivist application for online learning in music. *Research and Issues in Music Education*, *7*(1), 1–8. https://files.eric.ed.gov/fulltext/EJ894765.pdf

Kerr, B. A., & Huffman, J. M. (2018). Gender and talent development of gifted students. In S. I. Pfeiffer (Ed.), *Handbook of giftedness in children: Psychoeducational theory, research, and best practices* (2nd ed., pp. 115–128). Springer. https://doi.org/10.1007/978-3-319-77004-8_8

Kuntsche, E., Le Mével, L., & Berson, I. (2016). Development of the four-dimensional Motives for Listening to Music Questionnaire (MLMQ) and associations with health and social issues among adolescents. *Psychology of Music*, *44*(2), 219–233. https://doi.org/10.1177/0305735614562635

Loo, F. Y., & Loo, F. C. (2021, March 8–9). The Importance of a cultural-embedded learning model: Raising awareness of cultural heritage. In edited L. G. Chova, A. L. Martinez, & I. C. Torres (Eds.), *Proceedings of the International Technology, Education and Development Conference* (pp. 308–316). IATED Academy. https://doi.org/10.21125/inted.2021.0090

Loo, F. Y., Loo, F. C., & Chai, K. E. (2016). Learning traditional Malay folk song and tempo control by using an M-learning model design for beginner pianists. *Turkish Online Journal of Educational Technology* [special issue], 41–46.

Md Jais, I., & Azu Farhana, A. (2020). The significance of music to gifted students. *Quantum Journal of Social Sciences and Humanities*, *1*(4), 33–43. http://qjoest.com/qjssh/index.php/qjssh/article/view/21

Mills, M. (2000). Issues in implementing boys' programme in schools: Male teachers and empowerment. *Gender and Education*, *12*(2), 221–238. https://doi.org/10.1080/09540250050010027

Montacute, R., & Cullinane, C. (2018). *Parent power 2018: How parents use financial and cultural resources to boost their children's chances of success*. Sutton Trust. http://dera.ioe.ac.uk/id/eprint/32179

Noddings, N. (2012). The caring relation in teaching. *Oxford Review of Education*, *38*(6), 771–781. https://doi.org/10.1080/03054985.2012.745047

Pallant, J. 2011. *SPSS survival manual: A step by step guide to data analysis using SPSS* (4th ed.). Open University Press. https://searchworks.stanford.edu/view/8839418

Renzulli, J. S. (2016). The three-ring conception of giftedness: A developmental model for promoting creative productivity. In S. M. Reis (Ed.), *Reflections on gifted education: Critical works by Joseph S. Renzulli and colleagues* (pp. 55–90). Prufrock Press Inc. https://www.routledge.com/Reflections-on-Gifted-Education-Critical-Works-by-Joseph-S-Renzulli-and/Renzulli-Reis/p/book/9781618215055

Rogers, J., & Revesz, A. (2020). *Experimental and quasi-experimental designs*. Routledge.

Roscoe, J. T. (1975). *Fundamental research statistics for the behavioral sciences* (2nd ed.). Holt Rinehart & Winston.

Ruokonen, I., & Ruismäki, H. (2016). E-learning in music: A case study of learning group composing in a blended learning environment. *Procedia—Social and Behavioral Sciences*, *217*, 109–115. https://doi.org/10.1016/j.sbspro.2016.02.039

Ruthmann, S. A., & Hebert, D. G. (2012). Music learning and new media in virtual and online environments. In S. A. Ruthmann & D. G. Hebert (Eds)., *The Oxford handbook of music education*, *Vol. 2* (pp. 567–583). New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199928019.013.003

Schmidt, C. P. (2005). Relations among motivation, performance achievement, and music experience variables in secondary instrumental music students. *Journal of Research in Music Education*, *53*(2), 134–147. https://doi.org/10.2307/3345514

Song, L., Singleton, E. S., Hill, J. R., & Koh, M. H. (2004). Improving online learning: Student perceptions of useful and challenging characteristics. *The internet and higher education*, *7*(1), 59-70. https://doi.org/10.1016/j.iheduc.2003.11.003

Spencer, J. (2020, October 15). Empowering students in distance learning environments. *John Spencer*. https://spencerauthor.com/empower-distance-learning/

Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics*. Harper & Row.

Urruzola, M. V., & Bernaras, E. (2020). Music performance anxiety in 8-to 12-year-old children. *Revista de Psicodidáctica* (English ed.), *25*(1), 76–83. https://doi.org/10.1016/j.psicoe.2019.10.003

VanTassel-Baska, J. (1994). *Comprehensive curriculum for gifted learners*. Allyn & Bacon.

Wallace, P. (2005). Distance education for gifted students: Leveraging technology to expand academic options. *High Ability Studies*, *16*(1), 77–86. https://doi.org/10.1080/13598130500115288

Wen, K. Y. K., & Kim Hua, T. (2020). ESL teachers' intention in adopting online educational technologies during COVID-19 pandemic. *Journal of Education and e-Learning Research*, *7*(4), 387–394. https://doi.org/10.20448/journal.509.2020.74.387.394

Zhang, T., Wang, Z., Liu, G., & Shao, J. (2019). Teachers' caring behavior and problem behaviors in adolescents: The mediating roles of cognitive reappraisal and expressive suppression. *Personality and Individual Differences*, *142*(1), 270–275. https://doi.org/10.1016/j.paid.2018.10.005

Athabasca University

CC BY

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

"They Have to Combine the Future of the University and Their Own Future": OpenCourseWare (OCW) Authoring as an Academic Practice in Spain

Daniel Villar-Onrubia, Ph.D.
Coventry University, UK

Abstract

This study looks at OpenCourseWare (OCW) in Spain, a country where most public universities have tried to promote that particular model of open educational resources (OER) provision among academics. Using three universities with varying levels of OCW activity as a case study, this article examines key drivers behind the implementation of OCW initiatives and unpacks what it means, as an academic practice, to engage in OCW authoring. Following a qualitative case study approach and a multi-methods design, this study offers a basis for theoretical generalisations that can be useful for understanding similar dynamics taking place within different organisational contexts in Spain and beyond. The findings reveal a major disconnect between the drive to implement OCW initiatives in Spain and actual opportunities for academics to engage with them as part of their work. The author concludes that the extrapolation of a highly prescriptive model of OER provision into institutional realities different from the context where it was originally devised—in this case, the Massachusetts Institute of Technology in the United States—is rather problematic. The article also provides some recommendations to university leaders and policy makers, encouraging the creation of alternative models that are mindful of the institutional and cultural specificities of their own contexts and also to take into consideration the social and material realities of the communities they aim to provide with lifelong learning opportunities.

Keywords: OpenCourseWare, OCW, open educational resources, OER, open educational practices, OEP, open education, higher education, universities, Spain

Introduction

Over the last two decades, the creation of open educational resources (OER) in higher education (HE) has generated a considerable amount of attention worldwide. A major milestone in the history of OER was the launch of the OpenCourseWare (OCW) initiative by the Massachusetts Institute of Technology (MIT), conceived as a "free and open digital publication of high quality educational materials, organized as courses" (Carson, 2009, p. 27). Soon after, other universities in countries around the globe became interested in replicating that OER provision model (Carson & Forward, 2010), and Spain quickly stood out due to an extraordinarily high number of institutions in the country doing so (Aranzadi & Capdevila, 2011).

Most public universities in Spain implemented an OCW initiative between 2006 and 2010, but this fact may result in misleading conclusions about the actual importance and long-term uptake of OCW authoring in the Spanish higher education context. The number of Spanish universities in the OCW Consortium started to dwindle in 2014, and currently, the majority of OCW sites "are not up to date, they redirect to other university systems, or they are no longer operational" (Santos-Hermosa et al., 2020, p. 6).

The amount of research on OCW—and more generally OER—in Spain has been modest, and most studies were published several years ago. They include descriptive accounts of the implementation of OCW initiatives at particular universities (Clifton et al., 2013; Gallardo Paúls, 2008; Llorens et al., 2010; Ros et al., 2014) and comparative studies focused on technical features or the nature and structure of content (Borrás Gené, 2010; Llorente Cejudo et al., 2013). Two research projects on the impact of OCW in Spain and Latin America were funded by the Spanish Government (Frías-Navarro et al., 2010; Tovar et al., 2013).

The purpose of this study was to research the interplay between the OCW model and the institutional arrangements influencing academics' behaviour at Spanish universities. Using three universities as a case study, this article focuses on what it means to be an OCW author in those contexts and the role of OCW authoring as an academic practice, offering a basis for theoretical generalisations that can be useful for understanding similar dynamics at different organisational contexts in Spain and beyond. Additionally, the role of the Universia organisation and network of universities is considered to have played a key role in the proliferation of OCW initiatives in Spain by establishing a regional consortium (with HE institutions in Latin America and Portugal too) associated with the global OCW Consortium.

The contextualisation for this study is considered next, including the theoretical framing and identification of the research questions addressed by the case study approach, which is used to structure the remainder of the article, before moving to the discussion, conclusions, and recommendations.

Background

OCW as a Model of OER Provision

In 2001, MIT launched the first OCW initiative, which was aimed at providing the public with access to high-quality OER and covering the entire curriculum of a selection of courses taught to its students, with the ambition to eventually include all undergraduate and graduate courses (Abelson et al., 2021). A major concern during the design and implementation of the initiative was to ensure that it would not undermine in any way the service of MIT as a residential institution, which was always regarded as a high priority (Abelson, 2008). After a consultation process, it became clear that academics could only uptake OCW authoring if the publication process was offloaded to support staff, and therefore, "the plan for implementing OCW had to minimize faculty burden" (Lerman et al., 2008, p. 219). As Lerman and Miyagawa (2002, p. 27) stress, academics "operate essentially at capacity, and doing any new task inevitably means not doing something else." A sizeable budget was secured to enable the development of OCW courses at a scale (Abelson et al., 2021), and beyond financial resources, another key factor was that the initiative drew on a culture of open sharing that was already deeply entrenched in the MIT academic community (Lerman et al., 2008).

OCW as a Global Phenomenon

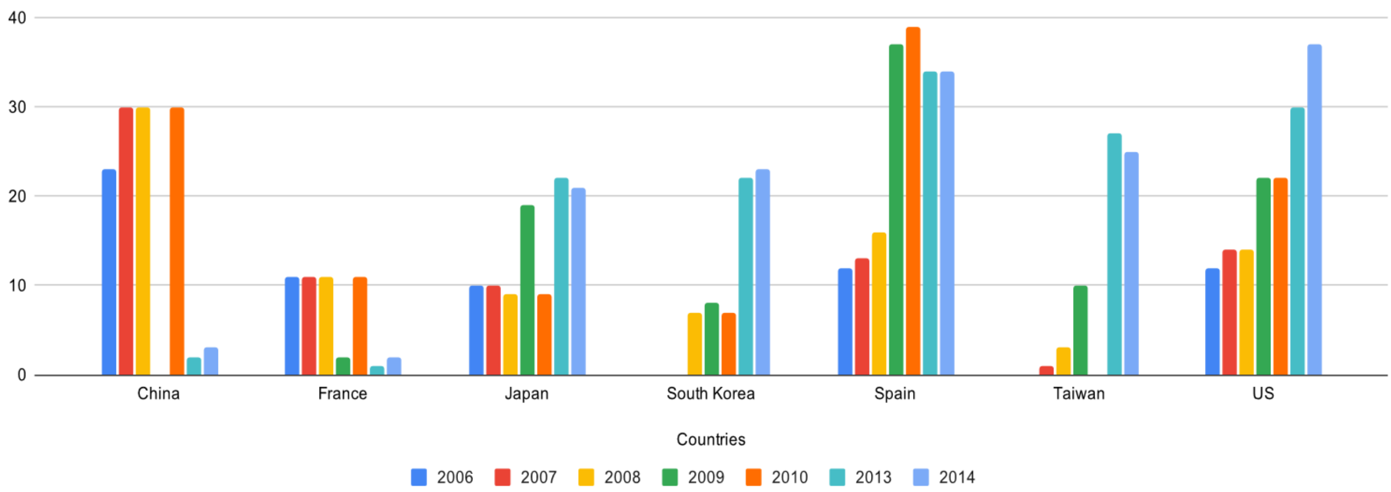
The OCW-MIT initiative was unique in capturing the imagination of policy makers, journalists, and opinion leaders all over the world. It contributed to redefining the *open* concept in education (Peter & Deimann, 2013) and inspired the coinage of OER as an established term (UNESCO, 2002). Moreover, it triggered other institutions' interest in launching their own OCW initiatives, and MIT collaborated with other universities and organisations to establish the OCW Consortium, which was officially launched in 2006, rebranded as the Open Education Consortium (OEC) in 2014, and again in 2019 as Open Education Global (OEG). Some studies have looked at key factors influencing OCW authoring in specific countries, such as Turkey (Kursun et al., 2014) and Taiwan (Wei & Chou, 2021).

OCW in Spain

Spain quickly stood out as one of the most active countries in terms of institutional members in the OCW Consortium (Figure 1). However, after 2014, the number of Spanish universities in the OCW Consortium decreased dramatically (Figure 2). In 2021, only six Spanish universities remain in OEG.

Figure 1

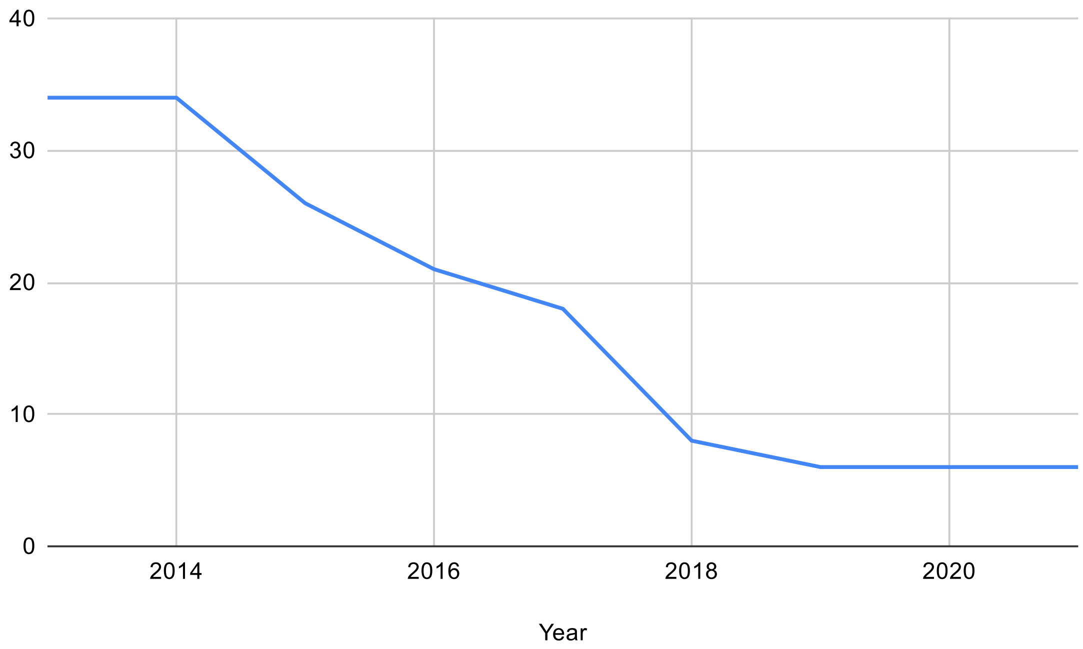
Higher Education Institutions in the OCW Consortium per Country



Note. Figure created by the author. It includes only those countries with more than 20 institutional members. Source data collected from the OCW Consortium website (http://ocwconsortium.org) and captures of that site by the Wayback Machine (https://web.archive.org).

Figure 2

Spanish Higher Education Institutions in the OCW Consortium (2013–2014), Open Education Consortium (2014–2018), and Open Education Global (2019–2021)



Note. Figure created by the author. Source data collected from the websites of the OCW Consortium (http://ocwconsortium.org/), the Open Education Consortium (http://oeconsortium.org/), Open Education Global (https://www.oeglobal.org) and captures of those sites by the Wayback Machine (https://web.archive.org). CC-BY.

Understanding the proliferation of OCW initiatives in Spain requires paying close attention to the advocacy efforts of Universia, an organisation launched in 2000 and sponsored by the Santander Bank—a key financial player with both philanthropic and commercial interests in the HE sector in Ibero-America (Lloyd, 2011). In 2002, Universia and MIT signed an agreement to translate a selection of OCW-MIT courses into Spanish and Portuguese. Later, Universia became one of the first sustaining members of the OCW Consortium and encouraged a dozen Spanish universities to join it, launching OCW-Universia as an associate consortium in 2007 (Aranzadi & Capdevila, 2011).

The lack of incentives associated with OCW authoring, a low level of awareness among academics, and the overall limited resources allocated by universities have been identified as key barriers to the implementation of OCW initiatives in Spain (Frías-Navarro et al., 2010; Tovar et al., 2013).

In addressing the above concerns, the following research questions guided this study:

1. What is the role of OCW initiatives with regard to the strategic orientation of universities?
2. To what extent does the implementation of OCW initiatives contribute to the adoption of OER-based practices and increase in OER awareness?
3. How do OCW authors perceive such activity in relation to their professional identities and activity?

Theoretical Framework

A socio-technical perspective in line with the principles of social informatics (Meyer et al., 2019) was used as a "theoretical foundation for addressing e-learning, deriving ... from the sociology of contemporary culture, particularly where it intersects with computing use by groups, organizations, communities, and societies" (Andrews & Haythornthwaite, 2007, p. 27). This approach was adopted to develop a more nuanced and context-aware understanding of technologies in use, offering an alternative to deterministic accounts often found in education and technology (EdTech) discussions (Oliver, 2011).

More specifically, the study viewed OCW initiatives as socio-technical interaction networks: "A Socio-Technical Interaction Network (STIN) is a network that includes people (including organizations), equipment, data, diverse resources (money, skill, status), documents and messages, legal arrangements and enforcement mechanisms, and resource flows" (Kling et al., 2003, p. 48). STIN models help one to grasp the intricacies of technology-mediated human behaviour by shedding light upon both technological systems and social factors, which are accounted for as highly enmeshed phenomena, and have proven relevant to the study of education and technology (Creanor & Walker, 2012; McCoy & Rosenbaum, 2019; White et al., 2020).

Instead of focusing on the analysis of technical platforms and the potential activities enabled by them, this study examines the manifold dynamics, processes, and practices that bring these socio-technical arrangements into being as composite networks made of OCW authors, technicians, policy makers, advocates, technical protocols, and others. The STIN framework provides a set of heuristics that can help to articulate the inquiry on socio-technical systems at various levels, prompting researchers to identify a "relevant population of system interactors, core interactor groups, incentives, excluded actors and undesired interactions, existing communication forums, system architectural choice points, resource flows [and then] map architectural choice points to socio-technical characteristics" (Kling et al., 2003, p. 57).

While the STIN strategy was used to map out key elements and relationships around OCW initiatives, the interpretive process—that is, making sense of the practices and perspectives of people involved in those initiatives—was grounded in a set of principles underpinning social informatics and cognate with a socio-materiality lens (Orlikowski & Scott, 2008).

Methodology

Research Design

Drawing on the STIN framework, the research adopted a qualitative case study approach and followed a multi-methods design (Denzin & Lincoln, 2011; Miles & Huberman, 1994), involving the analysis of discourse and behaviour as embodied in (a) documents of different kinds, (b) the accounts provided by research participants, and (c) online data. It addressed the three research questions by investigating OCW authoring within its real-life context, examining several organisations within a bounded system over time through in-depth data collection and drawing on diverse information sources (Creswell, 2007; Yin, 2003).

The study started with a review of documents and online data (phase 1) relating to OCW in Spain, followed by in-depth interviews (phase 2) with both OER experts in that country ($n = 4$) and participants involved in establishing OCW-Universia ($n = 3$). The analysis of data collected during the first stages informed the design of the interview guides used as part of the subsequent fieldwork, which involved semi-structured interviews with university leaders and professional staff at the three sites of the case study ($n = 23$) and at other universities in the same regional system ($n = 10$) (phase 3). The final stage of fieldwork (phase 4) entailed interviewing academics at the three sites ($n = 24$) who had been involved in authoring OCW courses and other kinds of learning resources available on the Web.

Sites

The overall aim to maximise "what we can learn" (Stake, 1995, p. 6) guided the selection of the three universities as well as the wider system in which they are embedded. A "purposeful maximal sampling" (Creswell, 2007, p. 75) strategy was used, aimed at generating a rich account of the interplay between the OCW model of OER provision and the specificities of different institutional settings. The three universities of the case study are anonymised here as UNI-1, UNI-2 and UNI-3, while other key universities analysed for contextual purposes are anonymised as UNI-a, UNI-b, and UNI-c.

Both UNI-1 and UNI-3 are large universities, while UNI-2 is medium sized. The OCW initiatives of those three institutions achieved different levels of activity in terms of number of OCW courses and number of OCW authors. Less than 1% of academics in UNI-1 (i.e., about 30), 3% in UNI-2 (i.e., about 60), and above 10% in UNI-3 (i.e., about 400) had contributed to their respective OCW initiatives.

Participants

The selection of research participants followed a purposive sampling strategy to cover different perspectives and to "establish a good correspondence between research questions and sampling" (Bryman, 2008, p. 458). The samples included three groups of key actors: (a) university leaders (i.e., senior management), (b) professional services staff, and (c) academics (Table 1). The academics recruited as participants had been particularly active in authoring OCW modules and, to a lesser extent, other types of OER and online learning materials.

With the aim of gaining insight into the wider context and its influence on the case study sites, interviews were also conducted with participants involved establishing OCW-Universia and university leaders and staff at other Spanish universities in the same regional HE system.

Table 1

Research Participants per Setting and Category

Organisation	University leaders	Professional staff	OCW authors	Authors of other kinds of learning resources on the Web	Others	Total
UNI-1	4	3	5	2	-	14
UNI-2	5	4	7	1	-	17
UNI-3	3	4	7	2	-	16
UNI-a	1	2	-	-	-	3
UNI-b	1	2	-	-	-	3
UNI-c	2	2	-	-	-	4
Others	-	-	-	-	7	7
Total	16	17	19	5	7	64

Note. OCW = OpenCourseWare.

Data Collection and Analysis

Desk research (phase 1) included analysing policy and strategy documents (statutes, bylaws, manifestos, plans) relevant to OCW authoring at each university and the whole Spanish HE sector. Other relevant sources (i.e., calls for participation, texts about the OCW initiatives available at their respective sites, brochures, press releases) were also analysed. Online data were collected to assess the scale of OCW initiatives, and Webometrics techniques were applied to gain insight into the overall level of attention paid to the notion of OER across Spanish universities (Villar-Onrubia, 2014).

In-depth interviewing was key to both tracing the origins and evolution of the OCW initiatives and gaining insight into the meaning of open educational practices (OEP) as an academic practice. The interviews with OER experts informed the design of the interview guides used with university leaders, staff, and academics, which were tailored to each group and tested before use.

A computer-assisted qualitative data analysis software package was used to manage and examine data sources (i.e., transcribed interviews, documents, Webpages) and involved descriptive and analytical coding stages, handling the texts interpretively, and focusing on core themes, emerging leitmotifs, and causal links (Miles & Huberman, 1994; Richards, 2009).

Quality Criteria

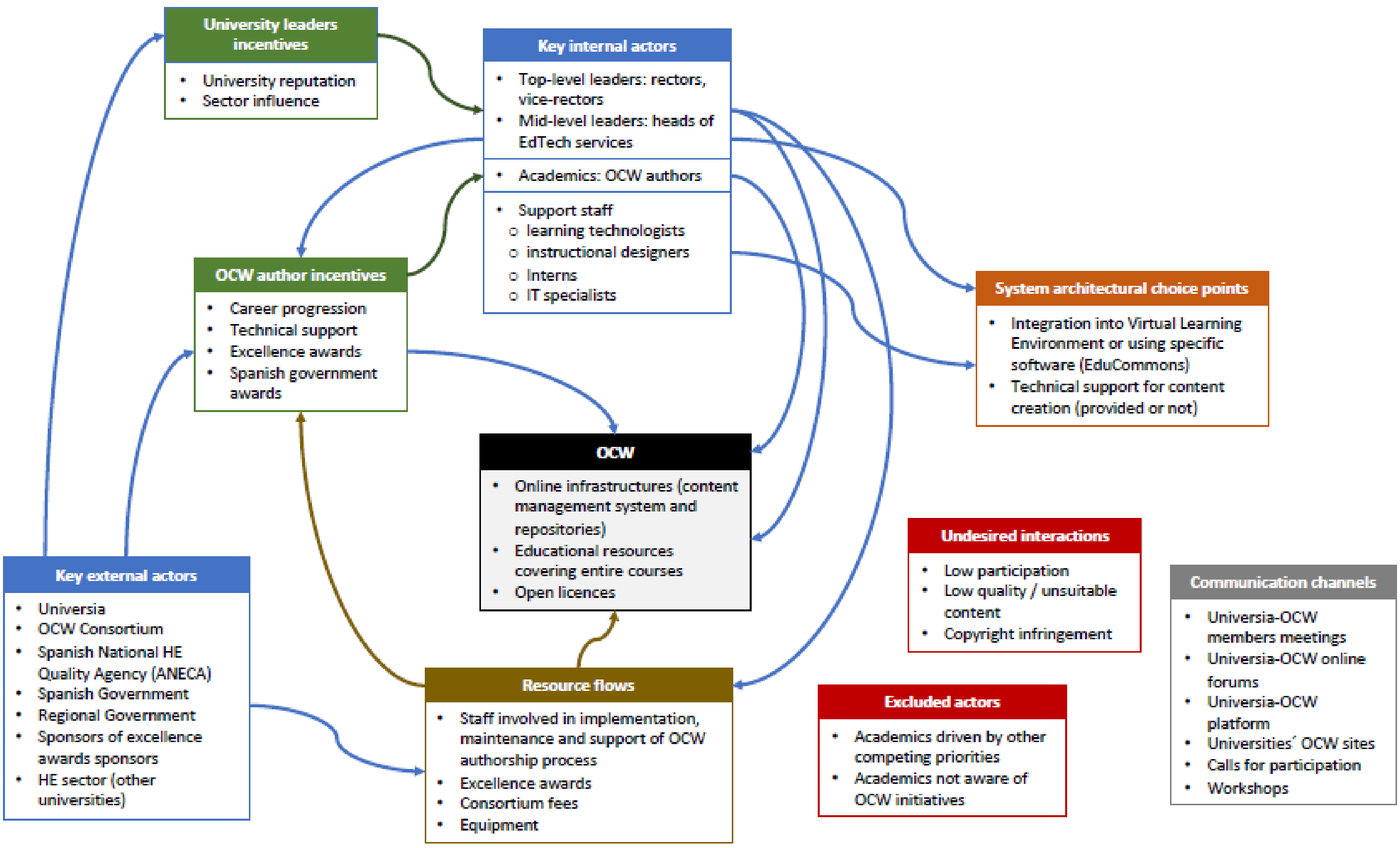
The study followed quality criteria for qualitative research to ensure rigour and credibility, mainly by grounding interpretive practices on multiple sources of evidence, establishing rapport with participants, returning to them for further clarifications when needed, keeping a clear line between their perspectives and researchers’ observations, and adopting a reflexive approach throughout the entire process of data collection and analysis (Tracy, 2010). Ethical approval was obtained before fieldwork.

Results

Based on the cross-comparative analysis of the three selected universities and their wider contexts, OCW initiatives were modelled as STINs, and the main potential elements and relationships at play are mapped out in Figure 3.

Figure 3

OCW Initiatives Modelled as Socio-Technical Interaction Networks



Note. Figure created by the author. EdTech = education and technology; OCW = OpenCourseWare; HE = higher education.

Despite the contextual similarities provided by the HE Spanish and regional systems, along with the requirements of the OCW model, there were considerable differences in the implementation of OCW initiatives across the three universities in the study. OCW was not embedded in the same ways and to the same extent into the preexisting institutional arrangements supporting EdTech at each of the three institutions. Table 2 summarises key differences.

Table 2

Key Differences in OpenCourseWare Implementation

Details	UNI-1	UNI-2	UNI-3
Launch date	2009	2008	2007
Key actors driving the implementation of OCW	Mid-level institutional leaders	Top-level institutional leaders	Top-level institutional leaders
OCW mentioned in top strategic documents	No	No	Yes
OCW formally included into EdTech plans	No	Yes	Yes
Certificates enabling recognition of OCW authoring in career progression	No	Yes	Yes
OCW excellence awards	Yes (only first year)	Yes (over two academic years)	No
Instructional design support to OCW authors	Yes (limited)	No	Yes (substantial)
Technology	Same as university virtual learning environment	Same as university virtual learning environment	Different from university virtual learning environment

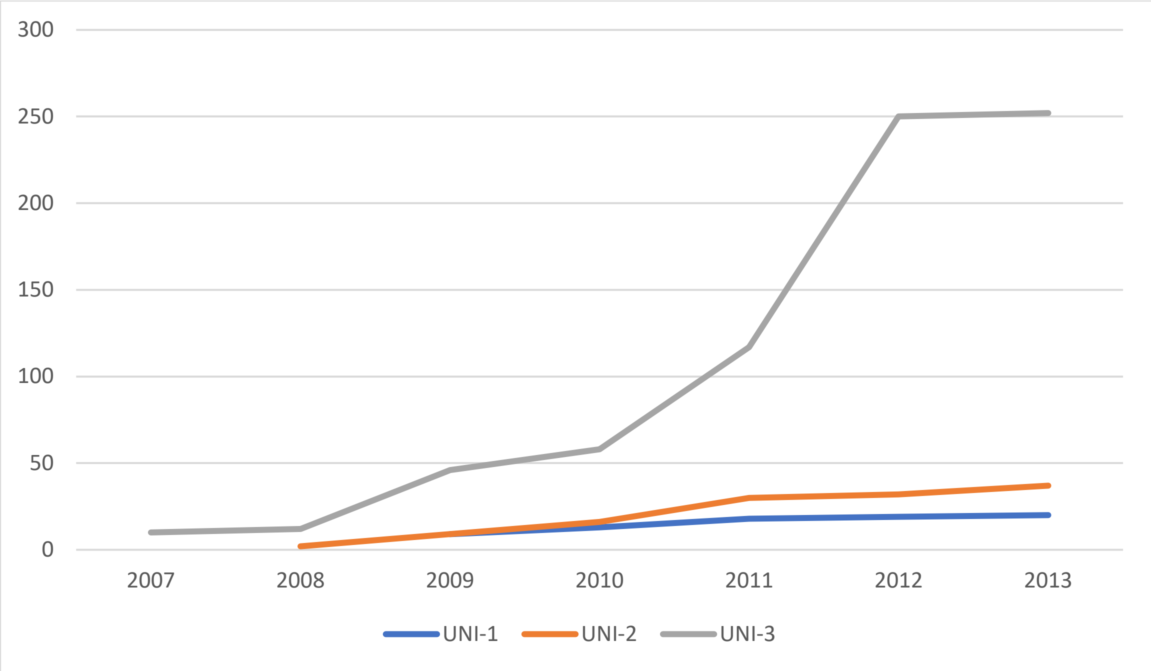
Note. OCW = OpenCourseWare; EdTech = education and technology.

OCW Production and Release

The release of OCW courses at the three universities grew fast during the period following the launch of each initiative, reaching a plateau within the first five years (Figure 4). Growth was negligible after 2013 for the OCW initiatives of both UNI-1 and UNI-3, the former being discontinued in 2017 (with 23 courses) and the latter in 2020 (with 253 courses). UNI-2 has released about 10 extra OCW courses since then, and, out of the three OCW sites, it is the only one still available. The creation of OCW modules tended to be one-off activities at the three universities, but some academics were credited as authors in more than one module—namely about 10% at both UNI-1 and UNI-2 and 20% at UNI-3.

Figure 4

Number of OpenCourseWare (OCW) Courses per University



Note. Figure created by the author. Source data collected from OCW sites of each institution and captures of them by the Wayback Machine (https://web.archive.org).

In the case of UNI-3, and to a lesser extent UNI-1, instructional designers were available to support the OCW authoring process. Indeed, the provision of that kind of support is key to understanding the large number of OCW courses produced at UNI-3, as any academic wishing to receive assistance in the creation of online resources in the academic year 2009–2010 was required to contribute the resulting content to the OCW initiative. Instructional design support was outsourced to an external company, and the number of OCW courses authored in 2009–2010 was so high that the courses had to be gradually released over the following years, due to limited capacity to process them. The end to instructional design support for OCW authors resulted in the release of just one new course after 2012:

> The regular release of OCW [courses] has been discontinued because the team of technicians was disbanded in June 2012. ... The programme depended on a vice-rector who is no longer in post after the appointment of our new rector, and nothing has replaced the actions that led to such a large number of OCW modules. (Senior learning technologist, UNI-3)

Integration of OCW into Institutional Arrangements

There were important differences across the three universities in terms of the extent to which OCW was present in strategic documents, plans, and support mechanisms relating to EdTech and educational innovation. While no key documents from UNI-1 mentioned OCW, both UNI-2 and UNI-3 included OCW in plans devoted to outlining priorities and support mechanisms in that regard. Moreover, top-level university leaders at UNI-3 mentioned OCW within documents defining their vision for the institution.

UNI-3 made OCW a core element in its teaching and learning innovation programme over a number of years. Most notably, in 2009–2010, access to instructional design was restricted to those willing and able to build their online learning resources for the virtual learning environment (VLE) as OCW courses. Over the next couple of years, the amount of support and incentives for OCW authoring at UNI-3 gradually decreased, and the innovation plan approved by a new leadership team in 2013 did not mention OCW at all.

OCW was also embedded into a programme aimed at supporting the creation of online learning resources at UNI-2. The authors of outstanding OCW courses were rewarded with an economic incentive in the 2008–2009 and 2009–2010 academic years, but coinciding with the appointment of a new leadership team, the allocation of budget for that specific purpose was discontinued. Certificates that could be used by academics for career progression purposes started to be issued to OCW authors at UNI-2 after the first year of activity, and they became the main way of fostering participation after economic incentives were discontinued.

The varying levels of importance from a strategic point of view also translated into significant differences in terms of resources allocated to operational costs and OCW authoring. For instance, while UNI-3 allocated substantial human resources to instructional design support and the implementation of required online infrastructures, UNI-1 and UNI-2 relied largely on the work of interns:

> We assigned too much responsibility to the intern, because we didn’t have any member of the staff available to do the job. So it was delayed. Indeed, there were [authors] who had finished their materials, but they were not available to the public on the Website. (Former top-level university leader, UNI-2)

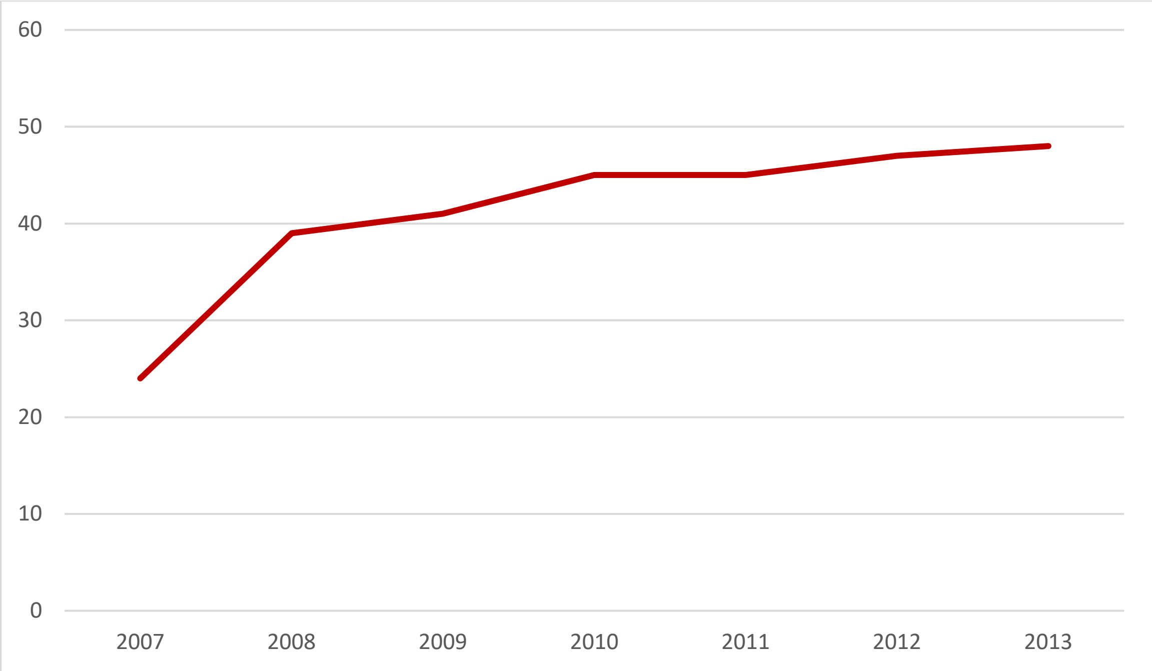
Optimising resources was also part of the rationale for UNI-1 and UNI-2 to use the same technical systems to run their OCW and VLE. While it required some tailoring to meet the requirements of the OCW model, that work could be done in-house as it was an open-source system, and both institutions had the required expertise. On the contrary, at UNI-3, it was not possible to integrate OCW with its proprietary VLE system, and a separate content management system was used instead.

Key Drivers to OCW Implementation

Universia had a strong influence on the implementation of OCW initiatives in Spain, especially among public universities. Most institutions in the OCW-Universia regional consortium joined it between 2007 and 2008, but the number kept growing until 2013 (Figure 5). In 2017, 40 public and 7 private Spanish universities were listed as members on the OCW-Universia site, which is no longer available.

Figure 5

Spanish Higher Education Institutions in OCW-Universia

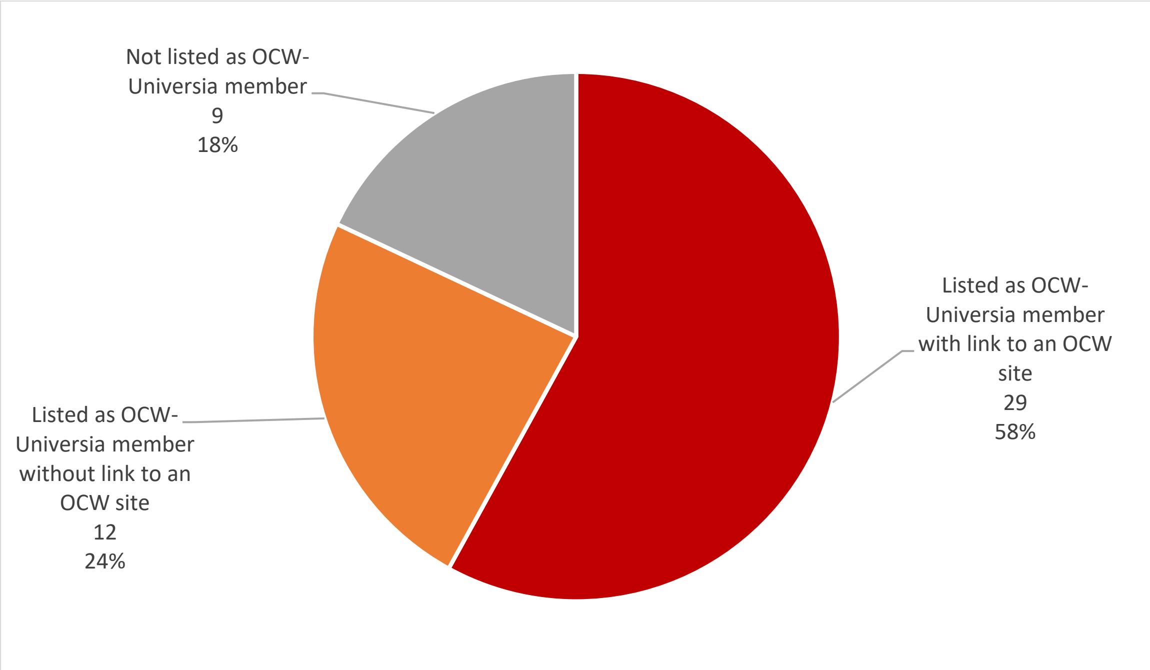


Note. Figure created by the author. Source data collected from the OCW-Universia site (http://ocw.universia.net) and captures of that website by the web.archive.org.

In order to join the OCW-Universia consortium, HE institutions were required to sign a memorandum of understanding and commit to implementing an OCW initiative with at least 10 courses in the first two years. However, not all of the member institutions completed the process of establishing their OCW sites In 2018, most public universities in Spain (more than 80%) were listed as members of the OCW-Universia on its Website, but only 29 active OCW sites (58% public universities) were linked from there (Figure 6).

Figure 6

Public Spanish Higher Education Institutions in OCW-Universia



Note. Figure created by the author. Source data collected from the OCW-Universia website, captured by the Wayback Machine (https://web.archive.org/web/20180126044851/http://ocw.universia.net/es/instituciones-integrantes-iberoamericanas-opencourseware.php). Total number of public higher education institutions in Spain provided by the Ministerio de Ciencia, Innovación y Universidades (2019). OCW = OpenCourseWare.

Top-level university leaders at both UNI-2 and UNI-3 had strong links with Universia, as indicated by the fact that representatives from both universities joined its board of directors around the time their institutions became affiliated with the OCW-Universia associate consortium. The decision for their universities to join the consortium must be understood in the context of the overall relationship between those institutions and the Universia organisation and network, as well as the Santander Group as its main sponsor. Besides the aims and values associated with the OCW initiative, cooperating with Universia was somehow a strategic movement on its own:

> There is a key factor in the implementation [of the OCW site]: the ongoing relationship between the university and the Santander Bank. The rector was one of the members of Universia’s board of directors, and Santander, through Universia, wanted [us] to implement that initiative. (Former top-level university leader, UNI-2)

The fact that so many Spanish universities had quickly joined OCW-Universia also operated as extra pressure, as not embracing OCW could be perceived as a failure to follow “cutting-edge” trends:

> It is also a bit like a domino effect. When you see someone else doing it you say: “Ah, that’s all right to do that, we’ll do it too.” Or because you get some complaints: “If X University is doing it, why is that you are not doing it?” (Top-level university leader, UNI-1)

OCW Authoring as an Academic Practice

The authoring of OCW courses does not necessary imply an understanding or even awareness of the basic principles of OER. While some interviewees were cognizant of the philosophical, legal, and technical aspects underpinning OCW as a type of OER, others simply equated OCW authoring with the creation on online learning resources. As an extreme example of this, one of the OCW authors interviewed at UNI-3 was disconcerted at realising during the interview that their OCW modules were accessible to people outside their university. Despite the scale of UNI-3's OCW and availability of instructional design support and training, the rationale behind OER provision was not clearly communicated to academics:

> It hasn't been [clearly] explained, saying, "Well, we will do either some technical training that includes a clear introduction to the context"—that is, what the values of the project are, to understand why we're doing this, [why] we're going to participate in this—or [even] simply some campaign. (Senior learning technologist, UNI-3)

Even though UNI-1 was the only participating university accompanying the launch of its OCW initiative with a series of workshops aimed at raising awareness of the historic, legal, technical, and philosophical implications of OER, dissemination efforts were rather limited:

> If you organise a few workshops but then you spend a whole year before circulating anything [again]or organising another workshop, then it is normal for people to think that [the project] is dead. Dissemination is very important and probably those other universities [that have been more successful in finding OCW authors] have done so. I must recognise that we haven't done enough publicity. Not just workshops, but sending e-mails, advertising it on the homepage of the university's Website. (Mid-level university leader, UNI-1)

OCW authors at the three institutions were particularly concerned about the creation of online learning resources for their students as a way to enhance learning and overcome the limitations of pedagogical approaches primarily based on academics giving lectures and students taking notes, which have historically dominated teaching and learning at Spanish universities. That level of dedication to authoring learning resources was not regarded by participants as a mainstream academic practice in Spain, but they considered it a core aspect of who they were as academics:

> I have created [learning] materials since I started. Always, always … whether I was working part-time or full-time … I've always elaborated my own [extended] notes. … It's absurd to have students taking notes during classes. … It looks to me like a waste of time. (Academic, UNI-2)

Likewise, academics in certain disciplines tend to base their teaching on original content created by them, and therefore amenable to be released under an open licence, while others heavily rely on the use of copyrighted materials (e.g., artworks). Indeed, the need to rely on copyrighted materials was perceived as an important barrier to OCW authoring. While OCW courses were largely a by-product of academics' ongoing work on the creation of learning resources, meeting the requirements of the OCW model required extra effort and was a clear barrier to participation: "The thing is that OCW does not require [creating] just a textbook, some notes. … It also involves creating exercises, self-assessment activities, etc. … Developing quizzes for self-assessment requires time too" (Academic, UNI-1).

Another impediment mentioned by participants was the perception that sharing teaching aids and other educational resources among peers was at odds with the prevailing academic culture at Spanish universities:

> Lecturers are very protective of the materials that they have and [the content] they deliver in their classes. … I think that it is something cultural, not to meddle in what other colleagues do, because it generates tensions and conflicts. These are sensitive issues. (Academic, UNI-3)

While participants considered the creation of learning resources as an important academic practice, they also recognised that its value for career progression purposes and appraisal was limited. Some felt that their excessive commitment to teaching and creating nonstandard scholarly publications (e.g., blogs, Websites, etc.) had a negative impact on their own opportunities for career development. This was particularly constraining to early-career academics:

> Any young academic who doesn't have a permanent position is currently listening to the following message: "You have to get accredited; you must publish a lot of research work." It's quite clear what she needs to do if she wants to have access to a [permanent] position in a reasonable time frame. Probably not to create OCW courses. … The thing is that within this population that cannot devote their time [to the creation of OER] there are more people convinced that this is the future of the university. But they have to combine the future of the university and their own future. (Former top-level university leader, UNI-2)

Even though the value attached to the creation of online learning resources for career progression—and more generally, teaching as compared to research—was relatively low, the effort of adapting content to the OCW requirements was more worthwhile at UNI-2 and UNI-3 thanks to the issuing of certificates to OCW authors:

> What they always say when you are planning to go for an accreditation is that you have to try and tick all the boxes, not to leave any gaps. … Everyone knows that research is the most important section. As for the teaching section, provided you have taught enough credits—there is a minimum—[you are fine], then the rest is just complementary. … You may meet the requirements of the teaching section just with your classes. … Innovation in teaching is far from being a decisive factor [but still counts]. (Academic, UNI-2)

While most senior academics would not themselves benefit directly from that kind of recognition, one motivation for some was to help younger colleagues by coauthoring OCW courses with them. Other motivations apart from direct incentives (e.g., support, certificates) were social pressure (e.g., being invited personally) or involvement in the implementation of OCW and having to therefore set an example.

Discussion

The findings from this case study are consistent with the cumulative body of knowledge generated in the social informatics literature (Meyer et al., 2019). In particular, this study uncovered some of the main difficulties that may arise when an EdTech initiative originally designed for a specific institution, namely OCW-MIT, is extrapolated to significantly different contexts. In doing so, it approached EdTech as socio-technical arrangements formed by a diverse range of interrelated elements, including artefacts (such as hardware and software), people, roles, values, practices, norms, and protocols (Kling, 2007).

This study’s results are also in line with the results of previous research on OEP that highlight the importance of institutional arrangements and cultures (Cachia et al., 2020; Cox, 2013; Hatakka, 2009; King et al., 2018) and complement, with qualitative insights, the results from previous studies on OCW in Spain (Frías-Navarro et al., 2010; Tovar et al., 2013).

An analysis of the meaning and value attached to OCW authoring as an academic practice across the institutions in this study revealed key aspects influencing the dispositions of academics towards OEP uptake, the long-term sustainability of those initiatives, and the more or less successful ways of fostering participation.

The analysis revealed five main differences between OCW-MIT and the OCW initiatives examined in this case study:

1. While the design and implementation of OCW-MIT was informed by a consultation with academics, the adoption of that model of OER provision at the Spanish universities in this case study followed a primarily top-down approach and did not draw on the views from the communities expected to participate as OCW authors.

2. A culture of creating learning resources and sharing with peers was already established at MIT before OCW was devised.

3. Unlike at MIT, the main driver in implementing OCW at Spanish institutions was an external actor, namely Universia, instead of an internal process aimed redefining their strategic approach towards lifelong learning.

4. MIT institutionalised OCW as a core component of its strategic orientation towards lifelong learning, ensuring its long-term continuity. By contrast, the sustainability of OCW at the universities in this case study was highly dependent on the personal support from members of their leadership teams, putting its continuity at risk when changes were made.

5. MIT allocated significant resources to permanently support OCW authoring in order to minimise the extra efforts required from academics.

Overall, a disconnect was observed between the drivers and rhetoric behind the implementation of OCW initiatives and the motivations for academics to create OCW courses. Different motives may intersect at the creation and release of OER, often resulting in tensions and conflicting priorities (Falconer et al., 2016). As noted by Selwyn (2013, p. 82), “it may be argued that the likelihood of teachers and learners involved in the production of open software or content is curtailed by the realities of institutionalized education—not least issues of time, technical expertise, interest and motivation.” Even when there is appetite to create OER, the day-to-day demands of established academic practices—including a host of administrative tasks—leave little room for doing so.

The lack of incentives is a clear barrier to OCW authoring in Spain (Frías-Navarro et al., 2010), and, beyond the particularities of each of the three selected universities and their own institutional arrangements, the value—for career progression—of producing online learning resources was minimal, as determined by performance criteria established at a national level for the entire Spanish HE system. This means that academics had to prioritise other kinds of academic activities, mainly publishing research, over OCW authoring if they were to advance in their careers.

However, the two universities in the case study that provided OCW authors with a certificate enabling them to get their work recognised in accreditation processes—even if the value was minimal—achieved higher levels of participation. Academics aspiring to be promoted to more senior positions would benefit from prioritising other types of outputs; but at the same time, an OCW certificate was more valuable to them than it was to senior colleagues. Beyond career progression implications, making access to instructional design support for the creation of online learning materials conditional to the release of the resulting resources as OCW courses made the biggest difference in terms of participation levels across the three universities. That type of support was key to the implementation of OCW-MIT and has been highlighted in previous studies as an important enabler of OER authoring (Henderson & Ostashewski, 2018; Ros et al., 2014; Wei & Chou, 2021).

Conclusions and Recommendations

Using three OCW initiatives in Spain as a case study, this article sheds light on the process of implementing a highly predefined model of OER provision within institutional contexts that are considerably different from where it was originally devised, namely MIT. The study reveals a major disconnect between the drive to implement the initiatives and the realities of adopting OCW authoring as an academic practice. Even though the availability of support and the issuing of certificates—which academics may use for career progression purposes—proved to be valuable mechanisms used by some of the universities as enablers of OCW authoring, the overall incentive to invest time and effort in OER provision was minimal due to the accreditation and performance criteria defined at a national level.

An important recommendation for university leaders and policy makers wishing to promote the authoring and release of OER is to devise models that take into consideration the specificities of their institutions and the wider HE system in which they operate, paying particular attention to ways of making participation a valuable academic practice in relation to other competing priorities. For those considering the implementation of a highly defined model of OER provision conceived elsewhere—such as OCW—it is advisable to invest time first into assessing the readiness of their communities for the uptake of the proposed OEP (Wei & Chou, 2021) and carefully examining their institutional contexts from a socio-technical perspective.

The implementation of OER initiatives, including OCW, and related learning opportunities, such as massive open online courses, is often driven by the honourable goal of widening access to education and enabling development opportunities and lifelong learning for all; however, despite good intentions, it may instead lead to reinforcing knowledge gaps and social inequalities (Knox, 2013; Rohs & Ganz, 2015) by wrongly putting too much emphasis on individual agency over social structures (Eynon & Malmberg, 2021). Therefore, university leaders, policy makers, and advocates should take a step back to consider the ultimate goal they want to achieve in relation to their institutions’ missions before deciding what type of OER initiative is the most suitable to implement or promote, and even whether OER provision is the best way of doing so.

Acknowledgement

The author acknowledges funding from Talentia Excellence Grants (Junta de Andalucia), the Oxford Internet Institute, the University of Oxford’s Santander Academic Travel Award and the Carr & Stahl Fund (St Antony’s College, University of Oxford) for the PhD research on which this paper is based. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Abelson, H. (2008). The creation of OpenCourseWare at MIT. *Journal of Science Education and Technology*, *17*(2), 164–174. https://doi.org/10.1007/s10956-007-9060-8

Abelson, H., Miyagawa, S., & Yue, D. K. P. (2021, May 24). On the 20th anniversary of OpenCourseWare: How it began. *MIT Faculty Newsletter*, *33*(5). https://fnl.mit.edu/may-june-2021/on-the-20th-anniversary-of-opencourseware-how-it-began/

Andrews, R., & Haythornthwaite, C. (Eds.). (2007). *The SAGE handbook of e-learning research*. Sage Publications. https://us.sagepub.com/en-us/nam/the-sage-handbook-of-e-learning-research/book243430

Aranzadi, P., & Capdevila, R. (2011). Open Course Ware, recursos compartidos y conocimiento distribuido. *La Cuestión Universitaria*, *7*, 126–133. http://polired.upm.es/index.php/lacuestionuniversitaria/article/view/3398

Borrás Gené, O. (2010). *Observatorio de plataformas para OCW*. Universidad Politécnica de Madrid.

Bryman, A. (2008). *Social research methods* (3rd ed.). Oxford University Press.

Cachia, R., Aldaoud, M., Eldeib, A., Hiari, O., Tweissi, A., Villar-Onrubia, D., Wimpenny, K., & Maya, I. (2020). Cultural diversity in the adoption of open education in the Mediterranean basin: Collectivist values and power distance in the universities of the Middle East. *Araucaria. Revista Iberoamericana de Filosofía, Política, Humanidades y Relaciones Internacionales*, *22*(44), 53–82. https://doi.org/10.12795/araucaria.2020.i44.03

Carson, S. (2009). The unwalled garden: Growth of the OpenCourseWare Consortium, 2001–2008. *Open Learning: The Journal of Open and Distance Learning*, *24*, 23–29. https://doi.org/10.1080/02680510802627787

Carson, S., & Forward, M. L. (2010). Development of the OCW Consortium. In *2010 IEEE Education Engineering (EDUCON)* (pp. 1657–1660). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/EDUCON.2010.5492401

Clifton, J., Díaz Fuentes, D., Fernández Gutiérrez, M., & Revuelta López, J. (2013). Difusión del proyecto OpenCourseWare—Universidad de Cantabria en asignaturas relativas a la economía mundial. In M. A. Bringas Gutiérrez, M. Fernández Gutiérrez, & M. Fernández Redondo (Eds.), *XV Reunión de Economía Mundial / Sociedad de Economía Mundial* (pp. 5–10). Universidad de Cantabria. http://hdl.handle.net/10902/4425

Cox, G. (2013). Researching resistance to open education resource contribution: An activity theory approach. *E-Learning and Digital Media*, *10*(2), 148–160. https://doi.org/10.2304/elea.2013.10.2.148

Creanor, L., & Walker, S. (2012). Learning technology in context: A case for the sociotechnical interaction framework as an analytical lens for networked learning research. In L. Dirckinck-

Holmfeld, V. Hodgson, & D. McConnell (Eds.), *Exploring the theory, pedagogy and practice of networked learning* (pp. 173–187). Springer. https://doi.org/10.1007/978-1-4614-0496-5_10

Creswell, J. W. (2007). *Qualitative inquiry & research design: Choosing among five approaches* (2nd ed.). Sage Publications. https://psycnet.apa.org/record/2006-13099-000

Denzin, N. K., & Lincoln, Y. S. (2011). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (4th ed., pp. 1–19). Sage Publications. https://www.sagepub.com/sites/default/files/upm-binaries/40425_Chapter1.pdf

Eynon, R., & Malmberg, L.-E. (2021). Lifelong learning and the Internet: Who benefits most from learning online? *British Journal of Educational Technology*, *52*(2), 569–583. https://doi.org/10.1111/bjet.13041

Falconer, I., Littlejohn, A., McGill, L., & Beetham, H. (2016). Motives and tensions in the release of open educational resources: The UKOER program. *Australasian Journal of Educational Technology*, *32*(4), Article 4. https://doi.org/10.14742/ajet.2258

Frías-Navarro, M. D., Pascual Llobell, J., Monterde i Bort, H., Pascual Soler, M., Badenes Ribera, L., & Pascual Mengual, J. (2010). *Impacto del Open Course Ware (OCW) en los docentes universitarios*. Universidad de Valencia/Ministerio de Educación. https://perma.cc/H49P-UFTZ

Gallardo Paúls, B. (2008). Teaching and open access: The Open Course Ware of the Universitat de València. *@tic. Revista d'innovació Educativa*, *1*, 16–25. http://ojs.uv.es/index.php/attic/article/view/45

Hatakka, M. (2009). Build it and they will come?—Inhibiting factors for reuse of open content in developing countries. *The Electronic Journal of Information Systems in Developing Countries*, *37*(1), 1–16. https://doi.org/10.1002/j.1681-4835.2009.tb00260.x

Henderson, S., & Ostashewski, N. (2018). Barriers, incentives, and benefits of the open educational resources (OER) movement: An exploration into instructor perspectives. *First Monday*, *23*(12). https://doi.org/10.5210/fm.v23i12.9172

King, M., Pegrum, M., & Forsey, M. (2018). MOOCs and OER in the Global South: Problems and potential. *The International Review of Research in Open and Distributed Learning*, *19*(5). https://doi.org/10.19173/irrodl.v19i5.3742

Kling, R. (2007). What is social informatics and why does it matter? *The Information Society*, *23*(4), 205–220. https://doi.org/10.1080/01972240701441556

Kling, R., McKim, G., & King, A. (2003). A bit more to it: Scholarly communication forums as socio-technical interaction networks. *Journal of the American Society for Information Science and Technology*, *54*(1), 47–67. https://doi.org/10.1002/asi.10154

Knox, J. (2013). Five critiques of the open educational resources movement. *Teaching in Higher Education*, *18*(8), 821–832. https://doi.org/10.1080/13562517.2013.774354

Kursun, E., Cagiltay, K., & Can, G. (2014). An investigation of faculty perspectives on barriers, incentives, and benefits of the OER movement in Turkey. *The International Review of Research in Open and Distributed Learning*, *15*(6). https://doi.org/10.19173/irrodl.v15i6.1914

Lerman, S. R., & Miyagawa, S. (2002). OpenCourseWare: A case study in institutional decision making. *Academe*, *88*(5), 23–27. https://eric.ed.gov/?id=EJ655807

Lerman, S. R., Miyagawa, S., & Margulies, A. H. (2008). OpenCourseWare: Building a culture of sharing. In T. Iiyoshi & M. S. V. Kumar (Eds.), *Opening up education: The collective advancement of education through open technology, open content, and open knowledge* (pp. 213–227). MIT Press. https://mitpress.mit.edu/books/opening-education

Llorens, F., Bayona, J. J., Gómez, J., & Sanguino, F. (2010). The University of Alicante's institutional strategy to promote the open dissemination of knowledge. *Online Information Review*, *34*(4), 565–582. https://doi.org/10.1108/14684521011072981

Llorente Cejudo, M. del C., Cabero Almenara, J., Vázquez Martínez, A. I., & Alducín Ochoa, J. M. (2013). *Proyecto OpenCourseWare y su implantación en universidades andaluzas*, *12*(1), 11–21. http://dehesa.unex.es/handle/10662/724

Lloyd, M. (2011, July 10). Billion-dollar benefactor for campuses in Latin America: A Spanish bank. *The Chronicle of Higher Education*. https://www.chronicle.com/article/billion-dollar-benefactor-for-campuses-in-latin-america-a-spanish-bank/

McCoy, C., & Rosenbaum, H. (2019). Uncovering unintended and shadow practices of users of decision support system dashboards in higher education institutions. *Journal of the Association for Information Science & Technology*, *70*(4), 370–384. https://doi.org/10.1002/asi.24131

Meyer, E. T., Shankar, K., Willis, M., Sharma, S., & Sawyer, S. (2019). The social informatics of knowledge. *Journal of the Association for Information Science and Technology*, *70*(4), 307–312. https://doi.org/10.1002/asi.24205

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Sage Publications. https://psycnet.apa.org/record/1995-97407-000

Ministerio de Ciencia, Innovación y Universidades. (2019). *Datos y cifras del sistema universitario Español 2018-2019 (Data and figures of the Spanish university system 2018-2019).* Gobierno de España. https://www.educacionyfp.gob.es/dam/jcr:2af709c9-9532-414e-9bad-c390d32998d4/datos-y-cifras-sue-2018-19.pdf

Oliver, M. (2011). Technological determinism in educational technology research: Some alternative ways of thinking about the relationship between learning and technology. *Journal of*

Computer Assisted Learning, *27*(5), 373–384. https://doi.org/10.1111/j.1365-2729.2011.00406.x

Orlikowski, W. J., & Scott, S. V. (2008). Sociomateriality: Challenging the separation of technology, work and organization. *The Academy of Management Annals*, *2*(1), 433–474. https://doi.org/10.1080/19416520802211644

Peter, S., & Deimann, M. (2013). On the role of openness in education: A historical reconstruction. *Open Praxis*, *5*(1), 7–14. https://doi.org/10.5944/openpraxis.5.1.23

Richards, L. (2009). *Handling qualitative data: A practical guide* (2nd ed.). Sage Publications.

Rohs, M., & Ganz, M. (2015). MOOCs and the claim of education for all: A disillusion by empirical data. *The International Review of Research in Open and Distributed Learning*, *16*(6). https://doi.org/10.19173/irrodl.v16i6.2033

Ros, S., Hernández, R., Read, T., Rodrı, M., Pastor, R., & Dı, G. (2014). UNED OER experience: From OCW to open UNED. *IEEE Transactions on Education*, *57*(4), 248–254. https://doi.org/10.1109/TE.2014.2331216

Santos-Hermosa, G., Estupinyà, E., Nonó-Rius, B., Paris-Folch, L., & Prats-Prat, J. (2020). Open educational resources (OER) in the Spanish universities. *Profesional de la Información*, *29*(6), Article 6. https://doi.org/10.3145/epi.2020.nov.37

Selwyn, N. (2013). *Distrusting educational technology: Critical questions for changing times*. Routledge. https://www.routledge.com/Distrusting-Educational-Technology-Critical-Questions-for-Changing-Times/Selwyn/p/book/9780415708005

Stake, R. E. (1995). *The art of case study research*. Sage Publications. https://us.sagepub.com/en-us/nam/the-art-of-case-study-research/book4954

Tovar, E., López, J., Piedra, N., Sancho, E., & Soto, Ó. (2013). Aplicación de tecnologías Web emergentes para el estudio del impacto de repositorios OpenCourseWare españoles y latinoamericanos en le Educación Superior. Universidad Politécnica de Madrid & Gobierno de España. http://ocw.uc3m.es/informe-final-ocw.pdf/at_download/file

Tracy, S. J. (2010). Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qualitative Inquiry*, *16*(10), 837–851. https://doi.org/10.1177/1077800410383121

UNESCO. (2002). *Forum on the impact of open courseware for higher education in developing countries—Final report*. http://unesdoc.unesco.org/images/0012/001285/128515e.pdf

Villar-Onrubia, D. (2014). The value of Web mentions as data: Mapping attention to the notion of OER in the HE arena. In M. Sánchez & E. Romero Frías (Eds.), *Ciencias sociales y humanidades digitales técnicas, herramientas y experiencias de e-research e investigación en colaboración* (pp. 163–181). Sociedad Latina de Comunicación Social. https://ora.ox.ac.uk/objects/uuid:0aac8efd-f100-4a06-9fe5-7a1d849120bf

Wei, H.-C., & Chou, C. (2021). Ready to do OpenCourseWare? A comparative study of Taiwan college faculty. *The International Review of Research in Open and Distributed Learning*, *22*(2), 118–142. https://doi.org/10.19173/irrodl.v22i2.5252

White, S., White, S., & Borthwick, K. (2020). Blended professionals, technology and online learning: Identifying a socio-technical third space in higher education. *Higher Education Quarterly*, *75*(1), 161–174. https://doi.org/10.1111/hequ.12252

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Sage Publications. https://www.worldcat.org/title/case-study-research-design-and-methods/oclc/50866947

Athabasca University

BY

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

Understanding the Relationship Among Self-Efficacy, Utility Value, and the Community of Inquiry Framework in Preservice Teacher Education

Mete Akcaoglu[1] and Mustafa Ozturk Akcaoglu[2]
[1]Georgia Southern University, [2]Kastamonu University

Abstract

School closures during the COVID-19 pandemic have shown the importance of distance education, and teachers have been tasked with designing and delivering online courses in a short amount of time without much preparation or deliberation. As the future generation of teachers, preservice teachers need to be prepared to teach online, and their motivation to do so is a key factor in how successfully they do it. The community of inquiry framework provides researchers and practitioners with a framework for designing and delivering online courses, while self-efficacy and utility value are important motivational constructs predicting future engagement and success in tasks. In this cross-sectional survey study, we investigated preservice teachers' (n = 344) perceptions of their self-efficacy, utility value, the importance of the three components of the community of inquiry framework: teaching presence, social presence, and cognitive presence. Our results show that overall, preservice teachers had high motivation to teach online and high perceptions of the three presences. Our regression analyses indicated that while preservice teachers' self-efficacy was a significant predictor of teaching presence, utility value only significantly predicted social presence. We discuss the implications of these findings for teacher education programs, including a holistic approach to teaching online learning and instructional design.

Keywords: distance education, teacher education, community of inquiry, self-efficacy, utility value

Understanding the Relationship Among Self-Efficacy, Utility Value, and the Community of Inquiry Framework in Preservice Teacher Education

Despite its long history, it was not until the COVID-19 pandemic that distance education became the primary mode of education for almost all educational institutions around the world. Before the pandemic, countries preferred traditional, in-person instruction. Particularly in K–12 settings, computer-supported distance learning was rarely used. As a result, most institutions did not have the required experiences and preparations to develop and deliver effective online learning experiences during the school closures of 2020. Furthermore, it has been revealed that neither learners nor teachers were fully prepared in terms of individual efficacies and using technological hardware and software facilities to teach and learn online (Mishra et al., 2020; Rapanta et al., 2020).

Given the increasing importance of teaching online, the future generation of teachers will be required to engage in designing and creating effective online learning environments. This necessitates that they are introduced to pedagogical and design-related aspects of online learning during their preservice education. Therefore, teacher education programs along with professional development activities carry the responsibility of preparing future and current teachers for teaching and learning online.

Previous studies regarding online learning support the community of inquiry (CoI) as a well-founded theoretical framework to understand the process and planning online learning in line with both instructors' and learners' experiences, interests, and needs (Garrison & Akyol, 2013; Garrison & Arbaugh, 2007). CoI has been one of the most used and cited theoretical frameworks in research on online teaching and distance education in the last decade (Bozkurt et al., 2015; Kim & Gurvitch, 2020; Valverde-Berrocoso et al., 2020). According to the CoI framework, there are three main components of regulating and preserving the effectiveness of online learning in educational settings: teaching presence, social presence, and cognitive presence (Garrison et al., 1999). The framework has been frequently tested in research studies focusing on online learning to improve students' learning experiences (Burgess et al., 2010; Garrison et al., 2010; Kazanidis et al., 2018; Rubin et al., 2013).

For both preservice and in-service teachers, motivation for effectively integrating technology is as essential as having the skills required for effective teaching (Ertmer et al., 2012). Creating and teaching online courses also operate on similar principles. To this end, one of the prominent motivation theories, expectancy–value theory (Wigfield, 1994; Wigfield & Eccles, 2000), has been used as a framework to understand individuals' task choices and success in those tasks. According to this theory, one's belief that they can do a task (i.e., expectancies) and the value they place on the task (e.g., utility value) are predictors of their success in the task (Wigfield et al., 2004). Such a task may be, for students, success in coursework or, for teachers, integration of technology or ability to teach online.

Aiming to understand both the underlying motivational processes and perceptions of the CoI framework, our purposes in this study were to (a) investigate the perceptions of preservice teachers in terms of their approaches to online teaching from the CoI perspective, and (b) examine the relationship between the components of CoI and some key motivational factors that influence preservice teachers' perceptions.

Background

The Community of Inquiry Framework

The CoI framework places community, critical thinking, and knowledge construction at the center of learning, especially in the online learning process (Garrison & Archer, 2000). The framework is based on Dewey's progressive education approach and is built on the social constructivist perspective (Kim & Gurvitch, 2020). Dewey (1959) thought that educational experiences should serve the common interests of the individuals and society, that individual development depends on the community, and that learners' inquiry process is at the center of educational experiences. Dewey viewed cooperative learning, constructivism, and practical inquiry as at the heart of the CoI framework; these are thought to guide the theory and practice to be used during the online learning process. It should be noted, however, that due to this specific pedagogical and epistemological emphasis, there may be situations where following the CoI framework may not be feasible or necessary.

The three components of the CoI framework (i.e., social, cognitive, and teaching presence) are based on experiences and enhance the quality of online learning (Garrison et al., 1999). Social presence includes affective expression, open communication, and group cohesion. It focuses on "the ability of participants to identify with the group or course of study, communicate purposefully in a trusting environment, and develop personal affective relationships progressively by way of projecting their individual personalities" (Garrison, 2009, p. 352). Social presence also focuses on the communication skills of learners and supports the promotion of a collaborative learning environment (Akyol & Garrison, 2011). It is regarded as a mediating variable between the other two components of the CoI (Garrison et al., 1999, 2010).

Cognitive presence refers to "the extent to which learners are able to construct and confirm meaning through sustained reflection and discourse in a critical community of inquiry" (Garrison et al., 2001, p. 11). Its focus is on students' development of meaningful knowledge and centers on four phases: (a) a triggering event, (b) exploration, (c) integration, and (d) resolution (Garrison et al., 2001). A triggering event can be the identification of a problem that requires extra inquiry; exploration involves critical reflection and discourse to investigate an issue; integration means to construct meaning based on the explored ideas; and resolution denotes applying the recently developed knowledge to the school environment.

Finally, teaching presence focuses on "the design, facilitation, and direction of cognitive and social processes for the purpose of realizing personally meaningful and educationally worthwhile learning outcomes" (Anderson et al., 2001, p. 5). According to Anderson et al. (2001), teaching presence has three subdimensions: (a) instructional design and organization, (b) facilitation of discourse, and (c) direct instruction. Research results show that teaching presence is necessary for creating and sustaining the CoI environment (Anderson et al., 2001; Joo et al., 2011; Pardo & Peñalvo, 2008; Pecka, 2014; Van Niekerk, 2015).

Self-Efficacy

Self-efficacy refers to one's perceptions or beliefs about one's perceived ability to learn or fulfill tasks at certain levels and an individual's belief in successfully performing a task related to learning or practice (Bandura, 1986). Studies examining the relationship between self-efficacy and academic achievement have revealed that self-efficacy predicts academic achievement and that individuals with high self-efficacy are more inclined to perform tasks, are more determined, and work harder (Ferede et al., 2016; Valentine et al., 2004; Vogel & Human-Vogel, 2016; Wang & Finch, 2018).

The sources of self-efficacy imply that experiences, both mastery and vicarious, play a crucial role in the formation of self-efficacy beliefs (Bandura, 1997). While mastery experiences are related to the gains we make when we take on a new task and successfully complete it, vicarious experiences are those in which self-efficacy is achieved by observing and imitating a role model who accurately completes a specific task. Therefore, self-efficacy alone can be thought of as both a cause and an effect, which is changed and affected by the educational experiences and collaborative environment in a community of inquiry (Akyol & Garrison, 2011).

Self-Efficacy in Online Teaching and Learning

In the online learning process, self-efficacy is an important element that encourages productive and self-directed learning while also contributing to learners overcoming the effect of being alone (Hodges, 2008; Ponton et al., 2005; Song & Hill, 2007). Furthermore, teacher self-efficacy is an essential variable in explaining the integration of technology in classroom activities (Kwon et al., 2019). For this reason, self-efficacy might be considered a prerequisite for success in online learning environments (Taipjutorus et al., 2012; Yavuzalp & Bahcivan, 2020). In addition, high self-efficacy is closely related to feeling able to work independently and able to self-regulate a learning process, which is very important in online learning environments (Busch, 1996; Putarek & Pavlin-Bernardić, 2020).

Self-efficacy in the context of online learning also has an essential role in determining students' confidence level to accomplish learning tasks. Therefore, we think that self-efficacy, required to describe and identify active and successful learners, could be a very important component for the development of a theoretical framework for online education, especially in the absence of a traditional classroom environment.

In each of the CoI framework's components, along with psychological features of learners (e.g., attitudes, efficacy, and motivation) and sociological aspects (e.g., collaboration and interaction), there are experiences Dewey (1986) advocates as the roots of learning and Bandura (1997) shows as the source of self-efficacy. Both psychologists have emphasized the importance of experiences and interaction in the learning process. With an emphasis on learning by doing and living, Dewey advocated the same thoughts as Bandura about students' experiencing and interacting with a concept so that they could learn.

A relationship exists between teacher self-efficacy and the intention to use technology (Joo et al., 2018; Park, 2009; Teo & Zhou, 2014; Valtonen et al., 2015). Similarly, research results (e.g., Anderson et al., 2011) have shown that preservice teachers' beliefs regarding the importance of using technology in the classroom significantly predicted their intention to use technology. Researchers also state that the intention to use information and communication technologies is positively affected by the self-efficacy of preservice teachers (Joo et al., 2018; Valtonen et al., 2015). Based on prior research, in this study, we propose that self-efficacy can function as an antecedent for supporting the components in the CoI framework.

Utility Value

As a component of the expectancy–value theory (Eccles, 1983), utility value refers to the value of a task in terms of its usefulness for one's future life. Utility value or the prospective relevance of a task can be in the form of, for example, a course's usefulness for a student's future career plans (Hulleman & Harackiewicz, 2009). These real-life connections may not be readily visible to individuals; therefore, support might be needed for them to find and understand these connections (Hulleman et al., 2017).

In the context of online teaching, especially considering the experiences of preservice teachers in their formal educational experiences, it may not always be possible for them to seek and understand the relevance of online teaching skills for their future teaching. Notably, learning about distance education conceptually may not also mean developing perceived value and interest in it to engage with this task in the future. In the context of online learning, we do not know how the utility value of distance learning (a) varies among preservice teachers and (b) relates to the CoI framework's specific components. Therefore, in this study, we also investigated the relationship between utility value and preservice teachers' perceptions toward online teaching in the context of CoI.

The Present Study

Given the increasing importance of distance education, in this study, our purpose was to investigate the perceptions of future teachers toward teaching online from a CoI framework perspective. Therefore, we first descriptively investigated the perceptions of preservice teachers:

1. What are preservice teachers' self-efficacy and utility value beliefs about distance education?
2. What do the preservice teachers feel about the three components of the CoI framework?

Since expectancies and value—utility value in specific—are strong predictors of future engagement and success, we were also interested in investigating their relationship with the preservice teachers' future distance teaching perspectives:

1. Does self-efficacy predict the preservice teachers' perceptions toward CoI components?
2. Does utility value predict the preservice teachers' perceptions toward CoI components?

Method

In order to answer our research questions, we conducted a cross-sectional survey study.

Participants

The participants in this study were teacher education students studying Extra-Curricular Activities in Education and Principles and Methods of Teaching courses at a midsized public university in the Western Black Sea Region of Turkey during fall 2020. A total of 360 students participated in the survey.

We identified outliers by creating a variable that calculated the mean of all items for each student. Outliers with a score of 4.8 and above (n = 16) were removed from the analyses as this score indicated that these students elected to choose the highest score for almost all survey items regardless of the question (i.e., maximum Likert scale option was 5). The analyses were conducted with the remaining 344 students (251 female, 93 male). The students came from various teacher education programs. Table 1 shows the distribution of students across different programs.

Table 1

Participants' Distribution Across Majors

Program	n	%
Art and Crafts Education	10	2.9
Computer Education and Instructional Technologies	17	4.9
Guidance and Psychological Counseling	90	26.2
Elementary Education	12	3.5
Elementary Mathematics Education	24	7
Early Childhood Education	67	19.5
Religious Culture and Ethics Education	65	18.9
Science Education	8	2.3
Social Studies Education	22	6.4
Turkish-Language Teaching	27	7.9
Missing	2	0.6

The majority of students were in their second year (54%), while 34% were in their third year. First- and fourth-year (11%) students represented a total of 12% of the participants. The age average of the participants was 21.7 (SD = 3.2).

Instruments and Measures

Participants' perceived utility value was measured by an adapted version of Hulleman et al. (2017)'s utility value scale (Example scale item: "The material in this class is useful in my everyday life"). These items were adapted for preservice teachers' perceptions of how distance education can be relevant for their future teaching careers (Example scale item: "When I become a teacher, knowing about distance education will be useful"). The reliability calculated from our data of the scale was high: Cronbach's alpha (α) = .85.

We measured participants' self-efficacy using patterns of adaptive learning scales (PALS) by Midgley et al. (2000). More specifically, we adapted the items in the academic efficacy scale (Example scale item: "I'm certain I can master the skills taught in class this year") so that they would apply to our participants' future teaching using distance education (Example scale item: "When I start teaching, I am certain I can master the necessary skills for distance education"), similar to the utility value items. The reliability calculated from our data of the scale was high: α = .88.

The participants' CoI perceptions were measured via an adapted version of a survey created by Arbaugh et al. (2008). To measure the three components of the CoI framework, Arbaugh et al. (2008) created a 34-item survey. The survey had three factors that matched with the three components of CoI: items 1–13 measured teaching presence (Example scale item: "The instructor clearly communicated important course topics"); items 14–22 measured social presence (Example scale item: "Getting to know other course participants gave me a sense of belonging in the course"); and items 23–34 measured cognitive presence (Example scale item: "Problems posed increased my interest in course issues"). Like the previous scales in this study, we also adapted these items to measure preservice teachers' perceptions of these components for their future distance education teaching (Example scale item: "In distance education, as a teacher, I need to clearly communicate important course topics to the students"). The reliability of the three scales calculated from the data in this study was moderate to high: α = .91, α = .79, and α = .92, respectively.

Procedures and Data Analysis

We created an online version of the survey and distributed it within the online courses taught by the coresearcher in this study. The survey remained available to the students for 10 days. There were no incentives for participation.

All statistical analyses were conducted using JASP software (JASP Team, Amsterdam, the Netherlands, 2020). JASP is an open-source free alternative to commercial statistical software that allows for robust statistical analyses and has a user-friendly interface (Love et al., 2019).

To answer the first and the second research questions, we obtained descriptive statistics. To answer the third and fourth research questions, we ran multiple regression analyses treating each CoI component as the dependent variable and the remaining CoI components and motivation variables as independent variables.

Results

Self-Efficacy, Utility Value, and CoI Components

First, we descriptively investigated preservice teachers' self-efficacy and utility value perceptions to get a sense of their preparedness for distance education. The means indicated that preservice teachers, in general, agreed with the utility value and self-efficacy statements, indicating that they had strong beliefs that they could teach online and that they believed teaching online would be valuable for their future careers.

Next, we conducted similar descriptive analyses to investigate the perceptions of the importance of the three components of the CoI framework. It is notable that for all three components, the preservice teachers tended to strongly agree with the statements regarding the components' importance in their future online teaching experiences, with self-efficacy being the lowest-rated construct (Table 2).

Table 2

Descriptive Statistics for Self-Efficacy and Utility Value

	Utility value	Self-efficacy	Teaching presence	Social presence	Cognitive presence
M	4.149	3.799	4.706	4.409	4.649
SD	0.763	0.714	0.366	0.485	0.443

Relationship Between Motivation Beliefs and CoI Components

Before conducting the correlation analyses, we checked the normality for the distribution of the variables. The Shapiro–Wilk test indicated that the variables violated the normality assumptions. Therefore, we used Spearman's rho correlation coefficient. Preliminary analyses indicated that the variables had significant positive correlations with one another, notably and expectedly among the three components of the CoI framework (Table 3).

Table 3

Correlations for Community of Inquiry Components and the Motivation Variables (Spearman's Rho)

Variable	UV	SE	TP	SP	CP
UV	-				
SE	0.393	-			

TP	0.441	0.320	-		
SP	0.398	0.302	0.647	-	
CP	0.384	0.286	0.777	0.662	-

Note. Correlations are significant at $p < .001$ level. UV = utility value, SE = self-efficacy, TP = teaching presence, SP = social presence, CP = cognitive presence.

Next, to understand the relationship among the preservice teachers' motivation beliefs (i.e., self-efficacy and utility value) and the three components of the CoI framework (teaching, social, and cognitive presence), we ran multiple hierarchical regressions using each component of the CoI framework as the dependent variable in each regression and using the remaining variables as independent variables.

To find out the predictors for teaching presence, first, we ran a stepwise regression with teaching presence as the dependent variable and utility value, self-efficacy, social presence, and cognitive presence as covariates. The regression model with self-efficacy, social presence, and cognitive presence explained the most variance: $R^2 = .66$. Utility value was not included in the model since it did not significantly increase the R^2. It should be noted, however, that self-efficacy was only slightly over the acceptance threshold of $p < .05$.

Next, we ran a stepwise regression with social presence as the dependent variable and utility value, self-efficacy, teaching presence, and cognitive presence as covariates. The model with utility value, teaching presence, and cognitive presence explained the most variance: $R^2 = .51$. Self-efficacy was not included in the model since it did not significantly increase the R^2.

Finally, we ran a stepwise regression with cognitive presence as the dependent variable and utility value, self-efficacy, teaching presence, and social presence as covariates. The model with teaching and social presence explained the most variance: $R^2 = .67$. Self-efficacy and utility value were not included in the model since they did not significantly increase the R^2. The results of the regression analyses can be found in Table 4. The quantile–quantile (Q-Q) plots of residual distribution for all regression analyses indicated normality.

Table 4

Regression Results Predicting CoI Components

Predictor	β	t	p
Teaching presence (DV)			
(Intercept)	-	11.427	< .001
CP	0.640	15.127	< .001
SP	0.200	4.611	< .001

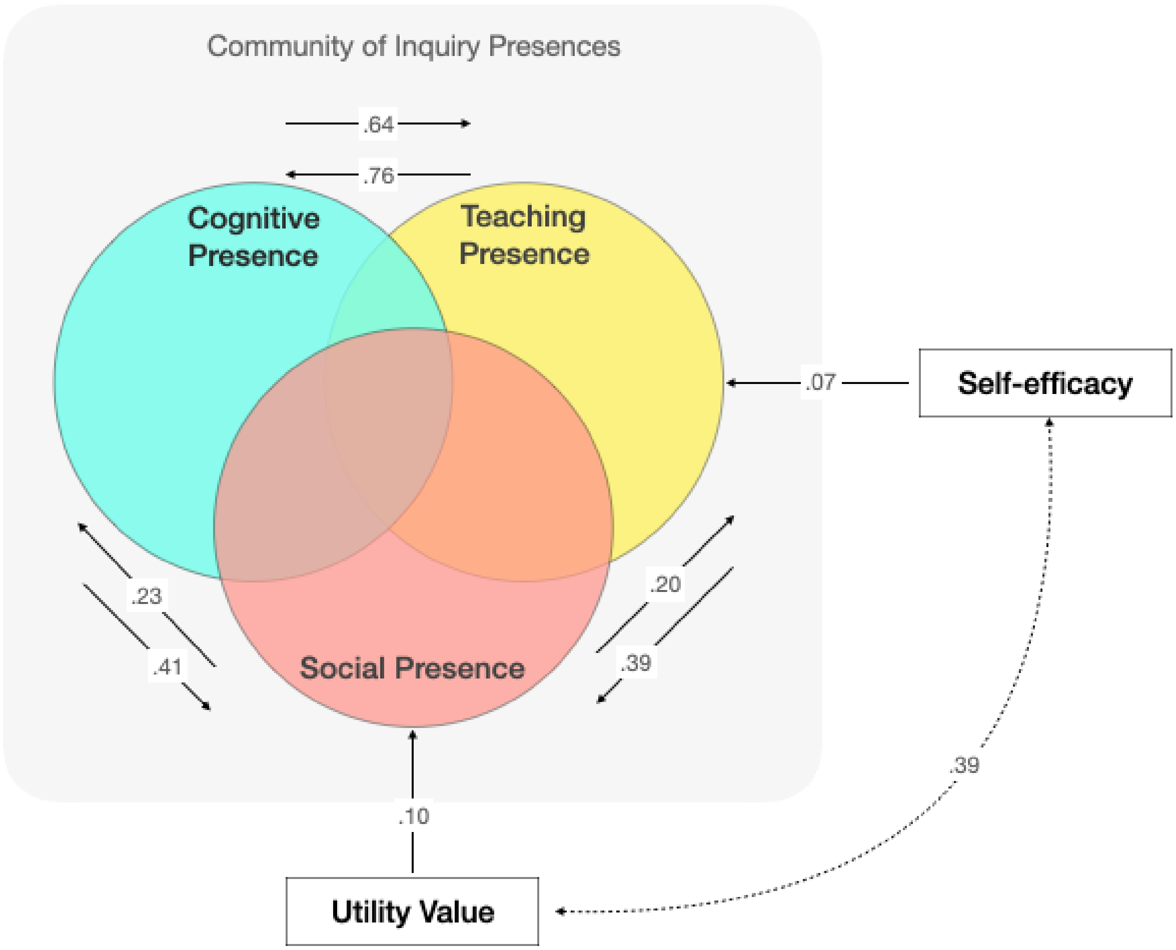
SE	0.068	1.975	.049
Social presence (DV)			
(Intercept)	-	0.962	.337
CP	0.412	5.951	< .001
TP	0.390	4.611	< .001
UV	0.104	3.969	< .001
Cognitive presence (DV)			
(Intercept)	-	0.177	.860
TP	0.763	15.364	< .001
SP	0.232	6.212	< .001

Note. DV=dependent variable; CP = cognitive presence; SP = social presence; SE = self-efficacy; TP = teaching presence; UV = utility value.

Based on the results of the regression and correlation analysis, Figure 1 shows a conceptual path model visualizing the network of relationships among the motivation and CoI variables.

Figure 1

Conceptual Path Model of the Self-Efficacy, Utility Value, and Community of Inquiry Components



Note. The dashed line represents correlation, while solid arrows represent regressions.

Discussion and Implications

The purpose of this research was twofold: first, we examined the preservice teachers' current levels of perceived self-efficacy and utility value toward distance teaching, as well as the importance they attribute, as future teachers, to the three dimensions of the CoI model. We also investigated the relationship between the motivation constructs and perceptions toward the importance attributed to the specific CoI components. Our results indicate that the participants had high perceptions of utility value and self-efficacy for distance education and the components of CoI. These results alone suggest that preservice teachers (a) feel ready for online teaching and (b) see the three distinct dimensions of the CoI as important aspects of teaching online.

More interesting, and the significant contribution of this study to the field, are the results of the investigation of the relationship between the motivation constructs and the CoI domains. Our regression analyses indicated that the CoI components predicted one another. This was an expected result, but it confirms the theoretical underpinnings of the model in that these components are interconnected (Garrison et al., 2010). More interestingly, however, in our analyses, we found that self-efficacy and utility value each predicted a different component of CoI. Specifically, while we found that self-efficacy predicted teaching presence, utility value predicted social presence.

The result from our regression analyses and the constructed conceptual path model are important and can be used to inform teacher education programs. These results can be interpreted in several ways. First, teaching presence refers to the overall design of the course, including selecting materials, organizing content, and facilitating the learning activities (Garrison et al., 1999). From this perspective, it seems that preservice teachers' self-efficacy to teach online is related to their perceptions of design, organization, and facilitation skills. Given that these tasks require knowledge of design, hardware, and software, these knowledge areas and practices seem to be directly linked with teachers' self-efficacy to teach online. Since self-efficacy can be supported through various experiences (e.g., mastery, vicarious) (Bandura, 1997), it then becomes essential for teacher education programs to offer such design and teaching experiences to preservice teachers to boost their self-efficacy to design, which then contributes to their perceptions of teaching presence (which then contributes to social and cognitive presence) (Garrison et al., 2010).

Second, our findings indicate that utility value, or one's perceptions of the relevance/usefulness of a task for one's future life or career, directly predicts preservice teachers' perceptions of the importance of social presence in CoI only. Social presence refers to one's perception of others in the learning environment as "real" people (Garrison et al., 1999). It involves building rapport and building personal connections among course participants. Extant research (e.g., Fryer & Ainley, 2019; Linnenbrink & Pintrich, 2003; Üner et al., 2020) indicates that utility value develops through interest development and is linked to self-efficacy. In other words, high levels of utility value are more likely to occur when one's interest and self-efficacy are also at higher levels (Bong, 2001; Linnenbrink & Pintrich, 2003; Wigfield & Eccles, 2000). Interpreting the findings from this perspective, we argue that for preservice teachers, social presence in online learning is considered the next step after the initial organization and delivery of course content. In other words, once preservice teachers reach a deeper level of utility value (through high self-efficacy and interest development), they begin to see social presence's importance. It should be noted, however, that these results are not definitive and should be interpreted with caution; we discuss this in more detail in the "Limitations and Future Research" section.

Based on our findings and suggested path model, we argue that teacher education programs should develop coursework and experiences to holistically support preservice teachers' perceived self-efficacy and utility value toward distance learning. Traditionally, the focus is on offering experiences that focus on the design, organization, and hardware/software aspects of online learning. Such traditional approaches may not be enough. Opportunities for these preservice teachers to understand the future relevance and connections of these experiences for their future teaching are also essential. Fortunately, extant research investigating interest and utility value development has identified that these motivational constructs can be targeted through simple classroom work (e.g., Akcaoglu et al., 2018; Kale & Akcaoglu., 2018; Hulleman et al., 2017;

Priniski et al., 2019). One such method is to give students chances to reflect on the utility value of classroom activities for their future careers and lives. For example, in their work, Hulleman et al. (2017) found that writing reflections helped students develop interest in and perceived value toward their coursework. Similarly, it can be argued that preservice teachers, during their undergraduate education, can be given opportunities not only to develop knowledge and skills related to the design and development of online learning but also to reflect on the usefulness of these experiences for their future teaching and for their students. Such a holistic approach would target teachers' self-efficacy and utility value perceptions, which in turn impact their perceptions of teaching, social, and cognitive presence, which are the key components of effective online learning.

Limitations and Future Research

There are several limitations to this study that might limit its generalizability. First, the data were collected from a public university in Turkey. Although university students all around the world had a chance to experience online education due to the COVID-19 pandemic, and therefore their perceptions of online learning can bear similarities to students in other contexts, it is possible that these students have characteristics that make them meaningfully different from other students in different contexts. Therefore, we believe that this research should be replicated in other settings to validate the generalizability of its findings.

Inherent limitations also exist in studies using cross-sectional surveys: self-reporting can introduce bias, and cross-sectional surveys present a one-time snapshot of a situation. Therefore, studies that incorporate other research methods, especially qualitative methods, should be conducted to further explore the relationships identified in our results. Other research designs that involve longitudinal data collection can also help us understand the developmental process and the relationships between preservice teachers' motivation and their perceptions of CoI components.

It should be noted that, although validated, the CoI survey used in this study (Arbaugh et al., 2008) may need revisions. Notably, the number of items for each sub-construct could be more balanced. There are also, as identified by previous CoI work (Anderson et al., 2001; Garrison et al., 2001), subcomponents of each CoI presence, and a more nuanced survey that considers these distinctions would provide a clearer picture of the participants' perceptions toward the CoI framework and its components and subcomponents. Research to develop and validate such a survey would be beneficial to both researchers and educators interested in designing, developing, and evaluating online learning environments.

Conclusions

The CoI framework has been widely used by researchers and educators in studying the design and implementation of online learning (Garrison & Cleveland-Innes, 2005; Kazanidis et al., 2018; Lee et al., 2020; Nagel & Kotzé, 2010; Popescu & Badea, 2020; Richardson et al., 2016; Tan et al., 2020). Results of previous research have consistently shown that the CoI presences are an effective framework to show factors affecting students' satisfaction with online learning. In this study, we shed light on an area that has not been studied before: the connection between preservice teachers' motivation and CoI components. We believe this important contribution can provide evidence for holistic approaches to undergraduate teacher education and provide clues about the need for experiences and activities that target not only self-efficacy but also interest and value development. We believe that through such a holistic approach, a future generation of teachers with an understanding of effective online learning can be guided.

The continuing uncertainty regarding the pandemic shows that the process of distance education may be implemented for a long time and can be a viable alternative to in-person instruction when schools are closed. School closures have given educational institutions a chance to see successful and unsuccessful examples of online teaching and the opportunity to consider online learning environments as alternatives to traditional learning environments in the long run. We believe teacher education institutions can ride the tailwinds of this momentum and introduce holistic courses that focus on skill and value development.

References

Akcaoglu, M., Rosenberg, J. M., Ranellucci, J., Schwarz, C. V. (2018). Outcomes from a self generated utility value intervention on fifth and sixth-grade students' value and interest in science. *International Journal of Educational Research. 87,* 67-77. https://doi.org/10.1016/j.ijer.2017.12.001

Akyol, Z., & Garrison, D. R. (2011). Assessing metacognition in an online community of inquiry. *The Internet and higher education*, *14*(3), 183–190. https://doi.org/10.1016/j.iheduc.2011.01.005

Anderson, S. E., Groulx, J. G., & Maninger, R. M. (2011). Relationships among preservice teachers' technology-related abilities, beliefs, and intentions to use technology in their future classrooms. *Journal of Educational Computing Research*, *45*(3), 321–338. https://doi.org/10.2190/EC.45.3.d

Anderson, T., Liam, R., 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

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall. https://psycnet.apa.org/record/1985-98423-000

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman. https://psycnet.apa.org/record/1997-08589-000

Bong, M. (2001). Role of self-efficacy and task-value in predicting college students' course performance and future enrollment intentions. *Contemporary Educational Psychology*, *26*(4), 553–570. https://doi.org/10.1006/ceps.2000.1048

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

Burgess, M. L., Slate, J. R., Rojas-LeBouef, A., & LaPrairie, K. (2010). Teaching and learning in *Second Life*: Using the community of inquiry (CoI) model to support online instruction with graduate students in instructional technology. *The Internet and Higher Education*, *13*(1–2), 84–88. https://doi.org/10.1016/j.iheduc.2009.12.003

Busch, T. (1996). Gender, group composition, cooperation, and self-efficacy in computer studies. *Journal of Educational Computing Research*, *15*(2), 125–135. https://doi.org/10.2190/KQJL-RTW1-VVUY-BHLG

Dewey, J. (1959). My pedagogic creed. In L. A. Cremin (Ed.), *Dewey on education: Selections*. Teachers College Press. https://www.tcpress.com/dewey-on-education-9780807776353

Dewey, J. (1986). Experience and education. *The Educational Forum*, *50*(3), 241–252. https://doi.org/10.1080/00131728609335764

Eccles, J. (1983). Expectancies, values and academic behaviors. Achievement and achievement motives. In J. T. Spence (Ed.), *Achievement and Achievement Motives: Psychological and Sociological Approaches* (pp. 75-136). W.H. Freeman and Company

Ertmer, P. A., Ottenbreit-Leftwich, A. T., Sadik, O., Sendurur, E., & Sendurur, P. (2012). Teacher beliefs and technology integration practices: A critical relationship. *Computers & Education*, *59*(2), 423–435. https://doi.org/10.1016/j.compedu.2012.02.001

Ferede, T., Melese, E., Tefera, E., & Mossie, A. (2016). Students' academic self-efficacy viz-a-viz their academic achievement: Jimma and Hawasa University students in focus. *Ethiopian Journal of Education and Sciences*, *11*(2), 1–16. https://www.ajol.info/index.php/ejesc/article/view/153661

Fryer, L. K., & Ainley, M. (2019). Supporting interest in a study domain: A longitudinal test of the interplay between interest, utility-value, and competence beliefs. *Learning and Instruction*, *60*, 252–262. https://doi.org/10.1016/j.learninstruc.2017.11.002

Garrison, D. R. (2009). Communities of inquiry in online learning. In P. L. Rogers, G. A. Berg, J. V. Boettcher, C. Howard, L. Justice, & K. D. Schenk (Eds.), *Encyclopedia of distance learning* (pp. 352–355). IGI Global.

Garrison, D. R., & Akyol, Z. (2013). The community of inquiry theoretical framework. In M. G. Moore (Ed.), *Handbook of Distance Education* (pp. 104–120). Routledge. https://www.routledgehandbooks.com/doi/10.4324/9780203803738.ch7

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

Garrison, D. R., 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., & Archer, W. (2000). *A transactional perspective on teaching and learning: A framework for adult and higher education*. Pergamon. https://eric.ed.gov/?id=ED451371

Garrison, D. R., & Cleveland-Innes, M. (2005). Facilitating cognitive presence in online learning: Interaction is not enough. *American Journal of Distance Education*, *19*(3), 133–148. https://doi.org/10.1207/s15389286ajde1903_2

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

Hodges, C. B. (2008). Self-efficacy in the context of online learning environments: A review of the literature and directions for research. *Performance Improvement Quarterly*, *20*(3–4), 7–25. https://doi.org/10.1002/piq.20001

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, *326*(5958), 1410–1412. https://doi.org/10.1126/science.1177067

Hulleman, C. S., Kosovich, J. J., Barron, K. E., & Daniel, D. B. (2017). Making connections: Replicating and extending the utility value intervention in the classroom. *Journal of Educational Psychology*, *109*(3), 387–404. https://doi.org/10.1037/edu0000146

Joo, Y. J., Lim, K. Y., & Kim, E. K. (2011). Online university students' satisfaction and persistence: Examining perceived level of presence, usefulness and ease of use as predictors in a structural model. *Computers & Education*, *57*(2), 1654–1664. https://doi.org/10.1016/j.compedu.2011.02.008

Joo, Y. J., Park, S., & Lim, E. (2018). Factors influencing preservice teachers' intention to use technology: TPACK, teacher self-efficacy, and technology acceptance model. *Journal of Educational Technology & Society*, *21*(3), 48–59. https://www.jstor.org/stable/26458506

Kale U., & Akcaoglu, M. (2018). The role of relevance in future teachers' utility value and interest toward technology. *Educational Technology Research & Development*. *66*(2), 283-311. https://doi.org/10.1007/s11423-017-9547-9.

Kazanidis, I., Pellas, N., Fotaris, P., & Tsinakos, A. (2018). Facebook and Moodle integration into instructional media design courses: A comparative analysis of students' learning experiences using the community of inquiry (CoI) model. *International Journal of Human–Computer Interaction*, *34*(10), 932–942. https://doi.org/10.1080/10447318.2018.1471574

Kim, G. C., & Gurvitch, R. (2020). Online education research adopting the community of inquiry framework: A systematic review. *Quest*, *72*(4), 395–409. https://doi.org/10.1080/00336297.2020.1761843

Kwon, K., Ottenbreit-Leftwich, A. T., Sari, A. R., Khlaif, Z., Zhu, M., Nadir, H., & Gok, F. (2019). Teachers' self-efficacy matters: Exploring the integration of mobile computing device in middle schools. *TechTrends*, *63*(6), 682–692. https://doi.org/10.1007/s11528-019-00402-5

Lee, R., Hoe Looi, K., Faulkner, M., & Neale, L. (2020). The moderating influence of environment factors in an extended community of inquiry model of e-learning. *Asia Pacific Journal of Education*, *41*(1), 1–15. https://doi.org/10.1080/02188791.2020.1758032

Linnenbrink, E. A., & Pintrich, P. R. (2003). The role of self-efficacy beliefs instudent engagement and learning in the classroom. *Reading & Writing Quarterly*, *19*(2), 119–137. https://doi.org/10.1080/10573560308223

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., Ly, A., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Wild, A., Knight, P., Rouder, J. N., Morey, R. D., & Wagenmakers, E.-J. (2019). JASP—Graphical statistical software for common statistical designs. *Journal of Statistical Software*, *88*(2), 1–17. https://doi.org/10.18637/jss.v088.i02

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., & Urdan, T. (2000). *Manual for the patterns of adaptive learning scales*. University of Michigan. http://websites.umich.edu/~pals/PALS%202000_V13Word97.pdf

Mishra, L., Gupta, T., & Shree, A. (2020). Online teaching-learning in higher education during lockdown period of COVID-19 pandemic. *International Journal of Educational Research Open*, *1*, Article 100012. https://doi.org/10.1016/j.ijedro.2020.100012

Nagel, L., & Kotzé, T. G. (2010). Supersizing e-learning: What a CoI survey reveals about teaching presence in a large online class. *The Internet and Higher Education*, *13*(1–2), 45–51. https://doi.org/10.1016/j.iheduc.2009.12.001

Pardo, A. M. S., & Peñalvo, F. J. G. (2008). Philosophical and epistemological basis for building a quality online training methodology. In F. J. G. Peñalvo (Ed.), *Advances in e-learning: Experiences and methodologies* (pp. 46–60). IGI Global. https://doi.org/10.4018/978-1-59904-756-0.ch003

Park, S. Y. (2009). An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. *Journal of Educational Technology & Society*, *12*(3), 150–162. https://www.jstor.org/stable/jeductechsoci.12.3.150

Pecka, S. L. (2014). *Evaluating higher order thinking in online discussions*. University of Nebraska Medical Center. https://mydigitalpublication.com/publication/?i=236271&article_id=1872237&view=articleBrowser&ver=html5

Ponton, M., Derrick, G., Hall, J. M., Rhea, N., & Carr, P. (2005). The relationship between self-efficacy and autonomous learning: The development of new instrumentation. *International Journal of Self-Directed Learning*, *2*(1), 50–61.

https://www.researchgate.net/publication/281323039_The_relationship_between_self-efficacy_and_autonomous_learning_The_development_of_new_instrumentation

Popescu, E., & Badea, G. (2020). Exploring a community of inquiry supported by a social media-based learning environment. *Educational Technology & Society*, *23*(2), 61–76. https://doi.org/10.2307/26921134

Priniski, S. J., Rosenzweig, E. Q., Canning, E. A., Hecht, C. A., Tibbetts, Y., Hyde, J. S., & Harackiewicz, J. M. (2019). The benefits of combining value for the self and others in utility-value interventions. *Journal of Educational Psychology*, *111*(8), 1478–1497. https://doi.org/10.1037/edu0000343

Putarek, V., & Pavlin-Bernardić, N. (2020). The role of self-efficacy for self-regulated learning, achievement goals, and engagement in academic cheating. *European Journal of Psychology of Education*, *35*(3), 647–671. https://doi.org/10.1007/s10212-019-00443-7

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

Richardson, J. C., Besser, E., Koehler, A., Lim, J., & Strait, M. (2016). Instructors' perceptions of instructor presence in online learning environments. *The International Review of Research in Open and Distributed Learning*, *17*(4), 82–104. https://doi.org/10.19173/irrodl.v17i4.2330

Rubin, B., Fernandes, R., & Avgerinou, M. D. (2013). The effects of technology on the community of inquiry and satisfaction with online courses. *The Internet and Higher Education*, *17*, 48–57. https://doi.org/10.1016/j.iheduc.2012.09.006

Song, L., & Hill, J. R. (2007). A conceptual model for understanding self-directed learning in online environments. *Journal of Interactive Online Learning*, *6*(1), 27–42. http://www.ncolr.org/jiol/issues/pdf/6.1.3.pdf

Taipjutorus, W., Hansen, S., & Brown, M. (2012). Investigating a relationship between learner control and self-efficacy in an online learning environment. *Journal of Open, Flexible, and Distance Learning*, *16*(1), 56–69. https://files.eric.ed.gov/fulltext/EJ1079899.pdf

Tan, H. R., Chng, W. H., Chonardo, C., Ng, M. T. T., & Fung, F. M. (2020). How chemists achieve active learning online during the COVID-19 pandemic: Using the community of inquiry (CoI) framework to support remote teaching. *Journal of Chemical Education*, *97*(9), 2512–2518. https://doi.org/10.1021/acs.jchemed.0c00541

Teo, T., & Zhou, M. (2014). Explaining the intention to use technology among university students: A structural equation modeling approach. *Journal of Computing in Higher Education*, *26*(2), 124–142. https://doi.org/10.1007/s12528-014-9080-3

Üner, A., Mouratidis, A., & Kalender, İ. (2020). Study efforts, learning strategies and test anxiety when striving for language competence: The role of utility value, self-efficacy, and reasons for learning English. *Educational Psychology*, *40*(6), 781–799. https://doi.org/10.1080/01443410.2020.1736989

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, *39*(2), 111–133. https://doi.org/10.1207/s15326985ep3902_3

Valtonen, T., Kukkonen, J., Kontkanen, S., Sormunen, K., Dillon, P., & Sointu, E. (2015). The impact of authentic learning experiences with ICT on pre-service teachers' intentions to use ICT for teaching and learning. *Computers & Education*, *81*, 49–58. https://doi.org/10.1016/j.compedu.2014.09.008

Valverde-Berrocoso, J., Garrido-Arroyo, M. D. C., Burgos-Videla, C., & Morales-Cevallos, M. B. (2020). Trends in educational research about e-learning: A systematic literature review (2009–2018). *Sustainability*, *12*(12), Article 5153. https://doi.org/10.3390/su12125153

Van Niekerk, M. P. (2015). Students' perceptions on IWB through the lens of the community of inquiry framework. *South African Journal of Education*, *35*(4), 1–10. https://doi.org/10.15700/saje.v35n4a1212

Vogel, F. R., & Human-Vogel, S. (2016). Academic commitment and self-efficacy as predictors of academic achievement in additional materials science. *Higher Education Research & Development*, *35*(6), 1298–1310. https://doi.org/10.1080/07294360.2016.1144574

Wang, L., & Finch, H. (2018). Motivation variables mediate the relationship between socioeconomic status and academic achievement. *Psychology and Education: An Interdisciplinary Journal*, *55*, 123–136. https://www.researchgate.net/publication/329516367_Motivation_Variables_Mediate_the_Relationship_between_Socioeconomic_Status_and_Academic_Achievement

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review*, *6*(1), 49–78. https://doi.org/10.1007/BF02209024

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wigfield, A., Tonks, S., & Eccles, J. S. (2004). Expectancy value theory in cross-cultural perspective. In D. M. McInerney & S. Van Etten (Eds.), *Big Theories Revisited* (pp. 165–198). Information Age Publishing.

Yavuzalp, N., & Bahcivan, E. (2020). The online learning self-efficacy scale: Its adaptation into Turkish and interpretation according to various variables. *Turkish Online Journal of Distance Education*, *21*(1), 31–44. https://files.eric.ed.gov/fulltext/EJ1238987.pdf

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

The Effects and Implications of Using Open Educational Resources in Secondary Schools

Paul Harvey, EdD and John Bond, EdD
Seattle Pacific University

Abstract

Open educational resources (OER) constitute a curriculum innovation that is considered revolutionary and has the potential to change the landscape of curriculum at all levels and content areas. OER have gained attention and widespread acceptance by educators and policy makers since 2002. The promise of OER is that they provide cost savings, promote collaboration, and are adaptable to the needs of teachers and students while providing a legitimate alternative to commercially produced print textbooks. Determining the relevance and viability of the movement to embrace OER requires an examination of theoretical foundations and empirical research to illuminate the effect of using OER as core curricula. While advocates promote the use of OER as a financially liberating model of curriculum and as a source of constructivist learning materials, more research is needed. The purpose of this study was to examine the relationship between OER and student learning. The study critically analyzed previous studies on OER and applied empirical analyses to the use of OER by a sample of middle schools. Twenty-eight middle schools from Washington State served as the subjects for the study. The study followed an ex post facto causal comparative model. Three research questions provided the focus for the study to investigate the effects of OER curriculum, duration of curriculum use, and other factors on student achievement in middle school mathematics. The results of the study found non-significant effects for OER use in relationship to school performance in mathematics, and significant effects on math scores for the variables of student poverty, curriculum duration, and cohort size.

Keywords: open educational resources, secondary schools, mathematics, state testing

Introduction

Considered to be an educational innovation with the potential to fundamentally change the nature of curricula from kindergarten through graduate school, open educational resources (OER) have gained widespread attention and acceptance by educators and policy makers over the past two decades (McKerlich et al., 2013; Smith & Casserly, 2006; United Nations Educational, Scientific and Cultural Organization [UNESCO], 2016). As educators, students, and policy makers become familiar with and adopt OER as an acceptable form of curricula, the need to assess the effects of OER on student learning has become increasingly important (Fisher et., 2015). Districts and teachers contemplating the use of OER as core curricula have reason to be concerned about the risks associated with abandoning familiar, mass-produced, and market-tested learning materials for resources that have open access and allow for liberal use, re-use, and repurposing. While advocates promote the use of OER as a financially liberating model of curriculum and as a source of constructivist learning materials, more research is needed.

This study was designed to examine the effects of using OER curriculum in secondary schools. It also examined the effect of other variables on student performance including cohort size, socioeconomic status, and duration of curriculum implementation. The study critically analyzed previous studies on OER implementation in order to provide insight and recommendations for future use of OER.

Development of OER

In 2002, UNESCO held a global forum that introduced the world to the concept of open educational resources. UNESCO developed an online community intended to provide a platform where educators and learners could access, copy, and change learning material without restrictions of copyright laws and economic resources. UNESCO stated that the OER movement had the potential to improve the quality of education and policy dialogue, as well as make it easier to share knowledge and build educational capacity. A product of the UNESCO summit held in 2002 was a working description of OER, which is now the generally accepted definition used by educational organizations. Open educational resources "are teaching and learning materials that reside in the public domain or have been released under an open license. These resources may be used free of charge, distributed without restriction, and modified without permission" (Office of Superintendent of Public Instruction [OSPI], 2015). Weiland (2015) provided additional qualifications for OER which included independent learning objects or content, and tools such as open software, collections, and licensing.

Current trends show increasing interest in and implementation of OER across the United States and across the world (McKerlich et al., 2013; Smith & Casserly, 2006; UNESCO, 2016). However popular the movement may be, there is a need to determine the value of OER to education as a whole. Determining the relevance and viability of the movement requires connecting policies and practices to theoretical foundations and empirical research in order to show the effectiveness of using OER as a main source curriculum.

In the area of learning reading, for example, an Internet search for OER lessons for reading elementary school yielded over 17 million articles, books, or reports. Narrowing the search to fourth grade reduced that bulk to approximately 1.3 million items. It is unrealistic to expect individual teachers to research and assess each item for quality. Even at a pace of reviewing OER daily, teachers would likely find the materials they deemed most useful modified or replaced by more relevant and useful OER. Such an exercise would be a time-consuming and frustrating endeavor. If OER are to be viable and credible as tools for teaching, then there is a need for demonstrations of positive impacts on learning and also an identification of barriers to their use.

Theoretical Foundations

The use of OER within curriculum and instruction can be connected to constructivist and progressivist ideals regarding learning. Dewey (1938) promoted experience, adaptation, and expansion of knowledge as key elements to authentic and meaningful learning. The perpetual re-creation and expansion of knowledge and meaning is at the core of what the OER movement has to offer. As Dewey described, the progressivist tradition emphasizes free activity that includes thinking and questioning and interacting, as well as acceptance and interaction with a changing world. Piaget (1971) promoted the ideal that prior knowledge is key to the development of new knowledge and understanding. OER constitute a curricular form of this process: prior knowledge is adapted and repurposed to form new knowledge in the form of new learning objects. Adaptation and repurposing can be done at the individual teacher or student level, or on a level involving whole school systems. Knox (2013) affirmed that the OER are "the building blocks of a constructivist-informed 'learning 2.0,' comprised of social learning, legitimate peripheral participation and learning through communities of practice" (p. 825).

The OER movement shows strong ties to Dewey's philosophy of experiential learning, as it meets his description of ideal curricula. "Scientific study leads to and enlarges experience, but this experience is educative only to the degree that it rests upon a continuity of significant knowledge and to the degree that knowledge modifies or 'modulates' the learner's outlook and attitude" (Dewey, 1938, p. xii). Furthermore, OER provide the continuity Dewey discussed, as they are borne by educational choices and they produce artifacts for future learning. "From this point of view, the principle of continuity of experience means that every experience both takes up something form those which have gone before and modifies in some way the quality of those which come after" (Dewey, 1938, p. 36).

Prior Research

Robinson et al. (2014) conducted a quantitative quasi-experimental study which investigated the impact on learning outcomes for secondary science students who did and did not use OER for learning. Their study sampled over 4,000 students and 43 teachers in science courses from a school district in Utah over a two-year period. Using the Utah State science test as an outcome variable, the analysis controlled for 10 covariates and compared students who learned from OER and those who learned from a traditional curriculum of printed textbooks. Robinson et al. also controlled for teacher effect. There was a significant difference between the treatment and control groups, and several of the covariates did affect Criterion Referenced Test [CRT] scores significantly. Data presented in their results showed that the predictor variable of teacher effect had a greater effect size and *t* value than did the use of OER. The greatest predictor for scores on the CRT were students' previous scores. As for the effect of the treatment, the use of OER texts resulted in better scores on the CRT for students in chemistry. However, this was not the case for students in earth systems and biology. Though there was not sufficient effect size to promote wholesale adoption of OER in science courses, the data supported the notion that OER use had, at worst, a neutral impact on student achievement. Therefore, it would be appropriate to promote OER adoption for the sake of cost savings. “These findings conformed with our belief that teacher efficacy and prior ability would play a much more important role in educational achievement than textbook selection” (Robinson et al., p. 346).

Wiley and Hilton (2012) examined both the hypotheses of cost savings and impact on student learning promised by the use of OER. Their study was conducted over two years and included 20 middle school and high school science teachers. Data supported assumptions of cost savings through the use of OER but did not validate assumptions of learning improvement. While much OER are accessed and used electronically, the study prompted further inquiry into the effect of OER format (e.g., paper vs. electronic) and its effect on student learning. Conclusions drawn from the data showed no support for the hypothesis that OER improved student learning. “Simply substituting open textbooks for traditional textbooks did not appear to have an effect on student test scores” (Wiley & Hilton, p. 212).

Effect on Teaching Practice

In addition to examining the effect of changing the licensing format of the curriculum, it is important to consider any change in pedagogy employed by teachers in the context of using OER. Dotson and Foley (2017) emphasized that the variable of curriculum, in itself, does not account for student learning outcomes. Rather, OER elevated the effect that teachers and their pedagogies had on student learning as important factors to be considered. Pearcy (2014) reaffirmed the important role of the teacher for providing feedback and guidance in the context of innovative curriculum and instruction.

Demographics as a Variable in Student Learning

This study tested the hypothesis that other factors, including socioeconomic status, affect student achievement within the context of curriculum licensing format. Dotson and Foley (2017) stated that student achievement has a negative correlation with the students’ degree of poverty—the poorer a student, the less likely they will score at standard in standardized tests. They emphasized that this statistical relationship is consistent with the available research on student poverty and is pervasive across regions in the US. “Schools with high levels of poverty score very low on current measures of effectiveness which are primarily based on standardized tests” (Dotson & Foley, 2017, p. 299).

Barriers to Use

Kelly (2014) referenced the social learning theory regarding user efficacy and the context of trying something new, specifically OER. Kelly referenced the theory to address perceptions, intentions, and actual technology use. Kelly’s research indicated that attitudes and perceptions about OER were strongly correlated to the attitudes regarding new technology and moderately correlated to actual use of OER. “This indicates that OER must be considered easy to use or the perceived utility of the resource will be negatively impacted” (Kelly, p. 37). A potential barrier to the use of OER by instructors could be the perception of difficulty in finding and using OER, or actual negative experiences by teachers and students who have used OER without success or with great difficulty.

Time and access to training are also factors related to acceptance of OER, as teachers who wish to explore and develop their own skills around the use of instructional innovations find little opportunity within their contracted work to do so. The fast rate at which change occurs, particularly in regard to technology, confounds these obstacles.

Method

This study was guided by the following three research questions:

1. Is there a significant difference in school-level achievement scores in mathematics between eighth grade students who use OER compared to those who use commercially published print curriculum?

2. Does the length of time a given curriculum is implemented have a significant effect on school-level achievement scores in mathematics among eighth grade students?

3. Are there additional effects on school-level achievement scores in mathematics of eighth grade students besides the variables of cohort size and socioeconomic status? Is there a difference in the effect of those variables between students who use OER compared to those who use commercially published non-OER curriculum?

A causal comparative study was selected as the research design for this study. While Gall et al. (2007) cautioned that such studies do not permit strong cause-and-effect conclusions, they are appropriate in exploratory investigations in which manipulating the independent variable is a challenge. In this study, this researcher was not in a position to direct school districts as to the type of curriculum they adopted, therefore, a causal comparative design was appropriate.

One motivating factor for this study was to provide empirical evidence that would help future researchers decide whether the variables had a strong enough relationship to warrant the expense and time required to conduct experimental research. Our purpose was to help them focus on more specific variables on which to base controlled experimental studies. Future studies could compare the effects on student learning using independent variables such as curriculum type, specific source or title of OER, students' grade level, or other student or school demographics.

Variables

The main independent variable was the type of curriculum format—traditional, commercially produced print curriculum or OER curriculum. Other independent variables included (a) cohort size of each school's eighth grade class, (b) duration of curriculum adoption, and (c) socioeconomic status as indicated by each school's percentage of students receiving free and reduced-price lunches. The dependent variable was the school-level scores on the mandatory state assessments in mathematics.

Participants

The subjects of the study were 28 public middle schools in Washington State, representing 6,984 students. The cohorts consisted exclusively of eighth grade students who completed the Smarter Balanced Assessment (SBA) for mathematics in the spring of 2017. Each school and its related demographic data was considered to be one case in the statistical analysis. The schools were divided into two equal groups to compare the percentage of students proficient on the math assessment based on their use of OER. The original pool of all schools in the study consisted of 32 schools that used OER for mathematics curriculum and 14 schools that used non-OER curriculum. Because it is preferred to have equal group sizes for comparison in the administration of *t* tests, a computerized random number generator program was used to select schools into the OER comparison group which resulted in the formation of two comparison groups with 14 schools in each group.

The 28 schools included in the sample represented 11 of the 295 school districts in the State of Washington. Many of the schools were within a common geographic area only 600 square miles large, and 5 of the 11 school districts bordered at least one other sampled school district. One of the selected districts was located over 500 miles away from the nearest school district included in the sample. The original intent of the study was to include schools in districts that were contiguous within a given geographical region. However, the number of schools in Washington was limited, and so was a constraint on the available districts from which to draw data.

Data Collection

Demographic data was collected on each school in both the OER and non-OER groups, including (a) the number eighth grade students who were assigned to complete SBA math examination, (b) the number of years which the school used its particular eighth grade math curriculum, and (c) the percentage of students who participated in the National School Lunch Program.

Ex post facto data were drawn from a convenience sample of eighth grade mathematics scores from the 2017 SBA results, obtained from OSPI and used to compare means of school scores between schools using OER and schools using non-OER curricula; this served as the criterion variable representing student achievement. The SBA is an assessment used by several states including Washington as a requirement to show schools and school district students' proficiency in mathematics and English language arts. Passing the SBA in mathematics is also an official pathway to graduation from high school in Washington State. This assessment provided valid, reliable, and fair assessments of the deep disciplinary understanding and higher-order thinking skills increasingly demanded by a knowledge-based global economy (Smarter Balanced Assessment Consortium, 2016, p. vi).

There were seven different curricula used by all schools in the sample. The OER group used four different open math curricula, and the non-OER group used three different commercially published textbooks. Duration of curriculum use was a variable constructed to measure the number of years a given curriculum was in use prior to the 2017 SBA eighth grade mathematics assessment. The outcome variable of this study was the school-level scores on the Smarter Balanced Assessment (SBA).

An anonymous, informal survey of nine questions was provided to teachers in schools which used OER. Five of the questions asked for responses regarding perceptions of the use of OER from the vantage point of their role as teachers and their interaction with the curriculum, particularly regarding the effects of OER on student learning and their own teaching practices. The remaining four questions were more general about the respondent's experience as a teacher, such as years of service.

Data Analysis

The dependent variable in this study was the school-level score on the SBA mathematics examination for eighth grade. The independent variables included the (a) licensing format of curriculum, (b) size of the student cohort, (c) duration that a curriculum had been used prior to testing, and (c) percentage of students enrolled in the free and reduced-price lunch program.

This study aimed to compare the effect of using OER as curriculum on student achievement to the effect of using commercially published print curriculum. Grade-level mathematics scores were examined regarding the variable of OER condition and demographic variables to complete the data set. The data were screened for outliers, missing values, and normality. Descriptive and inferential statistics were calculated. The reported data included means, standard deviations, and statistics of skewness and kurtosis. The *t* test analysis (Gall et al., 2007) was used to compare means between groups to determine whether any differences of a common variable were statistically significant. An independent-samples *t* test compared means between groups for the main effect of OER use on math scores.

Since other factors affect learning outcomes beside the licensing format of the curriculum, a multiple regression analysis was performed to determine the relationship of variables to school test scores, beyond the use of OER. For all inferential statistics tests, a value of .05 was set as the threshold for significance.

Results

From the quantitative and qualitative comparisons, a few trends emerged and are worth notice. The non-OER schools outperformed the OER schools by nearly 5% on the SBA mathematics test. The mean duration of curriculum use prior to the 2017 test was greater for the non-OER group by three years. Finally, the correlation for student poverty and test performance had an effect size of 77% ($\beta = .77$).

Results of Quantitative Analysis

All variables were normally distributed for both comparison groups. Table 1 shows the mean percentage of students who were proficient on the math assessment, by school, for both curriculum groups. The schools in the OER curriculum group had a lower percentage of students who passed the SBA mathematics test ($M = 39.66$, $SD = 10.24$), than did the schools that utilized the non-OER curriculum had ($M = 42.82$, $SD = 12.69$), though the differences were not statistically significant.

For all schools in the total sample ($N = 28$), each variable had a range of values. The range of percentage of students showing proficiency on the SBA mathematics examination was 23.3% to 67.9%. The range for cohort size of each school's eighth graders who took the examination was 154 to 385 students. The number of years of use of the particular curriculum by each school ranged from 1 to 8 years. Finally, the range of students enrolled in the free and reduced-price lunch program was between 24.7% and 87.9%.

Table 1

Descriptive Statistics for Percentage of Students Proficient on Eighth Grade SBA (Math)

Descriptor	OER group	Non-OER group
Number	14	14
Mean	39.66	42.82
SD	10.24	12.69
Skewness		
Statistic	.128	.450
Standard Error	.597	.597
Kurtosis		
Statistic	-.096	-.307
Standard Error	1.154	1.154

Duration of curriculum was the second predictor variable studied. The non-OER schools had a greater mean number of years using their respective curricula (M = 6.14, SD = 1.74) than did the schools using OER (M = 2.50, SD = .65). In an effort to further understand the relationship between duration of curriculum and curriculum format, a t test was run to determine if there was a significant difference between the groups on the variable of duration of curriculum. The result of that analysis showed that the non-OER group had on average significantly more time using their curricula than did the OER schools $t(1,16) = 7.309$, $p < .001$. The OER schools had a larger mean cohort size as measured by the number of eighth grade students who took the SBA for mathematics in 2017 (M = 276.85, SD = 62.83) than did the non-OER schools (M = 222.00, SD = 39.87). As a demographic statistic for measuring the sample schools' level of poverty, the non-OER schools (M = 51.62, SD = 17.61) had a higher mean percentage of students enrolled in the free and reduced-price lunch program than did OER schools (M = 49.22, SD = 13.30).

The main effect for the first hypothesis regarding OER use by schools was found by conducting a t test which compared the mean scores of OER schools against the non-OER schools on the measure of school-level math scores on the SBA. The results of the t test showed that the non-OER schools (M = 42.8, SD = 12.69) had a higher percentage of students meeting proficiency on the SBA mathematics test than did the OER schools (M = 39.66, SD = 10.24), however the difference was not statistically significant $t(26) = .726$, p = .474.

A simple regression analysis was run for the effect of duration of all schools' curriculum implementation. The effect for curriculum duration by itself was found to be not significant, $R^2 = .050$, $F(1, 26) = 1.36$, p =

.254, indicating that the variable of how many years a curriculum is used by a school, when isolated from other variables, does not have a significant effect on student achievement.

A multiple regression analysis was conducted to examine the correlation between the variables of duration of time curriculum was used, student cohort size, and the percentage of students in the free and reduced-price lunch program as these related to the percentage of students who were proficient on the SBA mathematics test. The multiple regression analysis was run in two models. The first model isolated the predictor variable of curriculum format, namely status of OER use. Statistics within the first model were used to infer that use or non-use of OER did not significantly correlate to a school's math score, $F(1,26) = .527$, $p = .474$. In the second model, the predictor variables of free and reduced-price lunch, cohort size, and curriculum duration were added in to determine if there was a relationship between those variables and math scores. The second model included data which allowed for rejection of the null hypothesis, namely that variables other than OER use do not correlate to a school's mathematics scores. The results of the second model in the multiple regression analysis, $R^2 = .633$, $F(4,23) = 9.93$, $p < .001$, provided detail as to each predictor variable's relationship with the schools' math scores (Table 2).

In order of largest to smallest effect size, each predictor variable was found to have a relationship to the criterion variable. The free and reduced-price lunch program's percentage of a school's population had a significant negative correlation ($\beta = -.611$, $p < .001$); the correlation for curriculum duration was positive and significant ($\beta = .604$, $p = .014$); and the correlation for cohort size was positive and significant ($\beta = .344$, $p < .001$). These results supported the notion that the three variables in combination have a significant effect on a school's SBA math scores. The results also supported the claim that there are variables other than curriculum licensing status which affect the outcome of school-level scores in the math assessment, and that 63% of a school's math scores can be explained by the three variables included in multiple regression analysis.

Table 2

Multiple Regression Analysis Variable Statistics Related to Student Proficiency Rates on the Eighth Grade SBA (Math)

Variable	Statistic	Significance (p)
Cohort size	$\beta = .344$	$< .001$
Curriculum duration	$\beta = .604$	.014
Free and reduced-price lunch percentage	$\beta = -.611$	$< .001$

Qualitative Analysis Results

Results from the informal survey showed mostly favorable views of using OER. While a change in teacher practice was noted by more than half of the respondents, this was not the experience for all teachers using OER. Commentary from teachers who believed OER changed their teaching practices included: "I look more closely at how math is applied to everyday life"; "The curriculum is designed around group/partner talk"; "The basic format of most chapters is inquiry (i.e., a problem that is approachable for students to teach them the concept, as opposed to the traditional methods of lecture and teach)."

Some teacher commentary indicated that OER was not effective in improving student learning. For example: (a) the OER curricula over-emphasized examples and practice problems; (b) having just consumables rather than a textbook has been an adjustment to get used to, and that adjustment was cause for dissatisfaction with OER; (c) grading the workbooks was a challenge; (d) the definitions of terms were found through the material, not just in one central location, which was perceived to be a challenge for students; (e) some of the content did not cover the standards required for the grade level; and (f) the OER curriculum did not provide enough basic practice for some of the concepts.

In general, the teachers' responses to the use of OER indicated that the curriculum provided more opportunities for inquiry, deeper understanding of math concepts, and more student-to-student interaction. These positive attributes of OER are relevant and timely, particularly in the context of the growing movement to include those attributes in pedagogy improvement efforts. This is particularly so within the frameworks of universal design for learning, with its emphasis on choice and having students find relevancy in the curriculum, and the habits of mind as well as of mathematical practice. The positive traits of OER described by the teachers in the study align well to the elements currently being promoted in discussions around student-centered educational reform. Future study of the use of OER versus non-OER curricula in combination with traditional and emerging pedagogical frameworks may reveal more useful data about the combined variable effects on teacher confidence in the curriculum and student learning outcomes.

Discussion

The purpose of this study was to determine if the use of OER math curricula had a statistically significant effect on student achievement as measured by the annual required state assessment in mathematics. The analysis of the data supported the hypothesis that the licensing format of curriculum as a single variable does not have a statistically significant effect on a school's SBA mathematics scores. Schools that used OER curriculum performed slightly lower in eighth grade mathematics than schools that used traditional commercially published print curriculum, but not at a significant level. This may be interpreted as a positive sign for schools that are considering adopting OER math curricula, as the data in this study showed there was no significant difference between student achievement at OER and non-OER schools. However, because schools are officially and casually rated by their tests scores, any difference in scores may be interpreted by the public as an indication of one curriculum format being better or worse than another. Lacking any evidence of harm to learning, a switch from commercially published curricula could save school districts tens of thousands of dollars in curriculum purchases. Such savings could be redistributed to implement intervention programs, hire teachers and other support staff, and purchase learning materials, all of which can successfully address achievement and opportunity gaps. In other words, the cost savings that is implied with the adoption of OER can be used to accelerate student achievement.

Other factors affect learning outcomes beside open-access status of the curriculum. As Wiley and Hilton (2012) discovered, teacher effect and previous performance on examinations had larger positive effect sizes than the use of OER curriculum in relation to student learning. The finding that curriculum licensing format did not significantly affect student learning outcomes, combined with the finding that poverty had a significant and negative correlation to student achievement, prompts a call to action on the latter. Findings from this current study supported the notion that putting effort into finding effective measures that reduce or eliminate the negative effects of poverty will have greater value over experimenting with different curriculum formats.

The discussion of poverty and its effect on student achievement is germane to this study for two reasons. First, it causes the researcher to consider student socioeconomic status along with other factors when studying student achievement. And, second, it is a cause for inquiry regarding whether a particular curriculum or level of curriculum accessibility, that is, purchased or open-access, is more or less effective in helping students in poverty perform at standard on standardized tests.

Recommendations for Further Research

Experimental studies are lacking in the available research regarding OER. Two areas that could be better addressed by more experimental studies are the effect of curriculum type on student learning and also the student experience, both of which were outside the scope of this study. There would be widespread interest in a large-scale study comparing the different sources or titles of OER.

Conclusion

The results of this study indicate that the use of OER curriculum by itself does not significantly affect student learning in mathematics. The findings can be used to promote OER as a viable source of curriculum, as its implementation will not significantly benefit nor harm student achievement scores in math. Other variables are statistically significant contributors to student achievement in mathematics and include size of student cohort, duration of curriculum use, and rate of poverty. These factors should continue to be included in future studies.

Reasonable caution should be employed when making any shift in curriculum. Selecting an OER curriculum simply because it is more accessible through ubiquitous technology will not yield significant changes in student achievement. Even with the exponential growth in access to OER, its lifespan to date has been relatively short in the history of whole curricula. As Pearcy (2014) explained "teachers, like all individuals, are more likely to adopt an innovation if it proves to be a more effective means to accomplish something, and has observable effects" (p. 180). In the face of new and widespread exposure to OER by the hundreds of districts and dozens of states affected by the global COVID-19 pandemic and massive school closures, there is still a need for more evidence that OER improves teaching and learning before recommending a wholesale shift away from traditional forms of curricula to an exclusively OER structure. The current condition of forced remote or distance learning and teacher collaboration may be fertile ground for the proliferation of OER development, use, and study.

References

Dewey, J. (1938). *Experience and education*. Touchstone.

Dotson, L. & Foley, V. (2017). Common Core, socioeconomic status, and middle level student achievement: Implications for teacher preparation programs in higher education. *Journal Of Education and Learning, 6*(4), 294–303. http://doi.org/10.5539/jel.v6n4p294

Fischer, L., Hilton, J., III, Robinson, T., & Wiley, D. (2015). A multi-institutional study of the impact of open textbook adoption on the learning outcomes of post-secondary students. *Journal of Computing in Higher Education, 27*(3), 159–172. https://doi.org/10.1007/s12528-015-9101-x

Gall, M., Gall, J., & Borg, W. (2007). *Educational research: An introduction*. Pearson.

Kelly, H. (2014). A path analysis to educator perceptions of open educational resources using the technology acceptance model. *International Review of Research in Open and Distributed Learning, 15*(2), 26–42. http://dx.doi.org/10.19173/irrodl.v15i2.1715

Knox, J. (2013). Five critiques of the open educational resources movement. *Teaching in Higher Education, 18*(8), 821–832. https://doi.org/10.1080/13562517.2013.774354

McKerlich, R., Ives, C., & McGreal, R., (2013). Measuring use and creation of open educational resources in higher education. *The International Review of Research in Open and Distributed Learning, 14*(4), 90–103. https://doi.org/10.19173/irrodl.v14i4.1573

Office of the Superintendent of Public Instruction. (2015). *Open educational resources*. http://digitallearning.k12.wa.us/oer/

Office of the Superintendent of Public Instruction. (2018). *Washington State report card*. http://reportcard.ospi.k12.wa.us/summary.aspx?groupLevel=District&schoolId=1&reportLevel=State&yrs=2016-17&year=2016-17

Pearcy, M. (2014). Student, teacher, professor: Three perspectives on online education. *History Teacher, 47*(2), 169–185.

Piaget, J. (1971). *Psychology and epistemology: Towards a theory of knowledge*. Viking.

Robinson, T., Fischer, L., Wiley, D., & Hilton, J., III. (2014). The impact of open textbooks on secondary science learning outcomes. *Educational Researcher, 43(7),* 341–351. https://doi.org/10.3102/0013189X14550275

Smarter Balanced Assessment Consortium. (2016). *Smarter Balanced 2014–15 technical report*. https://portal.smarterbalanced.org/library/en/2014-15-technical-report.pdf

Smith, M., & Casserly, C. (2006). The promise of open educational resources. *Change, 38*(5), 8–17. https://doi.org/10.3200/CHNG.38.5.8-17

United Nations Educational, Scientific and Cultural Organization. (2016). *Open educational resources*. http://www.unesco.org/new/en/communication-and-information/access-to-knowledge/open-educational-resources/

Weiland, S. (2015) Open educational resources: American ideals, global questions. *Global Education Review 2*(3), 4–23.

Wiley, D., & Hilton, J., III. (2012). A preliminary examination of the cost savings and learning impacts of using open textbooks in middle and high school science classes. *The International Review of Research in Open and Distributed Learning, 13*(3). https://doi.org/10.19173/irrodl.v13i3.1153

Athabasca University

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

Design and Validation of the Virtual Classroom Management Questionnaire

Mohsen Keshavarz[1], Zohrehsadat Mirmoghtadaie,[2*] and Somayyeh Nayyeri[3]
[1a]Department of Medical Sciences Education Development, Torbat Heydariyeh University of Medical Sciences, Torbat Heydariyeh, Iran. [1b]Department of e-Learning, School of Paramedical Sciences, Torbat Heydariyeh University of Medical Sciences, Torbat Heydariyeh, Iran. [2]Department of e-Learning, Virtual School of Medical Education and Management, Shahid Beheshti University of Medical Sciences (SBMU), Tehran, Iran.(Responsible author). [3]Department of Operating Room, School of Nursing and Midwifery, Torbat Heydariyeh University of Medical Sciences, Torbat Heydariyeh, Iran.*

Abstract

Effective classroom management methods are well known, but effective ways of managing classes of beginner teachers remain elusive. Classroom management refers to the wide range of skills and techniques that teachers use to ensure that classes are conducted without destructive student behavior. The present study is applied nonexperimental research. The purpose of this study was to design a tool to measure the effective management of the virtual classroom from the perspective of professors and students in e-learning and evaluate its validity and reliability. The research sample was taken randomly from all universities that make use of e-learning in Tehran, Iran, during the 2019–2020 semesters. The results show that the professional development of online classroom management is necessary for preparing teachers to teach in digital environments. The results of this research in the form of a validated questionnaire can be considered as an indicator for educators and students working in online environments, and this tool can be used for effective teaching and learning in the digital age.

Keywords: classroom environment, classroom management, questionnaire, virtual classroom management

Design and Validation of the Virtual Classroom Management Questionnaire

Today, in the digital age, one of the main attributes that learners need to have is the skills to learn in new digital environments. For this reason, teachers must be familiar with digital-age teaching skills and techniques to manage and lead online classes (Keshavarz & Ghoneim, 2021). Among the various elements and components of the educational system, priority is given to the teacher, because it is through the teacher that educational goals in different dimensions are achieved. Various factors can be considered as effective in the promotion of quality teaching–learning processes in the higher education system. One of these factors is classroom management. The classroom environment is of special importance and sanctity. Classroom management skills are the cornerstone of overall teaching success. Despite one's scientific abilities, if a teacher cannot use the skills of classroom management, the realization of effective teaching will be difficult to achieve. Research shows that not only classroom management but also classroom management dynamics are important for effective teaching and learning (Levin & Nolan, 2014). The result of effective management is the creation of a democratic community in the classroom and thus effective learning (Konti, 2011; Muijs & Reynolds, 2017).

Classroom management is an important skill for the academic, social-emotional, and motivational development of students and the health of teachers (Gold et al., 2021). *Classroom management* refers to a wide variety of skills and techniques that teachers use to keep students organized, focused, and academically productive during a class (Babadjanova, 2020). Teachers must have classroom management skills the to successfully build a secure and effective learning environment for pupils' quality education (Adedigba & Sulaiman, 2020). As mentioned by Malik et al. (2020, p. 260), "management of classroom is a strong combination of the management of content and also the conduct of the teacher. Effective management means the management of students' attitudes, personalities, vitalities, competencies and passions." When classroom management strategies are executed effectively, teachers minimize behaviors that impede learning for both individual students and groups while maximizing behaviors that facilitate or enhance learning (Hapsari, 2020).

Many teachers enter the field of education without the necessary skills to implement an effective classroom management program and respond appropriately to student behavior (Greenberg et al., 2014). Various research has been done on this topic. For example, Akman (2020) collected data with a classroom management scale developed by Özcan and Gülözer (2017). This scale contains 18 items and three factors (human management, course management, and behavior management). The study was aimed at analyzing the correlations between teachers' classroom management efficacies, students' confidence in teachers, and the perception of educational stress (Akman, 2020). Additionally, Berger and Girardet (2020) show that the more vocational teachers felt responsible for the quality of their teaching, the more they tended to endorse adaptive or beneficial classroom management styles.

Classroom management includes managing students in the classroom with discipline and creating a conducive environment to facilitate learning and behavior change (Ghiasvandian et al., 2017). This is a cyclical process that includes the following steps: advanced design, implementation, evaluation during implementation, and final evaluation. The instructor must also consider learners and the environment as factors. Curzon (2003) considers a teacher's managerial responsibilities to include creating and maintaining a classroom environment in which effective learning takes place, compiling and explaining outlines, articulating goals, teaching with appropriate methods, motivating the class, evaluating learner performance, and providing feedback.

Effective classroom management also improves students' disciplinary behaviors (Kayıkçı, 2009). Classroom management is the component of teaching and learning, and it seems to be the most common concern of both preservice and experienced teachers (Yılmaz & Çavaş, 2008). Classroom management is the strategy that teachers use to maintain the environment in which teaching and learning be accomplished (Wong & Wong, 2001).

Online learning refers to educational delivered over the Internet. It includes a wide range of online applications to access educational materials, as well as to facilitate teacher–student interaction. *Online learning* is a term that is often equivalent to *distance learning* and *e-learning* (Bakia et al., 2012). Langdon (1997) indicates that 58 percent of teachers reported that students were constantly disruptive, and 50 percent of teachers expressed concern about learners' disobedience. In an analysis of 135 student–teacher experiences, Tulley and Chiu (1995) report that 15 percent of students regularly break the rules, and an additional 5 percent are chronic offenders. As a result, teachers are actively seeking information on effective classroom management practices (Hardman & Smith, 2003).

The purpose of this study was to design a tool to measure the effective management of the virtual classroom from the perspective of professors and students in e-learning and evaluate its validity and reliability.

Methodology

The present study is applied research in terms of its purpose, and it is nonexperimental research in terms of data collection: the design and psychometric analysis of the questionnaire were done in four stages based on Waltz et al.'s (2010) method. In the first step, to get acquainted with the concept of online classroom management, a review of relevant texts and resources to identify components of classroom management was performed. Then, the initial questions and questionnaire items were designed; the initial version of the questionnaire included 64 items. The present questionnaire consists of two parts. The first part consists of demographic information, and the second part includes the questionnaire items. A total of 64 items were designed; they are measured on a 5-point Likert from very high (5) to very low (1). In the next step, to evaluate and determine the validity of the questionnaire, content validity and construct validity methods were used. To assess the validity of the content, the questions were examined by 15 experts in the field of e-learning, medical education, and educational technology. These experts were asked to comment on whether the questions were necessary and useful. Finally, their information and opinions were collected and the necessary changes were made. After reviewing the opinions of experts, 39 options were approved in the form of a questionnaire, and five factors were determined: managing supportive interactions and behaviors, course management, meta-cognitive skills management, conflict management, and time management. If the number obtained from the Lawshe (1975) table (to determine the minimum value of the index) was greater than 0.49 (based on the evaluation of 15 experts), this indicated that the presence of the items in this tool was necessary and important (Figure 1).

Figure 1

The Content Validity Ratio (CVR) Formula

$$CVR = \frac{n - \left(\frac{N}{2}\right)}{\frac{N}{2}}$$

Note. The CVR proposed by Lawshe (1975) is a linear transformation of a proportional level of agreement on how many experts within a panel rate an item as essential. n = total number of experts divided by the number of experts saying item essential; N = total number of experts on the panel. From “A Quantitative Approach to Content Validity,” by C. H. Lawshe, 1975, *Personnel Psychology*, *28*(4), 563–575 (https://doi.org/10.1111/j.1744-6570.1975.tb01393.x). Copyright 1975. by C. H. Lawshe, A paper presented at Content Validity II, A conference held at Bowling Green State University.

Exploratory factor analysis was used to evaluate the validity of the structure. For this purpose, our research community was divided into two groups: students who had been through at least two semesters virtually using the learning management system and professors who had at least two years of virtual teaching experience. Regarding the minimum size of the sample required for performing the factor analysis, the ratio of the variable to the subject should at least one to five: according to the number of items entered to perform the exploratory factor analysis, the sample size (35 items) for the study (190 people) was sufficient. To compensate for the loss of samples, 10 samples were added to the total number received. Therefore, 200 questionnaires were considered for completion. The research sample was taken randomly from all e-learning executor universities in Tehran city, Iran.

Finally, the reliability of the questions was determined. The questions were examined from two dimensions of internal and external reliability. Internal consistency was obtained by calculating Cronbach’s alpha (α), and the Pearson correlation coefficient was used to determine the external reliability of the questionnaire (Appendix). The SPSS 21 software package was used for statistical analysis of data. The significance level in this study is $p < 0.001$.

Results

The mean age of participants in the present study was 59.41, with a standard deviation was 32.8. Participants included 81 men and 119 women (40.5% and 59.5%, respectively). Regarding education level, 46% participants had a master’s degree, 40% had a specialized doctorate, and 9% had a specialist degree. Of the participants, 57% were faculty members, 38% were experts, 68% had teaching experience, and 32% had no teaching experience. Regarding whether participants had a history of e-learning, 65% did and 37% did not.

The correlation, mean, standard deviation, and Cronbach's α of each of the factors obtained in the present study are presented in Table 1.The correlation between each of factors of research with the other factors such as Anaging Supportive Interactions and Behaviors, Course Management, Meta-cognitive Skills Management, Conflict Management, and Time Management showed an appropriate and high-scale correlation. Therefore, the questions were not changed, and none of the questions were deleted. Cronbach's α of all components of the questionnaire is 0.7, and the reliability of the whole set of 35 questions is equal to 0.93, which indicates high reliability of the questionnaire.

Table 1

Agents' Descriptive Statistics and Cronbach's Alpha

Row	Factor	1	2	3	4	5	Mean	*SD*	Cronbach's α
1	Time management	1	-	-	-	-	17.79	7.99	0.74
2	Course management	0.43**	1	-	-	-	37.44	9.31	0.9
3	Conflict management	0.45**	0.67**	1	-	-	8.56	3.15	0.75
4	Managing supportive interactions and behaviors	0.49**	0.71**	0.79**	1	-	36.01	11.65	0.94
5	Meta-cognitive skills management	0.41**	0.63**	0.64**	0.85**	1	6.78	2.78	0.84

Note. ** indicate the significance level in this study is $p < 0.001$

After examining the statistical characteristics of the scales and their alpha, exploratory factor analysis was performed on the factors. Bartlett's sphericity test was used to perform the principal component analysis method and to show that the data correlation matrix was not zero in the population. The results were statistically significant ($p < 0.001$ and KMO = 0.763). It should be noted that the Kaiser-Meyer-Olkin (KMO) Test is a measure of how suited your data is for Factor Analysis. The test measures sampling adequacy for each variable in the model and for the complete model. The results of Bartlett's sphericity test show that that the implementation of factor analysis based on the obtained correlation matrix is explainable. To determine that the measurement tool under study (set of questions) is made up of several factors, three eigenvalue indices, the ratio of variance explained by each factor, and a rotated eigenvalue diagram were examined.

To extract the appropriate factors, factor analysis was performed several times. Finally, it was found that according to the main structure of the questionnaire and the results of exploratory factor analysis, the five-factor is more sufficient, and this five factor was used. The results showed that the eigenvalues (Table 2) of five factors are greater than one and explain the percentage of common variance coverage between variables, among which the first factor (management of interactions and supportive behaviors) with an eigenvalue of 7.82 about 20.07%, and the fifth factor (management of interactions and supportive behaviors) with a specific value of 3.65 explains about 9.37% of the total variance. A total of five factors explain 64.9% of the total variance of the variables. The values of special value, percentage of variance and percentage of the cumulative variance of the five factors are shown in Table 2.

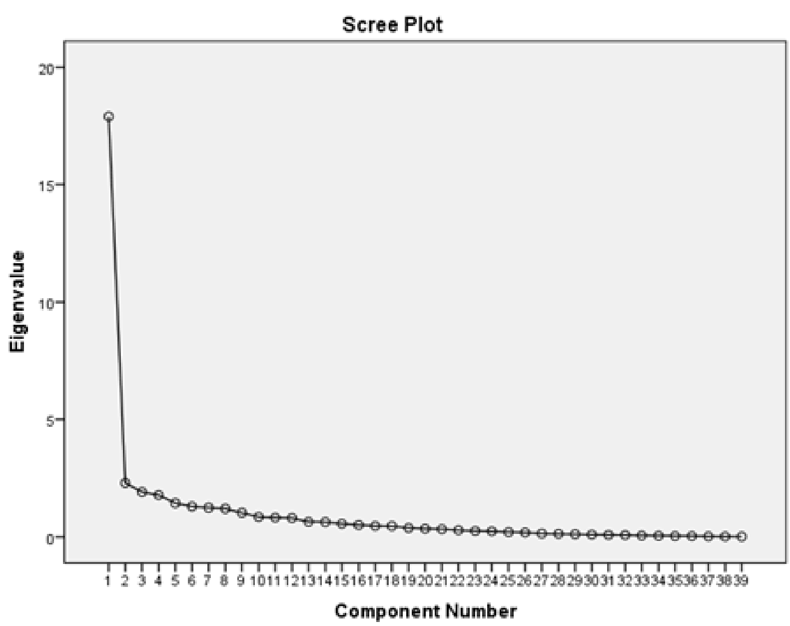
Table 2

Amounts of Special Value, Percentage of Variance, and Percentage of the Cumulative Variance

Factors	Special value	Variance (%)	Cumulative variance (%)
Time management	7.82	20.07	20.07
Course management	4.9	12.57	33.65
Meta-cognitive skills management	4.74	12.15	44.8
Conflict management	4.18	10.72	55.53
Managing supportive interactions and behaviors	3.65	9.37	64.9

Figure 2

Rotated Factor Special Value (Scree Plot)



The share of the first to fourth factors in the variance of the total variables is significant and different from the share of other factors (Figure 2). Also, from the fifth factor onward, the slope of the graph is cut and almost smoothed. The varimax rotation method was used to simplify the extraction of agents and their naming. The factor matrix created by the varimax rotation is shown in Table 3. Based on the results of factor analysis of the questionnaire materials, implementing the principal components analysis method, five factors were determined based on the final characteristics.

Table 3

Structure Matrix of 35 Questions With Varimax Method

Item	Time management	Conflict management	Meta-cognitive skills management	Course management	Managing supportive interactions and behaviors	CVR
32	-	-	-	-	0.79	0.80
34	-	-	-	-	0.76	0.75
35	-	-	-	-	0.73	0.75
33	-	-	-	-	0.71	0.87
36	-	-	-	-	0.66	1
29	-	-	-	-	0.55	1
26	-	-	-	-	0.54	1
25	-	-	-	-	0.51	0.62
28	-	-	-	-	0.50	0.75
31	-	-	-	-	0.49	0.80
30	-	-	-	-	0.45	0.87
27	-	-	-	-	0.3	1
22	-	-	-	0.7	-	1
16	-	-	-	0.68	-	0.75
13	-	-	-	0.61	-	0.75
15	-	-	-	0.61	-	0.75
21	-	-	-	0.6	-	0.87
10	-	-	-	0.58	-	0.80
20	-	-	-	0.56	-	0.62
14	-	-	-	0.52	-	0.80
19	-	-	-	0.39	-	0.87
17	-	-	-	0.38	-	1
12	-	-	-	0.31	-	0.80
38	-	-	0.64	-	-	0.75
39	-	-	0.64	-	-	0.87
37	-	-	0.33	-	-	1
18	-	0.69	-	-	-	0.62
24	-	0.48	-	-	-	0.75
23	-	0.32	-	-	-	1
3	0.72	-	-	-	-	0.87
2	0.67	-	-	-	-	1

5	0.63	-	-	-	-	0.62
1	0.6	-	-	-	-	0.75
6	0.48	-	-	-	-	1
4	0.33	-	-	-	-	0.75

Note. CVR = content validity ratio.

Discussion

The purpose of this study was to design a tool to measure the effective management of the virtual classroom from the perspective of professors and students in e-learning and to evaluate this tool's validity and reliability. Cronbach's α of all components of the questionnaire is 0.7, and the reliability of the whole set of 35 questions is equal to 0.93, which indicates the questionnaire has high reliability. The study revealed no significant difference in teachers' leadership styles based on gender or age; this is consistent with Adedigba and Sulaiman's (2020) results.

The results show that the eigenvalues of five factors are greater than one and explain the percentage of common variance coverage between variables, among which the first factor (management of interactions and supportive behaviors) with an eigenvalue of 7.82, about 20.07%, and the fifth factor (management of interactions and supportive behaviors) with a specific value of 3.65 explains about 9.37% of the total variance. A total of five factors explain 64.9% of the total variance of the variables. This corroborates Jones and Jone's (2012) assertion that effective classroom management is characterized by a safe environment and respect for pupils' opinions, creates confidence that students' ideas and opinions are valued, and gives them motivation to learn (Adedigba & Sulaiman, 2020). Lettink's (2020) results also indicate that the Classroom Management Questionnaire had a high construct validity as well as high reliability, in the sense that classes of students awarded their teachers similar scores. This is consistent with our study.

Classroom management guidelines are implemented to enhance students' behavior and increase academic achievements across all grade levels (Emmer & Sabornie, 2015). Online environments continue to be a popular career development option in education (Burkman, 2012; Herbert et al., 2016). The Internet and advances in information technology enable the creation of online spaces that provide instant access to research material and real-time interaction between faculty and students (Dash et al., 2012). Online professional development is flexible, allows participants to manage their educational and professional activities with personal responsibilities, and often increases access to resources that may not be otherwsie available (Vu et al., 2014).

Research on online professional development research shows that online environments are useful for those who take advantage of this opportunity and that there is demand for it (Acar & Yıldız, 2016; Baker et al., 2016). Teachers need efficient classroom management skills to cope with day-to-day challenges, but the time, money, and resources needed for comprehensive professional development are not always available. For this reason, many universities focus on the professional development of classroom management in their teacher training programs (Stobaugh & Houchens, 2014).

Conclusion

Classroom management refers to the wide range of skills and techniques that teachers use to organize, focus, perform tasks, and produce academically throughout the classroom (Glossary of Education Reform, 2018). Classroom management can be a difficult topic because the term implies that teachers must take on an authoritative role as a manager. As more and more classrooms embrace a blended learning environment, we must talk about strategies for classroom management. For the past few years, we hve seen a huge growth in distance learning. It has been suggested that despite the unique nature of the online learning environment, many of the same features that are essential to the success of a traditional classroom management plan also apply to the online classroom (Stewart, 2008). Virtual classes can be as productive and convenient as traditional classes. If possible, barriers are considered and preventive strategy management is used. These barriers include the special needs of students in cyberspace, the feelings of isolation, and the lack of face-to-face interaction with the teacher and other learners (Wilson, 2004). An essential aspect of teaching quality is classroom management, as it is a prerequisite for effective student learning. However, as far as we know, no instruments specifically measure this in Iran yet.

In this study, we were able to design a tool to measure the effective management of the virtual classroom from the perspective of professors and students in e-learning, and an approved tool for the use of educators and students in online environments was introduced. This tool is valid and reliable. It could be used during the teacher activities to indicate how to be efficient online classroom management is.

Acknowledgments

We would like to thank participants and all those who helped us to conduct this research.

References

Acar, I. K., & Yıldız, S. M. (2016). Professional development of elementary school teachers through online peer collaboration: A case study. *Turkish Online Journal of Qualitative Inquiry*, *7*(4), 422–439. https://doi.org/10.17569/tojqi.79480

Adedigba, O., & Sulaiman, F. R. (2020). Influence of teachers' classroom management style on pupils' motivation for learning and academic achievement in Kwara State. *International Journal of Educational Methodology*, *6*(2), 471–480. https://doi.org/10.12973/ijem.6.2.471

Akman, Y. (2020). The role of classroom management on confidence in teachers and educational stress. *International Journal of Contemporary Educational Research*, *7*(1), 335–345. https://doi.org/10.33200/ijcer.687109

Babadjanova, N. (2020). Effective classroom management techniques for curriculum of 21st century. *Science and Education*, *1*(7), 281–285. https://openscience.uz/index.php/sciedu/article/view/518

Baker, C., Gentry, J., & Larmer, W. (2016). A model for online support in classroom management: Perceptions of beginning teachers. *Connecting Education, Practice and Research*, *6*(1), 22–37. https://eric.ed.gov/?id=EJ1104337

Bakia, M., Shear, M., Toyama, Y., & Lasseter, A. (2012). *Understanding the implications of online learning for educational productivity*. US Department of Education, Office of Educational Technology. https://tech.ed.gov/files/2013/10/implications-online-learning.pdf

Berger, J. L., & Girardet, C. (2020). Vocational teachers' classroom management style: The role of motivation to teach and sense of responsibility. *European Journal of Teacher Education*, *44*(22), 1–17. https://doi.org/10.1080/02619768.2020.1764930

Burkman, A. (2012). Preparing novice teachers for success in elementary classrooms through professional development. *Delta Kappa Gamma Bulletin*, *78*(3), 23–33. http://www.deltakappagamma.org/NH/Spring%202012_Professional%20Development_2-27-12.pdf#page=24

Curzon, L.B. (2003). *Teaching in further education: An outline of principles and practice*. A&C Black Press.

Dash, S., de Kramer, R. M., O'Dwyer, L. M., Masters, J., & Russell, M. (2012). Impact of online professional development on teacher quality and student achievement in fifth-grade mathematics. *Journal of Research on for Technology in Education*, *45*(1), 1–26. https://doi.org/10.1080/15391523.2012.10782595

Emmer, E. T., & Sabornie, E. J. (2015). Introduction to the second edition. In E. T. Emmer & E. J. Sabornie (Eds.), *Handbook of classroom management* (2nd ed., pp. 3–12). Routledge.

https://www.routledge.com/Handbook-of-Classroom-Management/Emmer-Sabornie/p/book/9780415660334

Ghiasvandian, S., Varaei, S., Yadeghari, M. A., Pourrahimi, A., & Aghajanloo, A. (2017). Classroom management for the instruction of unmotivated students. *Journal of Medical Education and Development*, *12*(1–2), 13–26. http://jmed.ssu.ac.ir/article-1-769-en.html

Glossary of Education Reform. (2018). https://www.edglossary.org/

Gold, B., Pfirrmann, C., & Holodynski, M. (2021). Promoting a professional vision of classroom management through different analytic perspectives in video-based learning environments. *Journal of Teacher Education*, *72*(4), 431–447. https://doi.org/10.1177/0022487120963681

Greenberg, J., Putnam, H., & Walsh, K. (2014). *Training our future teachers: Classroom management*. National Council on Teacher Quality. https://eric.ed.gov/?id=ED556312

Hapsari, E. W. (2020). Error analysis in writing descriptive text of pre-service English teachers at Kutai Kartanegara University. *Intelegensia: Jurnal Pendidikan dan Pembelajaran*, *5*(1).

Hardman, E. L., & Smith, S. W. (2003). Analysis of classroom discipline-related content in elementary education journals. *Behavioral Disorders*, *28*(2),173-86. https://eric.ed.gov/?q=classroom+discipline&ff1=subTeacher+Role&id=EJ673113

Herbert, S., Campbell, C., & Loong, E. (2016). Online professional learning for rural teachers of mathematics and science. *Australasian Journal of Education Technology*, *32*(2), 99–114. https://doi.org/10.14742/ajet.2159

Jones, V. F., & Jones, L.S. (2012). *Comprehensive classroom management: Creating communities of support and solving problems* (10th ed.). Pearson.

Kayıkçı, K. (2009). The effect of classroom management skills of elementary school teachers on undesirable discipline behavior of students. *Procedia—Social and Behavioral Sciences*, *1*(1), 1215–1225. https://doi.org/10.1016/j.sbspro.2009.01.218

Keshavarz, M., & Ghoneim, A. (2021). Preparing educators to teach in a digital age. *The International Review of Research in Open and Distributed Learning*, *22*(1), 221–242. https://doi.org/10.19173/irrodl.v22i1.4910

Konti, F. (2011). Teachers and students perceptions towards teachers classroom management applications in primary schools. *Procedia—Social and Behavioral Sciences*, *15*, 4093–4097. https://doi.org/10.1016/j.sbspro.2011.04.503

Langdon, C. A. (1997). The fourth Phi Delta Kappa poll of teachers' attitudes toward schools. *Phi Delta Kappan*, *78*(3), 212–220. https://eric.ed.gov/?id=EJ555409

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Lettink, M. (2020). *Development and validation of a questionnaire on classroom management of Dutch secondary education teachers* [Master's thesis, University of Twente]. http://purl.utwente.nl/essays/85419

Levin, J., & Nolan, J. F. (2014). *Principles of classroom management: A professional decision-making model* (7th ed.). Pearson. https://eric.ed.gov/?id=ED577251

Malik, S., Azeem, M. F., & Paracha, A. T. (2020). Veiled and unveiled university faculty: A comparison on their needs for effective classroom management in relevance with the Islamic perspective. *Journal of Islamic Business and Management*, *10*(1), 260. https://doi.org/10.26501/jibm/2020.1001-016

Muijs, D., & Reynolds, D. (2017). *Effective teaching: Evidence and practice* (3rd ed.). Sage Publications.

Özcan, G., & Gülözer, K. (2017). Sınıf yönetimi ölçeğinin geliştirilmesi [Development of class management scale]. *Kastamonu Eğitim Dergisi*, 25(3), 1133-1146. Retrieved from https://dergipark.org.tr/en/pub/kefdergi/issue/29417/320749#article_cite

Stewart, D. (2008). Classroom management in the online environment. *Journal of Online Learning and Teaching*, *4*(3), 371–374. https://jolt.merlot.org/vol4no3/stewart_0908.pdf

Stobaugh, R., & Houchens, G. (2014). Preparing for success. *Principal Leadership*, *14*(7), 36–40.

Tulley, M., & Chiu, L. H. (1995). Student teachers and classroom discipline. *Journal of Educational Research*, *88*(3), 164–171. https://doi.org/10.1080/00220671.1995.9941295

Vu, P., Cao, V., Vu, L., & Cepero, J. (2014). Factors driving learner success in online professional development. *The International Review of Research in Open and Distance Learning*, *15*(3), 121–139. https://doi.org/10.19173/irrodl.v15i3.1714

Waltz, C. F., Strickland, O., & Lenz, E. R. (2010). *Measurement in nursing and health research*. Springer. https://dl.uswr.ac.ir/bitstream/Hannan/138859/1/9780826105080.pdf

Wilson, B. G. (2004). Designing e-learning environments for flexible activity and instruction. *Educational Technology Research and Development*, *52*(4), 77–84. https://doi.org/10.1007/BF02504720

Wong, H. K., & Wong, R. T. (2018). *How to be an effective teacher: The first days of school* (5th ed.). Harry K. Wong Publications. https://www.effectiveteaching.com/store/products/books/the-first-days-of-school-5th-edition-ebook

Yılmaz, H., & Çavaş, P. H. (2008). The effect of the teaching practice on pre-service elementary teacher's science teaching efficacy and classroom management beliefs. *Eurasia Journal of Mathematics,* *Science and Technology Education*, *4*(1), 45–54. https://doi.org/10.12973/ejmste/75305

Appendix

Questionnaire on How to Manage Virtual Classes

Biographical information

Age: ……………..

Gender: male/female

Degree: ……………..

Field of study: ……………..

Job status: employed/unemployed

Job category: bachelor's/master's/faculty member

Years of work: less than 5 years/between 5 and 10 years/more than 10 years

Have you ever had teaching experience at a university? Yes/No

Have you had an e-learning experience? Yes/No

Row	Item	Strongly Agree	Agree	No idea	Disagree	Strongly Disagree
1	Content is uploaded regularly and according to a predetermined schedule.					
2	A specific day and time for online responses are set by professors.					
3	A specific date has been set for face-to-face meetings with professors.					
4	There is a good time frame for submitting assignments.					
5	Course time is provided online and offline at the beginning of the semester.					
6	If the curriculum is changed, an alternative date will be announced.					
7	If you take the test online, the time will be determined in advance.					
8	Course topics are fully introduced and described.					

9	Teachers are well versed in the subjects.	
10	Prerequisite knowledge is assessed before the start of the course.	
11	The expectations and evaluation criteria of the student are announced orally at the beginning of the semester.	
12	The evaluation criteria at the end of each semester are clearly stated.	
13	Training is provided clearly and transparently.	
14	Content presented with a logical structure.	
15	The volume of content and assignments requested is proportional to the number of courses.	
16	The course content is presented by the different ways of learning of learners.	
17	New technologies (social networks, animation, simulation, etc.) are used for education.	
18	All student assignments are given feedback.	
19	The tutorials provided are useful and practical.	
20	A variety of educational resources are provided for learners to learn.	
21	The capabilities of the e-learning system are well used.	
22	Periodic assessments of learners' learning are done throughout the semester.	
23	The students can freely express their views on the e-learning course.	
24	Group conflict management is well done among learners during the course.	
25	Group activities and interactions in the e-learning environment are considered and encouraged.	
26	During the course, professors have friendly interactions and show supportive behaviors.	

27	Friendly interactions between professors and students are done through social networks, e-mails, wikis, etc.	
28	Professors provide guidance and design for educational activities.	
29	Professors make their personal information and educational records available to students.	
30	In the e-learning course, the real presence of the teacher–student is felt.	
31	There are incentive strategies to improve the quality of learners' learning.	
32	There are certain rules and regulations for managing the e-learning environment between student and teacher.	
33	All rules and requirements of virtual communication are observed.	
34	Professors and students are required to follow educational rules.	
35	Professors and students are required to observe ethical principles in the educational environment.	
36	I became acquainted with the principles of communication in cyberspace by the professors of the course.	
37	Teaching methods challenge professors and students.	
38	Homework is based on problem-solving and high-level mental activities.	
39	Student self-assessment is one of the final assessment criteria in e-learning courses.	

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

Ukrainian E-Learning Platforms for Schools: Evaluation of Their Functionality

Maryna Zhenchenko[1], Oksana Melnyk[2], Yaroslava Prykhoda[3], and Igor Zhenchenko[4]
[1,3]Taras Shevchenko National University of Kyiv, Ukraine; [2]Institute of Educational Content Modernization, Ukraine; [4]Borys Grinchenko Kyiv University, Ukraine

Abstract

This article defines 27 criteria for evaluating the functionality of e-learning platforms, grouped into three macro groups: (a) learning management, (b) learning content management, and (c) communications and collaboration tools. The proposed criteria can be used to evaluate any e-learning platform's functionality. They allow teachers and administrators to make conscious choices about the highest-quality e-learning platform for their schools and developers to improve e-learning platforms' functionality. The developed criteria became the basis for rating the functionality of Ukrainian developers' eight e-learning platforms' and determining the degree of support (in whole or partly) of e-learning components, categorized on the cognitive, social constructivist, motivation, and e-learning theories (CT, SCT, MT, and E-LT). The results indicate that the lack of communication and collaboration tools necessary to ensure quality distance learning is the main problem of Ukrainian e-leaning platforms. Comparative analysis of the functionality of e-learning platforms and components categorized on the learning theories helped determine that only three of the eight Ukrainian e-learning platforms (Accent [Mobischool], Class Assessment, My Class) fully follow the CT, SCT, and MT, but these platforms are all commercial products; therefore, they only partially support the E-LT. Solving this problem will be facilitated by developing e-learning platforms with open access, financed by the state budget in the context of the development of open and distance learning for Ukrainian students, as well as improving communication and collaboration tools in the context of conforming e-learning components to the social constructivist learning theory.

Keywords: e-learning platform, evaluation, functional suitability, open and distance education, learning theories

Introduction

The COVID-19 pandemic has affected all educational institutions in the world. The government of Ukraine, trying to restrain the spread of COVID-19, decided to close all educational institutions on March 12, 2020. This lockdown lasted until the end of the school year (May 2020). According to UNESCO's global monitoring, this nationwide closure affected 1,676,550 primary school students and 2,376,878 secondary school students in Ukraine (UNESCO, 2020), all of whom massively moved to distance learning.

Distance education is defined by the Association for Educational Communications and Technology as "institution-based, formal education where the learning group is separated, and where interactive telecommunications systems are used to connect learners, resources and instructors" (Parchure, 2016, p. 63). The distance learning format has actualized teachers' use of e-learning platforms that differ both in structure and offered functions. Piotrowski (2010, p. 20) defines *e-learning platforms* as "software that provides the technical infrastructure on which e-learning activities can take place." Ouadoud et al. (2016b, p. 582) emphasize that "a type LMS (Learning Management System) e-learning platform is a software including services to assist teachers in the management of their course" Ecoutin (2000, p. 5) describes *open and distance learning platforms* as software that assists in distance learning and combines the tools needed "for the three main users—teacher, student, administrator."

An important problem in the implementation of e-learning platforms in distance education is the lack of clear criteria for assessing their quality. In evaluating e-learning platforms, Tomczyk et al. (2020) looked at the following criteria in teachers' and students' surveys: general course quality, professionally prepared materials, content usefulness, visual design, and the innovative character of platforms. Pandu and Fajar (2019) and Abubakari et al. (2021) evaluated e-learning platforms via the User Experience Questionnaire (UEQ). The UEQ consists of six scales with 26 items reflecting the following basic components: (a) attractiveness, (b) dependability, (c) efficiency, (d) perspicuity, (e) novelty, and (f) stimulation (Abubakari et al., 2021, p 4; Schrepp, 2015).

However, surveys of students and teachers mostly provide their opinions about using e-learning platforms (Pandu & Fajar, 2019, p. 1) and are not sufficient in evaluating e-learning platforms' functionality and their ability to provide quality distance learning based on the e-learning components categorized by the following learning theories: cognitive theory (CT), social constructivist theory (SCT), motivation theory (MT), and e-learning theory (E-LT) (Kumar & Sharma, 2021; Schunk, 2020).

Defining criteria that can be used to rate the functionality of e-learning platforms, and to determine their effectiveness in the context of existing learning theories, will allow teachers and administrators to make a conscious choice about the highest-quality e-learning platform to use at their schools, and it will allow developers to see how to improve the functionality of their e-learning platforms.

The objectives of this study were twofold:

1. to define criteria for assessing the functionality of e-learning platforms; and
2. to perform a rating assessment of the functionality of Ukrainian developers' e-learning platforms and determine the degree of Ukrainian platforms' support (in whole or in part) of e-learning components categorized by the CT, SCT, MT, and E-LT learning theories.

Theoretical Framework for Evaluating E-Learning Platforms

Ouadoud et al. (2016a) developed the approach for the quality evaluation of e-learning platforms, which is based on "the quality model interactive systems" (Ouadoud et al., 2016a, p. 13) of International Organization for Standardization (ISO) standard number 25010: "The product quality model categorizes product quality properties into eight characteristics (functional suitability, reliability, performance efficiency, usability, security, compatibility, maintainability and portability). Each characteristic is composed of a set of related subcharacteristics" (ISO, 2011, s. 4.2). The researchers combined the characteristics presented in ISO standard 25010 into two categories—utility and usability—each of which was divided into subcategories (Ouadoud et al., 2016a, pp. 16–17, 19), which are presented in Tables 1 and 2.

Characteristics selected for evaluating an e-learning platform are developed via a software engineering approach with an emphasis on the technical aspects of the e-learning platform. In this study, we analyzed the quality of the e-learning platforms by one characteristic only: functional suitability. This helped us to determine how certain functionalities of a platform help with implementing e-learning components, defined according to different learning theories.

Table 1 summarizes the sub-characteristics of the functional suitability of e-learning platforms proposed by Ouadoud et al (2016a).

Table 1

Characteristics Selected for Evaluating the Functional Suitability of an E-Learning Platform

Functional completeness	Functional correctness	Functional appropriateness
• Forum • Synchronous causerie (cat/chat) • Virtual classroom (videoconferencing/webinar) • Sharing documents • Calendar • Awareness (list of connected people) • Tests management • Collaboration (Wikis) • Learners' management (registration, schedule, etc.) • Learners' management in working groups • Users' roles management • Customizable platform • Advancement scale or progression percentage in the course resources • Management (course) • Support of multiple authors	• Learners' and teachers' management of working time • Results and notes • Notes display • Course tracking statistics • Control connections (tracking of learners) • Reports on test results • Glossary • Reports on the frequency or use of a course	• Certification (certificate of training follow-up) • Foyer (family group) • Registration chat • Messaging • Plagiarism detection tools • RSS feed/podcast: means of distributing files (audio, video, other)

Note. Adapted from "Studying and Analyzing the Evaluation Dimensions of E-Learning Platforms Relying on a Software Engineering Approach," by M. Ouadoud, M. Y. Chkouri, A. Nejjari, and K. E. El Kadiri, 2016, *International Journal of Emerging Technologies in Learning*, *11*(1), pp. 16-17 (https://doi.org/10.3991/ijet.v11i01.4924). CC-BY.

The list of functional suitability sub-characteristics were modified and supplemented by functional characteristics, which, according to Colace et al. (2002), "must be absolutely present in an on-line learning platform." (Colace et al., 2002, p. 7) The following are the particular functional characteristics (Colace et al., 2002, pp. 6–7):

- progress tracking,
- multiple course management,
- student groups' creation and management,
- content inclusion in accordance with standards,
- content importation,
- new course creation in accordance with standards,
- course importation from other producers,
- reports on course frequency or use,
- test creation,
- course catalogue,
- multiple-choice tests,
- reports on test results, and
- automatic evaluation of tests.

Important services that are necessary for efficiently training the authors of the mentioned research online include "textual or vocal chat, whiteboard, live video stream, virtual classroom, application and file sharing" (Colace et al., 2002, p. 5).

Baggia et al. (2019, p. 53) combined various characteristics that "have to be considered when selecting the appropriate system for an individual case" into three main groups:

1. Learning content management: This includes content authoring, content storage and management, course libraries, compliance with standards for e-learning software Sharable Content Object Reference Model (SCORM) and Aviation Industry Computer-Based Training Committee (AICC) or Tin Can Application Programming Interface (API), and multimedia support.

2. Course management: This includes registration management, course catalogue management, course calendars, gradebooks, student and instructor portals, attendance tracking, proficiency testing, e-commerce capabilities (the ability to sell online courses), and virtual classrooms.

3. Social learning and collaboration: This includes support social learning with collaboration features (live chat, blog modules, Web conference integration, following concept, content sharing and rating, discussion boards, file sharing, integration with social media networks, profiling and expertise capabilities, and gamification tools).

The works of Baggia et al. (2019), Colace et al. (2002), and Ouadoud et al. (2016a) became the basis for developing criteria for assessing the functionality of e-learning platforms, presented in the results this study.

Kumar and Sharma (2021, p. 11) emphasize that "e-learning components, identified from the learning theories are very much important for any platform. If these components are not integrated in the platforms, the success of learning cannot be guaranteed." In analyzing the theoretical perspective of e-learning pedagogy, Kumar and Sharma (2021) derived the following characteristics for a successful e-learning framework: learner-centered, eco-sustainability, socioeconomic/cost-effectiveness, connectivity/networking, increased accessibility, on-demand availability, interaction, participation, cooperation, collaboration, motivation, engaging, communication, intrinsic motivation, extrinsic motivation, intriguing ideas, self-determination, competence, autonomy, relatedness, cognitive effectiveness, convenient, reliability, efficiency, achievement, personalization, self-pacing, constructive alignment, higher learning outcomes, learner satisfaction, confidence, peer review, evaluation/assessment/feedback from instructors, improved tracking, flexibility, skills and knowledge improvement, and learner satisfaction. These characteristics can be further divided according to the four major learning theories (CT, SCT, MT, and E-LT) (Kumar & Sharma, 2021, p. 4), shown in Table 2.

Table 2

E-Learning Components Categorized by Learning Theory

Cognitive theory	Social constructivist theory	Motivation theory	E-learning theory
• Learner satisfaction • Higher learning outcome • Cognitive effectiveness • Individual learning • Personalization • Achievement • Self-efficacy • Efficiency improvement • Skills and knowledge Improvement	• Collaboration • Interaction • Participation • Cooperation • Engaging • Communication • Constructive alignment • Peer review • Evaluation /assessment/feedback	• Motivation • Intrinsic motivation • Extrinsic motivation • Intriguing ideas • Self-determination • Competence • Autonomy • Relatedness • Confidence	• Learner-centered • Eco-sustainability • Socioeconomic/cost-effectiveness • Connectivity/ networking • Increased accessibility • On-demand availability • Convenience • Flexibility • Self-pacing • Improved tracking

Note. Components of e-learning are from "E-Learning Theories, Components, and Cloud Computing-Based Learning Platforms," by V. Kumar and D. Sharma, 2021, *International Journal of Web-Based Learning and Teaching Technologies*, *16*(3) p. 5 (https://doi.org/10.4018/IJWLTT.20210501.oa1). Copyright 2021 by IGI Global.

Our study of functionality of an e-learning platform in the context of their compliance of e-learning components, identified from the learning theories, will answer our main question: How well does an e-learning platform help to achieve the goals of the learning process?

Methods

This study was conducted in four stages: (a) searching scientific sources devoted to the problem of evaluating e-learning platforms, a literature analysis, and determining the criteria for evaluation of the functionality of e-learning platforms; (b) selecting e-learning platforms of Ukrainian developers for assessment; (c) rating assessment of the functionality of e-learning platforms of Ukrainian developers; and (d) evaluating Ukrainian e-learning platforms on a pedagogical approach and determining the degree of support (in whole or in part) of e-learning components, categorized by learning theories: CT, SCT, MT, and E-LT. At each stage, the research methods described below were used.

Stage 1

At the first stage of the research, we searched the Scopus database for e-learning platforms' problems of quality assessment and their compliance with e-learning theories with the following queries: *e-learning platforms*, *evaluation e-learning platforms*, and *e-learning theories* (only open-access journals were searched). The range of selected articles was expanded by analyzing references in the articles found with the specified search queries.

Further study of scientific articles allowed us to identify several studies whose results became the basis for the development of criteria for assessing e-learning platforms' functionality—in particular, Baggia et al. (2019), Colace et al. (2002), and Ouadoud et al. (2016a, 2016b).

Stage 2

Computerization and informatization of the Ukrainian education system are accompanied by the development of electronic educational resources (EERs) and e-learning platforms. The commission on informatization of educational institutions was established by order of Ukraine's Ministry of Education and Science, Youth and Sports, dated November 25, 2011, No. 1364 (as amended by the order of the Ministry of Education and Science, Youth and Sports of Ukraine, dated November 29, 2012, No. 1341) for state examination, granting permission for the use of EERs and e-learning platforms in the educational process in all secondary schools of Ukraine.

For the selection of e-learning platforms for evaluation, the protocols of the commission for 2016–2019 were analyzed, as the permission to use the platforms is granted for five years. Qualitative analysis of the protocols was based on the search phrases *educational platform*, *educational system*, and *online platform*, which allowed us to single out e-learning platforms among EERs of different types. Using an online information retrieval method, using the search query *distance learning platforms and services*, the All-Ukrainian School Online platform—developed in late 2020 with support from the Ukraine Ministry of Education and Science in response to distance learning challenges due to the COVID-19 pandemic—was added to our list of e-learning platforms.

Note that highly specialized platforms that are focused on in-depth study of a particular discipline—such as Lingva.Skills for the social project for learning foreign languages, Indigo Mental Training Club, and GIOS for learning mathematics—remained outside the scope of this study.

Stage 3

Developing the criteria for assessing e-learning platforms' functionality became the basis for qualitative analysis of e-learning platforms that were included in the list identified in the second stage. Each function was coded with one tag (+ or –) to remark on the presence or absence of a particular feature (i.e., the tag *learners' management* + is used to indicate the presence of the learners' management function).

The presence or absence of certain functionality of e-learning platforms was studied by qualitative analysis of information provided by developers on the sites of e-learning platforms (description of functionality, video presentation, etc.).

The final results consist of a set of evaluations composed of numerical ratings expressed in a range from 1 to 5, depending on the number of available functionalities in each of the three units. The maximum score a platform can receive is 15 points.

Stage 4

Comparative analysis of the functionality of e-learning platforms and e-learning components categorized on the learning allowed us to determine the degree of support (fully or partially) Ukrainian e-learning platforms of different e-learning components categorized on the CT, SCT, MT and E-LT. Mathematical methods were used for processing the survey results, and graphical methods were used to construct diagrams and tables.

Results

Criteria for Evaluating E-Learning Platforms' Functionality

The criteria proposed by Colace et al. (2002), Ouadoud et al. (2016a), and Baggia et al. (2019) were summarized, clarified, and grouped into three categories: (a) learning management, (b) learning content management, and (c) communications and collaboration tools (Table 3).

A learning content management system includes "all the functions enabling creation, description, importation or exportation of contents as well as their reuse and sharing" (Colace et al., 2002, p. 2). "Set of Tools represents all the services that manage teaching processes and interactions among users" (Colace et al., 2002, p. 9). Whereas distance education technologies are divided into two modes of delivery, namely, synchronous learning (all participants are present at the same time) and asynchronous learning (participants access course materials flexibly on their own schedule) (Parchure, 2016), communication tools between teachers and students are divided into "two fundamental categories: asynchronous communication tools and synchronous communication tools" (Colace et al., 2002, p. 5). Therefore, the described approach made it possible to propose 27 criteria for evaluating the functionality of e-learning platforms, presented in Table 3.

Table 3

Criteria of Evaluating E-Learning Platforms' Functionality

Learning management	Learning content management	Communication and collaboration tools
• Awareness (list of connected people) • Learners' management (registration, schedule, etc.) • Learners' management in working groups (student groups' creation and management) • Users' roles management • Advancement scale or progression percentage in the course resource • Management course (course catalogue, multiple course management) • Tracking of learners (progress tracking, reports on course frequency or use) • Management of tests (auto-evaluation tests, reports on test results) • Learning outcomes management (electronic class register, electronic diary) • Certification (certificate of follow-up training)	• Multimedia content (audio, video, flash, etc.) • Ready content from a developer • Content inclusion in accordance with standards • Constructor for creating teachers' content • Test constructor • Course importation from other producers • Content importation • Plagiarism detection • Sharing documents	Asynchronous communication tools • Forum • E-mail Synchronous communication tools • Textual or voice chat • Live video stream • Virtual classroom (videoconference/ webinar) • Application sharing • Whiteboard • Gamification tools

Rating Assessment of Ukrainian E-Learning Platforms' Functionality and Determination of the Degree of Their Support of Different Learning Theories

Table 4 presents the e-learning platforms selected for assessment and created by Ukrainian developers over the last five years. The platforms are arranged chronologically depending on the year of development. Each platform is assigned a code number (Table 4), which presents the results of a qualitative analysis of the functionality of the selected e-learning platforms in accordance with the evaluation criteria proposed in Table 3.

Table 4

E-learning Platforms of Ukrainian Developers

Platform code	Platform name	URL
1	Accent (Mobischool)	http://mobischool.ac-cent.com/
2	Class Assessment	https://klasnaocinka.com.ua/
3	My Class	https://miyklas.com.ua/
4	Pidruchnyk.ua	http://www.gutenbergz.com/ua/pidruchnyk.html
5	Euclid	https://www.euclidlms.com/
6	Classtime	https://www.classtime.com/uk/
7	The Only School	https://eschool-ua.com/
8	All-Ukrainian School Online	https://lms.e-school.net.ua/

Analysis results (Table 5) show that most of the considered e-platforms have similar functionality for learning management and learning content management. There is a lack of course importation from other producers and plagiarism detection functions across almost all platforms. The latter function, plagiarism detection, is especially important given the problem of academic integrity in the educational environment. Only the Classtime platform has the anti-cheating function. All commercial platforms are recommended for use in the educational process marked without content. The developers of the platforms offer teachers and students their own content (a set of test tasks, interactive exercises, theoretical materials on individual topics, etc.), but this content has not passed state examination. Only electronic versions of textbooks are recommended by the Ukraine Ministry of Education and Science on the Pidruchnyk.ua platform, and the All-Ukrainian School Online platform hosts electronic courses certified by experts from the Ministry of Education and Science of Ukraine.

Table 5

Analysis of the Functionality of E-Learning Platforms

Functionality criteria	E-learning platform							
	1	2	3	4	5	6	7	8
Learning management								
Management course (course catalogue, multiple course management)	+	+	+	+	+	+	+	+
Awareness (list of connected people)	+	+	+	+	+	+	+	+
Learners' management (registration, schedule, etc.)	+	+	+	+	+	+	+	+
Learners' management in working groups (student groups' creation and management)	+	+	+	–	+	+	+	–
Advancement scale or progression percentage in the course resource	+	+	+	–	+	–	+	–
Management of user roles	+	+	+	+	+	+	+	+
Management of tests (auto-evaluation of tests, reports on test results)	+	+	+	–	+	+	+	–
Learning outcomes management (electronic class register, electronic diary)	+	+	+	+	–	–	+	–
Tracking of learners (progress tracking, reports on course frequency or use)	+	+	+	–	+	+	+	+
Certification (certificate of follow-up training)	–	–	–	–	–	+	+	–
Learning content management								
Multimedia content	+	+	–	+	+	+	+	+
Ready content from a developer	+	+	+	+	–	+	–	+
Content inclusion in accordance with standards	–	–	–	+	–	–	–	+
Constructor for creating teachers' content	+	+	+	–	+	+	+	–

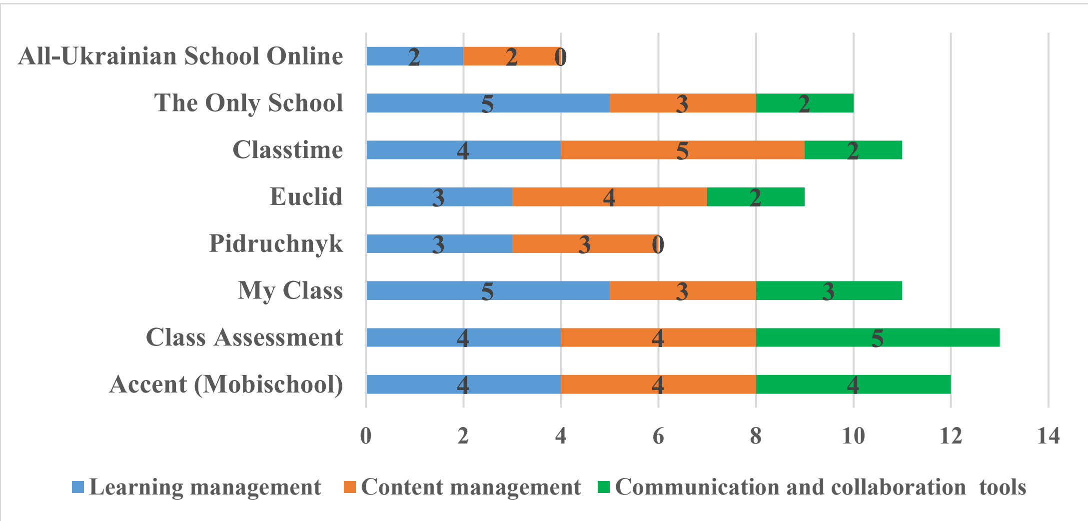
Tests constructor	+	+	+	–	+	+	+	–
Course importation from other producers	–	–	–	–	–	+	–	–
Content importation	+	+	+	-	+	+	+	–
Plagiarism detection	–	–	–	–	–	+	–	–
Sharing documents	+	+	–	–	+	+	–	–
Communication and collaboration tools								
Asynchronous communication tools								
E-mail	+	–	+	–	+	–	+	–
Forum	–	+	–	–	–	–	–	–
Synchronous communication tools								
Textual or voice chat	+	+	+	–	+	+	+	–
Whiteboard	+	+	–	–	–	–	–	–
Live video stream	–	–	–	–	–	–	–	–
Virtual classroom	+	+	+	–	–	–	–	–
Application sharing	–	–	–	–	–	–	–	–
Gamification tools	–	+	–	–	–	+	–	–

Note. 1 = Accent (Mobischool); 2 = Class Assessment; 3 = My Class; 4 = Pidruchnyk.ua; 5 = Euclid; 6 = Classtime; 7 = The Only School; 8 = All-Ukrainian School Online.

The top five e-learning platforms in terms of functionality are Class Assessment, Accent (Mobischool), My Class, Classtime, and The Only School (Figure 1).

Figure 1

Ratings of Analyzed Platforms



The availability of communication and collaboration tools—such as whiteboards and virtual classrooms, which are necessary to ensure quality distance learning—is most problematic in these e-learning platforms. The application sharing function, which is important during studying such subjects as computer science and technology, is not supported by any of the analyzed platforms. Only two analyzed platforms, Class Assessment and Classtime, support gamification tools. The first platform allows a creation of quizzes for students; the second allows a creation of team games and puzzles. Insufficient attention from developers of Ukrainian e-learning platforms to communication and collaboration tools negatively affects the quality of distance learning with these platforms: "Collaboration during problem-solving is one of the skills best promoted by modern e-learning" (Abubakari et al., 2021, p. 3).

An evaluation of the functionality of the e-learning platforms according to the criteria (Table 3) and the e-learning components categorized on the learning theories (Table 2) helped us determine how the analyzed platforms support the four learning theories (completely or partially). Only three of the analyzed commercial e-learning platforms' (Accent [Mobischool], Class Assessment, and My Class) support functions aimed at ensuring collaboration, interaction, extrinsic participation, cooperation, engagement, communication, constructive alignment, peer review, and evaluation/assessment/feedback—that is, they support all the components of SC theory: "The platforms that follow the Social Constructivist theory pedagogy will in turn deliver motivational and cognitive components" (Kumar & Sharma, 2021, p. 6).

The Accent (Mobischool), Class Assessment, and My Class e-learning platforms follow all three learning theory pedagogies completely. However, since these are commercial products, they only partially support the E-LT, which assumes increased accessibility and on-demand availability.

The Pidruchnyk.ua platform supports all learning theories only partially, as it provides access to electronic textbooks only and supports the functions of managing learning outcomes (electronic class register, electronic diary). The platform does not have tools for collaboration and group communication, for example.

The Euclid, Classtime, and The Only School platforms support all learning theories only partially as well, as they do not have sufficient collaboration, participation, and cooperation tools, and access requires payment.

The All-Ukrainian Online School platform, developed by order of the state, corresponds to E-LT, as it is free. However, it is an example of a mass open online course for middle and high school students, not a full-fledged e-learning platform, because it does not have many functionalities for learning management and content management, tools for communication between students and teachers, and teamwork organization. This platform therefore does not adhere the CT, SCT and MT:

> The educational content of the platform contains lessons in 18 main subjects: Ukrainian literature, Ukrainian language, Biology, Biology and Ecology, Geography, World History, History of Ukraine, Mathematics, Algebra, Geometry, Art, Basics of Law, Science, Physics, Chemistry, English and Foreign literature. Once launched, the content of the platform will be gradually supplemented according to the calendar plan. With the assistance of the International Renaissance Foundation, a mobile application of the All-Ukrainian School Online will be created soon and the functionality of the platform will be expanded, which will allow teachers to adapt teaching materials to the students' individual needs. (Ministry of Education and Science of Ukraine, 2020)

A significant disadvantage of the All-Ukrainian Online School platform is not only limited functionality, in particular, the lack of communication and collaboration tools, but also the focus on middle and high school students only. To develop effective distance education in school, it is necessary to place on this platform all 23 electronic textbooks for students of the first, second, fifth, and sixth grades, developed by Ukrainian publishers in 2018–2019 and recommended by the Ministry of Education and Science of Ukraine for use in the educational process (Zhenchenko et al., 2020, p. 732).

Discussion

Ouadoud et al. (2016a, 2016b) studied and analyzed the evaluation dimensions of e-learning platforms relying on a software engineering approach based on the quality model interactive systems of ISO standard no. 25010, which takes into account all technical aspects of interactive systems of e-learning platforms. In the context of our study, this model was used partly (functional suitability category). To evaluate e-learning platforms by criteria combined into the categories of performance efficiency, compatibility, security, maintainability, portability, and usability, more detailed technical information is needed. Colace et al. (2002, p. 7) distinguished, among the various functionalities of e-learning platforms, a representative number of the functionalities that must be absolutely present in any online e-learning platform. We accounted for the functionalities of e-learning platforms described by Colace et al. (2002) during the development of the e-learning platforms evaluation criteria.

Colace et al. (2002, p. 8) consider that "in order to accurately evaluate the potentials of an online learning platform, it is important to pay attention to its three main components: Learning Management System; Learning Content Management System; Virtual environment for teaching and services associated with it." Baggia et al. (2019) divide the functional characteristics of e-learning platforms into three major groups: (a) learning content management, (b) course management, and (c) social learning and collaboration. With this in mind, we have grouped the evaluation criteria of e-learning platforms into three macro groups: (a) learning management, (b) learning content management, and (c) communications and collaboration tools.

Assessing the functionality of e-learning platforms in the context of compliance e-learning components categorized on the learning theories (Kumar & Sharma, 2021) will allow developers to develop e-learning platforms that follow all four learning theories (CT, SCT, MT, E-LT) completely.

Various aspects of e-learning platforms' usability need further research via the UEQ (Pandu & Fajar, 2019, Abubakari et al., 2021) to improve teachers' and students' ability to use them. An in-depth assessment of student–teacher interaction through e-learning platforms can be based on Responsive Interactions for Learning (RIFL) measures—educator (RIFL-Ed) version (Rodrigues et al., 2021).

Conclusion and Implications

To assess the functionality of an e-learning platform, 27 criteria have been defined. They were grouped into three macro groups: (a) learning management, (b) learning content management, and (c) communications and collaboration tools. These criteria became the basis for rating assessment and determining the degree of support for various learning theories of the seven Ukrainian commercial platforms (Accent [Mobischool], Class Assessment, My Class, Pidruchnyk.ua, Euclid, Classtime, and The Only School) and the free platform All-Ukrainian School Online, developed in December 2021 by the Ministry of Education and Science of Ukraine to solve the problem of accessibility within quality distance education in Ukraine.

The main problem with Ukrainian e-leaning platforms is the lack of communication and collaboration tools necessary to ensure quality distance learning. The most common means of communication that support an e-learning platform are e-mails and chats. Only two platforms (Accent [Mobischool] and Class Assessment) provide whiteboard and virtual classroom functions, and two platforms (Class Assessment and Classtime) have gamification tools. Therefore, only three of the eight e-learning platforms follow the CT, SCT, and MT theories completely, but these are commercial products; hence, they support E-LT only partially.

The proposed criteria for assessing the functionality of e-learning platforms in the pedagogical aspect, taking into account the support of e-learning components according to the e-learning theories, can be used to assess and test functionality in developing new e-learning platforms and improving functionality in already-existing ones.

References

Abubakari, M. S., Nurkhamid, & Hungilo, G. (2021). Evaluating an e-learning platform at graduate school based on user experience evaluation technique. *c*, *1737*, Article 012019. https://doi.org/10.1088/1742-6596/1737/1/012019

Baggia, A., Leskovar, R., Rajkovič, U., Šmitek, B., Pesjak, K., Saša, M., Ravnik, S., Valjavec, M., Kingston, L., Murphy, F., Coffey, A., Narduzzi, B., Lillo-Crespo, M., & Riquelme, J. (2019, March 20–22). Selecting the e-learning platform for nursing education. In P. Šprajc, I. Podbregar, D. Maletič, & M Radovanović (Eds.), *38th International Conference on Organizational Science Development: Ecosystem of organizations in the digital age: Conference proceedings* (pp. 49–63). University of Maribor, University Press. https://doi.org/10.18690/978-961-286-250-3

Colace, F., De Santo, M., & Vento, M. (2002). Evaluating on-line learning platforms: A case study. In *Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS'03)* (pp. 1–9). IEEE Computer Society. https://www.computer.org/csdl/pds/api/csdl/proceedings/download-article/12OmNASILFE/pdf

Ecoutin, E. (2000). *Étude comparative technique et pédagogique des plates-formes pour la formation ouverte et à distance*. Fonds Social Européen. https://edutice.archives-ouvertes.fr/file/index/docid/ 1712/filename/synt08.pdf

International Organization for Standardization (ISO). (2011). *Systems and software engineering—Systems and software quality requirements and evaluation (SQuaRE)—System and software quality models* (ISO standard no. 25010:2011). ISO/International Electrotechnical Commission. https://www.iso.org/obp/ui/#iso:std:iso-iec:25010:ed-1:v1:en

Kumar, V., & Sharma, D. (2021). E-learning theories, components, and cloud computing-based learning platforms. *International Journal of Web-Based Learning and Teaching Technologies*, *16*(3), Article 1. https://doi.org/10.4018/IJWLTT.20210501.oa1

Ministry of Education and Science of Ukraine. (2020). *All-Ukrainian Online School*. https://mon.gov.ua/ua/tag/vseukrayinska-shkola-onlajn

Ouadoud, M., Chkouri, M. Y., Nejjari, A., & El Kadiri, K. E. (2016a). Studying and analyzing the evaluation dimensions of e-learning platforms relying on a software engineering approach. *International Journal of Emerging Technologies in Learning*, *11*(1), 11–20. https://doi.org/10.3991/ijet.v11i01.4924

Ouadoud, M., Chkouri, M. Y., Nejjari A., & El Kadiri, K. E. (2016b). Studying and comparing the free e-learning platforms. In *4th IEEE International Colloquium on Information Science and Technology (CiSt)* (pp. 581–586). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/CIST.2016.7804953

Pandu, F., & Fajar, A. N. (2019). E-learning implementation using user experience questionnaire. *Journal of Physics: Conference Series*, *1367*, Article 012015. https://doi.org/10.1088/1742-

6596/1367/1/012015

Parchure, N. P. (2016). "E-learning"—A boon for distance education. *International Journal of Management and Applied Science*, *2*(10), 63–66. http://www.iraj.in/journal/journal_file/journal_pdf/14-308-148007275163-66.pdf

Piotrowski, M. (2010). What is an e-learning platform? In Y. Kats (Ed.), *Learning management system technologies and software solutions for online teaching: Tools and applications* (pp. 20–36). IGI Global. https://doi.org/10.4018/978-1-61520-853-1.ch002

Rodrigues, M., Schneider, A., Sokolovic, N., Brunsek, A., Oré, B., Perlman, M., & Jenkins, J. (2021). Development and evaluation of an open-source, online training for the measurement of adult–child responsivity at home and in early childhood education and care settings. *The International Review of Research in Open and Distributed Learning*, *22*(3), 1–18. https://doi.org/10.19173/irrodl.v22i3.5472

Schrepp, M. (2015). *User experience questionnaire handbook*. https://doi.org/10.13140/RG.2.1.2815.0245

Schunk, D. H. (2020). *Learning theories. An educational perspective*. Pearson. https://www.pearson.com/us/higher-education/program/Schunk-Learning-Theories-An-Educational-Perspective-8th-Edition/PGM1996609.html

Tomczyk, Ł., Włoch, A., Potyrała, K., Wn˛ek-Gozdek, J. & Demeshkant, N. (2020). Evaluation of the functionality of a new e-learning platform vs. previous experiences in e-learning and the self-assessment of own digital literacy. *Sustainability*, *12*(23), Article 10219. https://doi.org/10.3390/su122310219

UNESCO. (2020). Education: From disruption to recovery. https://en.unesco.org/covid19/educationresponse

Zhenchenko, M., Melnyk, O., Miroshnychenko, V., and Zhenchenko, V. (2020). Electronic textbooks for Ukrainian education: Statistics, models of development, quality problems. *CEUR Workshop Proceedings* (pp. 721–733). http://ceur-ws.org/Vol-2732/20200721.pdf

Athabasca University

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

Using the Critical Incident Questionnaire as a Formative Evaluation Tool to Inform Online Course Design: A Qualitative Study

Anita Samuel[1] and Simone C. O. Conceição[2]
[1]Uniformed Services University of Health Sciences, Bethesda, Maryland, USA; [2]University of Wisconsin-Milwaukee, Milwaukee, Wisconsin, USA

Abstract

The online instructor plays a prominent role in influencing how students respond to an online course, from designing the course structure, course activities, and assignments to encouraging interaction. Therefore, to develop effective online courses, instructors need robust feedback on their design strategies. Student evaluation of teaching (SET) functions as a summative evaluation of the course design and delivery. Yet, the feedback from SETs can only be integrated into the next iteration of the course, thereby failing to benefit the students who provide the feedback. One suggestion is to use midsemester formative evaluation to inform course design in real time. A qualitative research study was conducted to explore whether the Critical Incident Questionnaire (CIQ) could be an effective formative evaluative tool to inform real-time online course design and delivery. Thematic analysis was conducted on the midcourse evaluations obtained from 70 students in six fully online master's level courses. There are three key findings from this study. First, CIQ use can provide opportunities for real-time adjustments to online course design and inform future redesign of online courses. Second, responses received via the CIQ prioritize the student voice and experience by focusing on factors that are critical to them. Finally, this deep-dive analysis reinforces the enduring factors that contribute to effective online course design and delivery. A recommendation for practice is to use the CIQ as an effective tool to gather formative feedback from students. This feedback can then be used to adjust course design as needed.

Keywords: student evaluation of teaching, Critical Incident Questionnaire, online course design, formative assessment

Using the Critical Incident Questionnaire as a Formative Evaluation Tool to Inform Online Course Design: A Qualitative Study

Student evaluation of teaching (SET) is standard practice in higher education. Evaluations are administered, usually at the end of an academic semester, attempting to measure teaching effectiveness; they function as a summative evaluation of the course design and delivery. SETs have gained importance as they inform tenure, reappointment, and promotion decisions (Uttl et al., 2017). Unfortunately, the SET is a flawed tool, and issues of bias associated with SETs are well documented (Boring et al., 2016; Mitchell & Martin, 2018; Reid, 2010). In the context of course design, there are two flaws: (a) the feedback from SETs cannot be used to make changes to course design in real time, and (b) SETs predominantly use surveys to gather quantitative data (Uttl et al., 2017).

SETs provide feedback that is intended to inform and enhance the design and delivery of a course. Yet, as a summative evaluation tool, the feedback from SETs can only be integrated into the next iteration of a course. The experiences of the students who provide the feedback are used to inform the design of the course for another group of learners, who might have very different responses to the course design (Gehringer, 2010). Moreover, the students who complete the SETs do not benefit from this course redesign (Gehringer, 2010). One way to address this is to use midsemester formative evaluation to inform course design in real time.

SETs are predominantly conducted via surveys that usually provide quantitative data (Erikson et al., 2016). Surveys limit the responses that students can provide. However, qualitative feedback tools allow students to go beyond predefined responses, encouraging them to delve deeper into their ideas about teaching and learning (Steyn et al., 2019). The Critical Incident Questionnaire (CIQ), a five-question open-ended questionnaire, has been used extensively in education and organizations as a tool of critical reflection and evaluation. Nevertheless, more research into how the CIQ can enhance online course design is needed (Keefer, 2009).

A qualitative research study using thematic analysis of the CIQ responses was conducted to explore whether the CIQ could be an effective formative evaluative tool to inform online course design and delivery in real time. This study answers the following research question: In what ways does the use of the CIQ as a formative evaluation tool contribute to online course design and delivery?

Literature Review

This study lies at the intersection of three concepts: SETs, the CIQ, and online course design. In this section, we explore literature related to these three concepts.

Student Evaluation of Teaching

SETs were originally intended to serve as a measure of teaching effectiveness (McKeachie et al., 1971; Rodin & Rodin, 1973) and have gained popularity since the 1960s (Rodin & Rodin, 1973). However, concerns about SETs range from the quality of the tool, the tool's lack of standardization, and cost implications of conducting these evaluations (Fisher & Miller, 2008). In addition, there are persistent issues with race and gender biases (Boring et al., 2016; Mitchell & Martin, 2018; Reid, 2010), and SETs primarily measure student perceptions of teaching rather than teaching effectiveness (Stark & Freishtat, 2014). Furthermore, SETs are summative evaluations provided at the end of a course. Any course design changes can only be made in the next iteration of the course. Gehringer (2010) notes that the feedback provided by summative SETs is too infrequent and broad in scope to effectively inform course design.

Formative midsemester student evaluations are more effective in providing actionable, real-time feedback for faculty during a course. There is an immediacy to the feedback, since it is provided while students are experiencing the course, giving it authenticity. Based on midsemester feedback, faculty can make changes to course design where possible and manage student expectations during the course (Veeck et al., 2016). These actions positively impact semester-end SETs, and faculty who receive midsemester feedback tend to receive higher ratings on end-of-semester evaluations (Cohen, 1980).

Formative midsemester evaluations have been conducted in various ways. Gehringer (2010) introduced a Google form evaluation tool that he used to gather student feedback on a daily basis after his face-to-face lectures. Fisher and Miller (2008) used a quantitative and qualitative midsemester assessment tool. Their qualitative questions focused on student expectations that elicited student responses that course instructors had not expected. Finelli et al. (2008) and Hurney et al. (2014) discuss the use of instructional consultants to gather midsemester feedback. The consultants conducted focus groups with the students and summarized their findings for the instructor.

The choice of evaluation tool used impacts the midsemester evaluation's effectiveness. Quantitative surveys tend to reflect the assumptions and biases of the designer, limit student responses, and not provide space for unique student experiences and contexts to emerge (Rao & Woolcock, 2003). Fisher and Miller's (2008) qualitative evaluation tool was designed for their specific context and focused primarily on student expectations, though they did get feedback about the design of the course as well. Focus groups with instructional consultants (Finelli et al., 2008; Hurney et al., 2014) are not anonymous and can be susceptible to groupthink. Furthermore, instructors were not receiving information directly from their students. Rather, they received a summarized version filtered through the instructional consultant's personal lens. An effective, robust, easy-to-use tool is needed to inform instructors about how their students are experiencing an online course.

Critical Incident Questionnaire

The CIQ, a five-question open-ended questionnaire, was proposed by Stephen Brookfield (1998) as a qualitative tool to elicit student feedback and to develop critically reflective practice in educators. Brookfield (1998) contends that educators' course design and pedagogical choices will be incomplete and ill-informed if they do not account for the student's voice. He states that the quality of teaching can only improve when instructors understand how their students are experiencing the course and the difficulties they are struggling with. The CIQ tool, as used initially by Brookfield, consists of five questions (see Figure 1) that are administered at the end of every class. The questions are general and focus on students' perceptions; students are not asked to identify what they liked or disliked in the class. Instead, the CIQ asks them to reflect on the class critically, and the responses students provide are guided by incidents that were significant for them in the class.

Figure 1

Critical Incident Questionnaire

1. At what moment in the class this week were you most engaged as a learner?
2. At what moment in the class this week were you most distanced as a learner?
3. What action that anyone in the room took this week did you find most affirming or helpful?
4. What action that anyone in the room took this week did you find most puzzling or confusing?
5. What surprised you most about this class?

Note. From S. Brookfield, "Critically reflective practice," 1998, *Journal of Continuing Education in the Health Professions*, *18*(4), p. 201 (https://doi.org/10.1002/chp.1340180402)

The CIQ has been used as a formative evaluation tool and also as a reflective tool. Linstrum et al. (2012) used the CIQ in a graduate-level course to obtain formative assessment data over two years. Their study identified four themes of course design: impact of instructor, student personal awareness, discussion mode teaching, and practical and applicable activities. Jacobs (2015) used the CIQ to evaluate course design and make changes to the course. Hessler and Taggart (2011) adopted the CIQ as a formative assessment and reflective tool for their writing courses. They found the CIQ insufficient for their needs and adapted it to include two more relevant questions to their context. Rather than only focusing on student feedback about the course, they used the CIQ to encourage students to reflect on their learning practices as well. In all these instances, the CIQ has been used in traditional face-to-face environments.

As college courses started moving online, in 2006, Brookfield adapted the CIQ for critical reflection in the online environment. But minimal research has been published on the use of the CIQ as an evaluative tool in online courses. Keefer (2009) conducted a literature review on the use of the CIQ and identified only two studies that used the CIQ in the online environment: Glowacki-Dudka and Barnett (2007) used the CIQ to study group development in online asynchronous graduate courses; and Phelan (2012) used the CIQ to explore students' perceptions of their online learning experiences. However, Glowacki-Dudka and Barnett (2007) and Phelan (2012) did not use the CIQ to inform course design, focusing instead on group and community development among the students. While anecdotal evidence exists that CIQ is used as a formative evaluation tool in online courses, research studies in this area are lacking.

Online Course Design

Best practices in online course design have been informed by various theories and models. Transactional distance, "a psychological and communications space" rather than a physical or temporal space (Moore, 1997, p. 22), was the defining theory of distance education. Moore (2013, p. 88) identifies three dimensions of distance education: (1) program or course "structure," (2) "dialogue" (interaction between instructor and student), and (3) "autonomy" of the learner. These three dimensions have been foundational to the various elements of online course design identified in the ensuing years.

Garrison et al. (1999) developed the community of inquiry (COI) framework to define a "worthwhile educational experience" (p. 88) in online education. The COI framework integrates social, cognitive, and teaching presence. Social presence encompasses the dialogue that Moore referred to, and teaching presence includes structure. Through cognitive presence, Garrison et al. (1999) address issues of student agency and motivation. Several other models and theories of online learning have been proposed, including Anderson's (2011) online learning model, Harasim's (2017) online collaborative learning theory, and Picciano's (2017) multimodal model for online education. These models and theories now include evaluation, reflection, learning resources, and learning modality.

Based on these models, several course evaluation instruments have been developed, such as the following:

- Blackboard Exemplary Course Program Rubric (Blackboard; Blackboard, 2017),
- Open SUNY Course Quality Review Rubric (OSCQR; State University of New York, 2018),
- Quality Learning and Teaching (QLT; California State University, 2019), and
- Quality Matters Higher Education Rubric (QM; Quality Matters, 2018).

These instruments vary in some ways, but they also share many course design elements that have been identified as best practices in online course design, including course structure and design, interaction, student activities, content or resources, course technology(ies), and assessment.

In addition to these evaluation tools, universities adopt their own evaluations of course design. The common feature of all these tools is that they are administered as summative evaluations or checklists prior to starting a course. These tools are not used for formative assessment of course design. Moreover, these tools do not prioritize student feedback.

The three intersecting bodies of literature—on student evaluations of teaching, the CIQ, and online course design—indicate that formative feedback, conducted via an appropriate tool, has the potential to provide instructors with real-time feedback on online course design.

Methodology

Study Context

This study was conducted at a public four-year university in the Midwestern United States, offering fully online courses. To study the effectiveness of the CIQ as an evaluation tool in online courses, the CIQ was incorporated as a midsemester evaluation tool in six fully online graduate-level courses in adult education and technology.

Different instructors taught the courses. However, the overall course design of all six courses was similar. All the courses were fully asynchronous with optional synchronous sessions with the instructor or other students. In addition, the courses included collaborative learning in the form of team projects and asynchronous group discussions. The asynchronous discussions in the courses were directed by student-generated discussion prompts that could generate meaningful dialogue. The students in these courses were practicing professionals and identified as adult students. Their experience with online learning ranged from no experience to having participated in multiple online courses.

The CIQ was distributed as a midsemester evaluation. The questions' phrasing was slightly adapted to account for the online environment and deployment of the tool once during the semester:

1. At what moment in the semester did you feel most engaged with what was happening? Why?
2. At what moment in the semester did you feel most distanced from what was happening? Why?
3. What action that anyone (you or anyone in your group or class) took in the online environment did you find most affirming and helpful? Why?
4. What action that anyone (you or anyone in your group or class) took in the online environment did you find most puzzling or confusing? Why?
5. What about the online environment during the semester surprised you the most? Why?

These evaluation questions were distributed via an anonymous online survey during week 7 of a 15-week academic semester. Institutional review board clearance was obtained, and participants were informed of their participation at the beginning of the midsemester evaluation survey. In total, 70 responses were received. The data from the surveys were entered into Microsoft Excel 365 and organized under the five CIQ questions.

While the original intention of the CIQ was to use it after every teaching session, in this study, the CIQ was only used once during an academic semester to avoid students experiencing feedback fatigue (Brookfield, 1998) at the end of the course.

Data Analysis

A semantic thematic analysis from a constructivist epistemology was conducted. The data were inductively analyzed following the six phases of thematic analysis outlined by Braun and Clarke (2006).

Phase one was familiarizing with the data. The data from the six courses were compiled and organized under the individual CIQ questions. The authors familiarized themselves with the data by reading and rereading the data, looking for patterns. Phase two was generating initial codes. As patterns were identified, the authors independently generated initial codes across the complete data set. Table 1 shows a sample of data extracted and the initial codes applied by the authors individually.

Table 1

Initial Coding

Data excerpt	Author 1 code	Author 2 code
The module discussions in which we were sharing experiences.	Sharing experiences	Sharing experiences
I met with my group on Skype[.] I felt most engaged because it felt more real.	Synchronous meeting	Synchronous face-to-face interaction
My group members are proving to be responsible for tasks, and willing to help with clarifications.	Surprised by responsive group members' accountability	Interaction with students, affirmative/supportive tone

Phase three was searching for themes. The authors met to review their individual codes. They organized the codes into potential themes and collated the data relevant to these themes. Phase four was reviewing themes. They checked to see if the themes worked at both the discrete extracted data and entire data set levels. Once the themes were confirmed, the authors developed a thematic map (see Table 2).

Table 2

Sample of Thematic Map

Themes	Sub-themes
Student–student communication	Support Sharing experiences Collaboration Feedback
Experience with online learning	Experienced Inexperienced
Type of communication	Asynchronous Synchronous

Phase five was defining and naming themes. The authors continued to refine the themes through categorization and renaming. Phase six was producing the report. Finally, the authors selected compelling extracts and rechecked them against the themes and the research question. See Table 3 for a sample.

Table 3

Sample Extracts of Student Comments

Subtheme	Code	Extracts from Student Feedback
Student–Instructor	Feedback	The professor's feedback is concise and [comes] in a timely manner. When the group or thread needs constructive criticism or more clarification, the professor jumps in to emphasize the need for more or better information.
Student–student	Collaboration	I felt most engaged when I was put into a group and start[ed] gaining different task[s] to do within my group.

Findings

The data analysis of student feedback received through the CIQ revealed five broad themes: interactions, expectations, course design, experience with online learning, and learners' sense of agency. These five factors affected students in different ways. The findings are presented using the CIQ questions as a framework. See Table 4 for a summary of findings.

Table 4

Summary of Findings

Factors	Findings
Engaging factors	• Student–student interaction • Robust communication • Relevant content • Learner sense of agency
Distancing factors	• Course design • Unclear expectations • Lack of peer interaction

Affirming factors	• Student–student interaction
	• Student–instructor interaction
	• Group dynamics
	  ○ Supportive
	  ○ Helpful
	• Learner sense of agency

Puzzling factors	• Peer interaction • Course design ○ Too many moving parts
Surprising factors	• Unexpected elements • Course design • Interaction with technology

Engaging Factors

The first question posed in the CIQ was "At what moment in the semester did you feel most engaged with what was happening? Why?" Interactions with peers and course content, quality of interactions, and learner sense of agency emerged as the key factors for engagement.

The participants in this study were overwhelmingly engaged by interactions with their peers. They identified both synchronous and asynchronous peer interactions as being engaging. Asynchronous group discussions were repeatedly mentioned, for example, "I feel engaged when I am responding to posts within my small group." These peer interactions were related to discussions regarding course content and, as mentioned by a student, "working on my group project with my group members." However, the peer interactions were effective within small groups rather than in the large class setting because "the general course discussion was overwhelming once more and more posts were added."

The quality of these peer interactions was also an influencing factor; one student mentioned "in-depth discussions that have been meaningful and thorough." Participants appreciated thoughtful and timely posts from their group members; one participant identified a turning point when "conversations took a turn into deeper analytical discussions, did I really feel engaged in the class and learning." Participants also appreciated "module discussions in which we were sharing experiences," creating "group discussions [that] were more conversational."

Synchronous activities were optional in the courses surveyed and were identified as engaging in the courses where students had participated in them. Participants mentioned they felt engaged "when I am Skyping with my group." The physicality of interactions in these synchronous meetings were specifically identified: "Now I know how they look, the way they talk etc. It is easier for me to relate to these people now"; "I felt most engaged because it felt more real. I think that having a real discussion and being able to hear someone talk are really important." The real-time immediacy of feedback in these synchronous sessions was also noted.

Participants also felt engaged when they found the course content relevant to their practice since "this brought the information full circle and to life, rather than just a theory" and "I was able to apply [what] I was learning first-hand." As one participant put it: "I think I was most engaged because I find these topics to be very interesting and where I would like to focus my research."

When participants took control of the learning environment and guided the direction of tasks and interactions, they found the experience engaging. One participant "felt engaged early on, as I took the responsibility for leading the first module and discussion on the readings."

Distancing Factors

Three factors caused students to feel distanced in the course: the course design, unclear expectations, and lack of peer interaction. Course design emerged as an oft-mentioned factor that created a feeling of distance within the course. Specifically, course design related to workload issues was a major contributor to participants' experience of distance. One participant said, "Reading from different texts, doing the book review, trying to get the tech plan. It seems like a little of everything all at once."

The distancing aspects of course design were exacerbated by unclear course instructions. When participants were unsure of what was expected of them, they expressed feeling distanced in the course. As one participant succinctly put it, "When I am confused about what I need to do or what is expected of me. I feel like just turning off." Another participant clarified: "I prefer to have very specific instructions, and at times I felt I needed more direction and felt distanced."

While peer interaction enabled participants to feel engaged with the course, lack of peer interaction and lack of a cohesive group dynamic distanced them from the course. But there was also an element of too much of a good thing, as some participants felt distanced when there was too much interaction: "I found myself feeling overwhelmed with the number of comments in the first module's discussion threads."

Affirming Factors

Interactions with peers, group dynamics, interactions with instructors, and learner sense of agency were identified as affirming actions. Group dynamics were repeatedly mentioned: "In general, the conversations my group has is affirming and helpful because everyone is very open, honest and complimentary." Candid conversations were appreciated and explicitly noted as this participant comments:

> One member of my group started out all the discussions with how he likes the DQs [discussion questions] to feel conversational. It has encouraged many in our group to follow suit. It has made the discussions much more lively and personable. Because of this, there are many times which we are supporting each other through sharing experiences and relating it to not only the text, but each other.

Meaningful communication was highly valued and noted by participants: "There were some very thoughtful and helpful comments and that was most helpful." The supportive nature of group dynamics was also identified as an affirming factor in the courses: "When the modules first began, I appreciated the fact that [the peer group] facilitator reached out to me to help me remember when postings were and requirements were to keep me a part of the group." One participant explained that "one of my group members was so helpful. … They encouraged me when I was getting anxious about our poor group participation. They also took on more responsibility within the group which made a positive impact on me."

Interactions with the course instructor were also noted as being affirming and helpful: "I like that my professor is involved and responds so quickly and very often." The timely nature of instructor responses and feedback was highlighted. This instructor presence "made me feel that the teacher actually is interested in what I had to say. It was nice to know she was 'there.'"

When learners exercised agency, they felt affirmed. "Stepping up to be the leader" was noted by participants as an affirming action, and "taking action and making a plan and schedule was something that really helped me."

Puzzling Factors

Just as lack of peer interaction was experienced as a distancing factor, it was also identified as the most puzzling part of the course: "Having group members not participate on a consistent basis" or "[w]hen members did not respond" confused some participants. Specific group members' actions were remarked upon as confusing and puzzling: for example, “The group that I was in for the book review seemed disinterested in the topic we had.”

Course design was also identified as confusing or puzzling. When there were too many moving pieces in the course, participants spoke about it as a confusing environment: “I am super confused by the different roles that we have in groups. This is mainly because we have both class discussions and a class project going on at the same time.” Another participant felt that “the discussion threads are confusing and overwhelming in this course. There are so many that I often find myself losing track of what we are talking about.”

Surprising Factors

Participants expressed surprise when something fell outside their expectations or what they were used to. Participants who had taken other online courses reported, “I have taken many online courses now and know what to expect and how to pace myself.” However, when participants encountered something different, some examples of their comments were as follows: “I was surprised by my reaction to the discussion board. I have taken classes that don’t involve the discussion board as much” and “I am spending much more time in the online class than traditional f2f [face-to-face]. I never thought it would require greater time commitment.”

Elements of the course design caused some participants to be either pleasantly or unpleasantly surprised: “I have enough time to think about the discussion and formulate a re[s]ponse that I feel good about posting,” said one participant, while another commented about “how the interface on D2L (the learning management system) did not look like other D2L class sites. It was a happy surprise.” One participant said, “The way threads are posted makes everything seem garbled together. Not enough separation yet too many places to check.”

Participants also expressed surprise when they had positive interactions with the technology: “This was the first time I used video in a response” and “I have not utilized OneDrive previously and am enjoying the benefits it provides with group tasks.” It was a negative surprise when they felt challenged by the technology: “I was surprised at the drastic changes [in the learning management system] and I’m still surprised that I can’t seem to adapt to this new environment.”

Discussion

Formative course evaluations are intended to help course instructors make changes to the online course in real time and enhance the student experience. However, for faculty to make changes to their courses, formative evaluations need to be robust and provide useful data. Findings from this study, conducted across six different fully online graduate-level courses, show that the CIQ can provide useful and actionable information for course instructors.

This study sought to answer the question, in what ways does the use of the CIQ contribute to online course design and delivery? There are three key findings from this study. First, the use of the CIQ for formative evaluation can provide opportunities for real-time adjustments to online course design and inform future redesign of online courses. Second, responses received via the CIQ prioritize the student voice and experience by focusing on factors that are critical to students. Last, this deep-dive analysis reinforces the enduring factors that contribute to effective online course design and delivery.

Informing Course Design

A key element of good online course design is consistency (Subramanian & Budhrani, 2020). This makes it challenging to implement changes to course design and delivery in real time. However, there are changes that instructors can make in real time to enhance student learning experiences.

The students in these courses highlighted the need for clear instructions and expectations. A lack of clarity in different activities, including assignment expectations, group interaction expectations, and instructor expectations, led to feelings of distance and confusion (Baldwin & Trespalacios, 2017). This is important feedback and easy for instructors to act on. After reviewing the midsemester evaluation, instructors can easily address points of confusion and clarify expectations (Gehringer, 2010).

Similarly, when students identify specific activities as unclear or elements of the course site as challenging to navigate, the instructor can correct that in real time by providing the necessary clarifications. Making these adjustments during the course has the most meaning for the students as it directly affects them. Acknowledging student experiences and trying to address concerns show students that the instructor is hearing them and is invested in their success (Dancer & Kamvounias, 2005).

Interactions within the course environment were frequently mentioned by the participants in this study. These students responded positively to proactive, timely feedback from instructors. This is positive reinforcement, and course instructors should make a note to actively maintain this form of interaction with the students.

The most challenging interaction to facilitate is student–student interaction, since it lies outside the instructor's locus of control (Samuel, 2020). Yet, instructors can encourage student–student interaction. When students engage positively with synchronous sessions, instructors can ensure that they offer more opportunities for this through the remainder of the course. Midsemester, instructors can provide appropriate feedback and encourage students who are less active to participate more. When responses from the CIQ midsemester evaluation is summarized and shared with students, highlighting student–student interaction could also encourage participation.

The graduate adult students in this study appreciated agency over their learning and valued readings and assignments that they found practical and applicable to their lives (Linstrum et al., 2012). Positive comments received through the CIQ reinforce the course design decisions of the instructor. Instructors also have the opportunity to assess their courses midsemester and ensure sufficient opportunities for students to exercise agency over their learning. Instructors might consider giving students a choice over the course content they engage with.

As illustrated above, some feedback can be acted upon in real time, but some changes can only be implemented in a future iteration of the course (Gehringer, 2010). The participants in this study clearly expressed that when a course design had too many moving parts, such as overlapping assignments, they experienced cognitive overload. The specific comments help instructors identify course elements that need to be adjusted. Changes to course timelines and assignments are not feasible in real time. However, this feedback, provided by students as they are experiencing the course, has immediacy and authenticity. This is valuable feedback for future course redesign.

Prioritizing Student Voices

The CIQ as a critically reflective tool was developed to remove the hegemony of the course instructor and create a democratic learning environment where students have a voice in their learning process. Using the CIQ as an anonymous midsemester evaluation tool reduces the power dynamic between student and instructor and elicits honest feedback about a course (Brookfield, 1998). The CIQ's open-ended questions allow students to speak to online course experiences that were critical to them. Reviewing comments received through the CIQ and modifying course design demonstrates to students that their experiences are meaningful and valued by the instructor.

The literature identifies three main types of interactions in online courses: student–student, student–instructor, and student–content (Moore, 1989). Relying on researcher-generated quantitative surveys, some studies have shown student–instructor interaction to be the most important (Kyei-Blankson et al., 2019; Linstrum et al., 2012; Martin & Bolliger, 2018; Swan, 2002). Other studies have shown that student–student interaction is more important (Bernard et al., 2009). In this study, the use of the CIQ revealed student–student interaction to be the most impactful interaction for participants. Using the CIQ helps bring clarity to instructors when research provides unclear findings. This is important for instructors as it shows them exactly what their students in a specific course are expecting and experiencing. Even though the findings might not align with the literature, acknowledging and addressing them as valid and unique to these participants is important.

Reinforcing Enduring Factors of Effective Online Course Design

All of this study's findings reinforce the factors that have been identified for effective online course design. Interaction appeared as a factor that engages and affirms students; but it could also make them feel distanced. Student–instructor interaction was highlighted as an affirming factor in this study, and student–instructor interaction has been repeatedly shown to have a significant effect on student learning, including predicting student success in a course (Crews et al., 2015; Jaggars & Xu, 2016; Martin & Bolliger, 2018). Student–student interactions were repeatedly mentioned by the participants in this study as impactful (Bernard et al., 2009).

Chunking course content (Martin et al., 2019; Subramanian & Budhrani, 2020) and maintaining consistency in the presentation and rhythm to the course (Subramanian & Budhrani, 2020) are important in an online course. The study participants noted that course sites that are difficult to navigate or that had too many elements were overwhelming and added to their cognitive load.

The study participants appreciated meaningful tasks, readings, and course content that had practical applications and were relevant to them (Linstrum et al., 2012). In addition, expectations for assignments need to be stated explicitly. Clear expectations also influence the quality of student–student interactions (Jaggars & Xu, 2016; Martin et al., 2019).

These findings echo the literature on best practices for online course design and validate the use of the CIQ tool. Using the CIQ can give instructors important information on where their courses might be deviating from best practices, offering them an opportunity to reflect on their course and reassess its design.

In this study, elements of course design were mentioned by students as a significant factor in their negative experiences of the course only. A lack of comments about the course design is an indicator to instructors that their course is functioning as expected.

Limitations of the Study

While the responses from the CIQ can be informative and guide course design, Keefer (2009) notes that the findings are not comprehensive. Students will only focus on incidents that have been impactful for them and will not necessarily provide a holistic review of the course. In this study, the nature of the CIQ questions guided the framing of students' responses, especially the third and fourth questions, which expressly referred to actions taken by course participants. This limits the scope of responses that the students provide to events that happened *to* them (Hessler & Taggart, 2011).

Furthermore, a participant self-selection bias is present in these surveys; usually, students with strong negative or positive responses to a course tend to respond to evaluation surveys (Wolbring & Treischel, 2015). It should also be noted that this study was limited to one department within one university, and the participants were graduate-level students. This might have affected the quality of their responses to the CIQ.

Conclusion and Next Steps

The online instructor plays a prominent role in influencing how students respond to an online course, from designing the course structure, course activities, and assignments to encouraging interaction. Therefore, to develop effective online courses, instructors need robust feedback on their design strategies. This study shows that the CIQ can be used as a tool to elicit useful formative evaluation feedback from students. Instructors can use this feedback to make changes to a course, both in real time and in future interactions to enhance the student learning experience. Since student evaluations of teaching impact tenure and promotion decisions, tools such as midsemester evaluations that can positively influence SETs should be used. However, these evaluations are meaningful only if implemented through a robust tool such as the CIQ that can facilitate concrete action based on student feedback. Future research should consider using the CIQ at two points in the course: in the middle and at the end. The first deployment of the CIQ can help the instructor identify issues with the course. Then, using CIQ again at the end of the course can help highlight whether the instructor's changes had any impact. The advantage of using the CIQ is that it highlights factors that specifically affect a particular group of students.[i]

References

Anderson, T. (2011). *The theory and practice of online learning* (5th ed.). Athabasca University Press. https://auspace.athabascau.ca/bitstream/handle/2149/411/?sequence=1

Baldwin, S. J., & Trespalacios, J. (2017). Evaluation instruments and good practices in online education. *Online Learning*, *21*(2). https://olj.onlinelearningconsortium.org/index.php/olj/article/view/913

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

Blackboard. (2017, April 10). Blackboard exemplary course program rubric. https://www.blackboard.com/resources/are-your-courses-exemplary

Boring, A., Ottoboni, K., & Stark, P. (2016, January 7). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *ScienceOpen Research*, 1–11. https://www.scienceopen.com/document/read?vid=818d8ec0-5908-47d8-86b4-5dc38f04b23e

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brookfield, S. (1998). Critically reflective practice. *Journal of Continuing Education in the Health Professions*, *18*(4), 197–205. https://doi.org/10.1002/chp.1340180402

California State University. (2019). CSU QLT course review instrument. https://docs.google.com/document/d/1ilqtDHjYfuJfjq1f8lG_bGhH9Xskcad-2a8aaPmnHG8/edit

Cohen, P. A. (1980). Effectiveness of student-rating feedback for improving college instruction: A meta-analysis of findings. *Research in higher education, 13*(4), 321-341.

Crews, T. B., Wilkinson, K., & Neill, J. K. (2015). Principles for good practice in undergraduate education: Effective online course design to assist students' success. *Journal of Online Learning and Teaching*, *11*(1), 87–103. https://uscdmc.sc.edu/about/offices_and_divisions/cte/instructional_design/docs/principles_good_practice_undergraduate_education_crews.pdf

Dancer, D., & Kamvounias, P. (2005). Student involvement in assessment: A project designed to assess class participation fairly and reliably. *Assessment and Evaluation in Higher Education 30*(4), 445–454. https://doi.org/10.1080/02602930500099235

Erikson, M., Erikson, M. G., & Punzi, E. (2016). Student responses to a reflexive course evaluation. *Reflective Practice*, *17*(6), 663–675. https://doi.org/10.1080/14623943.2016.1206877

Finelli, C. J., Ott, M., Gottfried, A. C., Hershock, C., O'Neal, C., & Kaplan, M. (2008). Utilizing instructional consultations to enhance the teaching performance of engineering faculty. *Journal of Engineering Education, 97*(4), 397–411. https://doi.org/10.1002/j.2168-9830.2008.tb00989.x

Fisher, R., & Miller, D. (2008). Responding to student expectations: A partnership approach to course evaluation. *Assessment & Evaluation in Higher Education, 33*(2), 191–202. https://doi.org/10.1080/02602930701292514

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*, 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Gehringer, E. (2010, June 20–23). *Daily course evaluation with Google forms* [Paper presentation]. 2010 American Society for Engineering Education Annual Conference & Exposition, Louisville, KY, United States. https://doi.org/10.18260/1-2--16350

Glowacki-Dudka, M., & Barnett, N. (2007). Connecting critical reflection and group development in online adult education classrooms. *International Journal of Teaching and Learning in Higher Education*, *19*(1), 43–52. https://eric.ed.gov/?id=EJ901286

Harasim, L. (2017). *Learning theory and online technologies*. Taylor & Francis. https://doi.org/10.4324/9781315716831

Hessler, H. B., & Taggart, A. R. (2011). What's stalling learning? Using a formative assessment tool to address critical incidents in class. *International Journal for the Scholarship of Teaching and Learning*, *5*(1), Article 9. https://doi.org/10.20429/ijsotl.2011.050109

Hurney, C., Harris, N., Bates Prins, S., & Kruck, S. E. (2014). The impact of a learner-centered, mid-semester course evaluation on students. *The Journal of Faculty Development*, *28*(3), 55–62. https://cetl.uni.edu/sites/default/files/impact_of_learner-centered_by_hurney_1.pdf

Jacobs, M. A. (2015). By their pupils they'll be taught: Using Critical Incident Questionnaire as feedback. *Journal of Invitational Theory and Practice*, *21*, 9–22. https://eric.ed.gov/?id=EJ1163006

Jaggars, S. S., & Xu, D. (2016). How do online course design features influence student performance? *Computers & Education*, *95*, 270–284. https://doi.org/10.1016/j.compedu.2016.01.014

Keefer, J. M. (2009). The Critical Incident Questionnaire (CIQ): From research to practice and back again. In R. L. Lawrence (Ed.), *Proceedings of the 50th Annual Adult Education Research Conference* (pp. 177–182). https://digitalcommons.nl.edu/cgi/viewcontent.cgi?referer=https://scholar.google.com/&httpsredir=1&article=1000&context=ace_aerc#page=191

Kyei-Blankson, L., Ntuli, E., & Donnelly, H. (2019). Establishing the importance of interaction and presence to student learning in online environments. *Journal of Interactive Learning Research*, *30*(4), 539–560. https://www.learntechlib.org/p/161956/

Linstrum, K. S., Ballard, G., & Shelby, T. (2012). Formative evaluation: Using the Critical Incident Questionnaire in a graduate counseling course on cultural diversity. *Journal of Intercultural Disciplines*, *10*, 94–102. https://www.proquest.com/docview/1033777245

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., 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

McKeachie, W. J., Lin, Y.-G., & Mann, W. (1971). Student ratings of teacher effectiveness: Validity studies. *American Educational Research Journal*, *8*(3), 435–445. https://doi.org/10.3102%2F00028312008003435

Mitchell, K. M., & Martin, J. (2018). Gender bias in student evaluations. *PS: Political Science & Politics*, *51*(3), 648–652. https://doi.org/10.1017/S104909651800001X

Moore, M. (1997). Theory of transactional distance. In D. Keegan(Ed.) Theoretical principles of distance education (pp.22-38) Routledge. http://www.c3l.uni-oldenburg.de/cde/found/moore93.pdf

Moore, M. G. (2013). *Handbook of distance education*. Routledge.

Phelan, L. (2012). Interrogating students' perceptions of their online learning experiences with Brookfield's Critical Incident Questionnaire. *Distance Education*, *33*(1), 31–44. https://doi.org/10.1080/01587919.2012.667958

Picciano, A. G. (2017). Theories and frameworks for online education: Seeking an integrated model. *Online Learning*, *21*(3). https://doi.org/10.24059/olj.v21i3.1225

Quality Matters. (2018). Specific review standards from the QM higher education rubric, sixth edition. https://www.qualitymatters.org/qa-resources/rubric-standards/higher-ed-rubric

Rao, V., & Woolcock, M. (2003). Integrating qualitative and quantitative approaches in program evaluation. In F. Bourguignon, & L. A. Pereira da Silva, (Eds.), *The impact of economic policies on poverty and income distribution: Evaluation techniques and tools* (pp. 165–190). World Bank and Oxford University Press.

Reid, L. D. (2010). The role of perceived race and gender in the evaluation of college teaching on RateMyProfessors.Com. *Journal of Diversity in higher Education*, *3*(3), 137–152. https://doi.org/10.1037/a0019865

Rodin, M., & Rodin, B. (1973). Student evaluations of teachers. *The Journal of Economic Education*, *5*(1), 5–9. https://www.jstor.org/stable/1734252

Samuel, A. (2020). Zones of agency: Understanding online faculty experiences of presence. *The International Review of Research in Open and Distributed Learning, 21*(4), 79–95. https://doi.org/10.19173/irrodl.v21i4.4905

Stark, P., & Freishtat, R. (2014, September 29). An evaluation of course evaluations. *ScienceOpen Research*, 1–7. https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AOFRQA.v1

State University of New York. (2018). The SUNY online course quality review rubric: OSCQR. https://oscqr.suny.edu/

Steyn, C., Davies, C., & Sambo, A. (2019). Eliciting student feedback for course development: The application of a qualitative course evaluation tool among business research students. *Assessment & Evaluation in Higher Education*, *44*(1), 11–24. https://doi.org/10.1080/02602938.2018.1466266

Subramanian, K., & Budhrani, K. (2020, February). Influence of course design on student engagement and motivation in an online course. In *SIGCSE '20: Proceedings of the 51st ACM Technical Symposium on Computer Science Education* (pp. 303–308). Association for Computing Machinery. https://dl.acm.org/doi/abs/10.1145/3328778.3366828

Swan, K. (2002). Building learning communities in online courses: The importance of interaction. *Education, Communication & Information*, *2*(1), 23–49. https://doi.org/10.1080/1463631022000005016

Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, *54*, 22–42. https://doi.org/10.1016/j.stueduc.2016.08.007

Veeck, A., O'Reilly, K., MacMillan, A., & Yu, H. (2016). The use of collaborative midterm student evaluations to provide actionable results. *Journal of Marketing Education*, *38*(3), 157–169. http://doi.org/10.1177/0273475315619652

Wolbring, T., & Treischl, E. (2016). Selection bias in students' evaluation of teaching. *Research in Higher Education, 57*(1), 51-71. https://link.springer.com/content/pdf/10.1007/s11162-015-9378-7.pdf

[i] The opinions and assertions expressed herein are those of the author(s) and do not necessarily reflect the official policy or position of the Uniformed Services University or the Department of Defense.

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

Fine-Tuned BERT Model for Large Scale and Cognitive Classification of MOOCs

Hanane Sebbaq and Nour-eddine El Faddouli
RIME Team, MASI Laboratory, E3S Research Center, Mohammadia School of Engineers (EMI), Mohammed V University, Rabat, Morocco

Abstract

The quality assurance of MOOCs focuses on improving their pedagogical quality. However, the tools that allow reflection on and assistance regarding the pedagogical aspects of MOOCs are limited. The pedagogical classification of MOOCs is a difficult task, given the variability of MOOCs' content, structure, and designs. Pedagogical researchers have adopted several approaches to examine these variations and identify the pedagogical models of MOOCs, but these approaches are manual and operate on a small scale. Furthermore, MOOCs do not contain any metadata on their pedagogical aspects. Our objective in this research work was the automatic and large-scale classification of MOOCs based on their learning objectives and Bloom's taxonomy. However, the main challenge of our work was the lack of annotated data. We created a dataset of 2,394 learning objectives. Due to the limited size of our dataset, we adopted transfer learning via bidirectional encoder representations from Transformers (BERT). The contributions of our approach are twofold. First, we automated the pedagogical annotation of MOOCs on a large scale and based on the cognitive levels of Bloom's taxonomy. Second, we fine-tuned BERT via different architectures. In addition to applying a simple softmax classifier, we chose prevalent neural networks long short-term memory (LSTM) and Bi-directional long short-term memory (Bi-LSTM). The results of our experiments showed, on the one hand, that choosing a more complex classifier does not boost the performance of classification. On the other hand, using a model based on dense layers upon BERT in combination with dropout and the rectified linear unit (ReLU) activation function enabled us to reach the highest accuracy value.

Keywords: cognitive MOOC classification, BERT, LSTM, transfer learning

Introduction

At the end of 2019, the spread of COVID-19 has caused a worldwide change in teaching from face-to-face to virtual or semi-virtual models. To ensure the continuity and efficacy of the learning process, many universities have turned to e-learning and especially MOOCs, and many professors use MOOCs to provide their academic courses. However, there are challenges to overcome, such as course design and quality. Quality, as it relates to the pedagogical framework of e-learning systems is the cornerstone for learning success and effectiveness (Conole, 2016) and it must be subject to constant monitoring and improvement. The quality assurance of e-learning systems can be guaranteed by applying quality instructional design (Kopp & Lackner, 2014) and by analyzing an important aspect of the teaching and learning process, namely the definition of learning objectives (LOs) associated with modules and programs. LOs are central to teaching and learning in many higher education institutions. However, teachers have limited tools to help them reflect on the LOs in the courses they create (Swart & Daneti, 2019). Our research aimed to assist teachers by recommending appropriate content based on their research (Sebbaq & al, 2020) and, importantly, on the cognitive level of their LOs. In our previous work (Sebbaq & Faddouli, 2021), we proposed a pedagogically enriched massive online open course (MOOC) ontology, that served as a standard to unify the representation of MOOCS and facilitate interoperability between MOOCs platforms. We enriched this ontology with metadata about learning objectives classified according to Bloom's taxonomy. This ontology served as a basis for the design and implementation of a linked data repository. We automatically extracted semantically rich descriptive metadata from different MOOC providers and integrate this metadata into a repository accessible through a simple protocol and RDF query language (SPARQL) endpoint. This repository served as the basis for our recommendation framework.

To concretize the epistemological position of the pedagogical dimension of MOOCs, we mapped MOOC learning objectives and Bloom's taxonomy levels. However, one of the limitations of our previous research work has been that the process of MOOC classification remained manual, which was tedious given its large scale. Therefore, the automation of classifying MOOCs according to their cognitive level remained an open research question. In this work, we proposed an approach for the cognitive classification of MOOCs. To the best of our knowledge and according to our literature review, there is no relevant study on automatic and large-scale pedagogical classification of MOOCs.

To study the pedagogies most suited to large-scale learning and teaching, and to highlight the special characteristics and properties of these pedagogies, several studies have compared existing pedagogies and case studies on MOOCs. Those studies analyzed the pedagogies used and proposed mechanisms and guidelines to improve the quality of MOOCs' pedagogical design. Analysis and classification of MOOCS was a difficult task, given the variability of MOOC structures, contents, designs, platforms, providers, and learner profiles. According to our literature review, the pedagogical classification of MOOCs has often been manual, as well as restricted to a limited number of MOOCs whose metadata are extracted manually and on a reduced scale. The objective of our work was to propose an automatic and large-scale pedagogical classification system for MOOCs according to their learning objectives. As we relied on the learning objectives for classification, we adopted the six cognitive levels of Bloom's taxonomy. The main constraint related to our study was the absence of annotated data—there is no dataset of annotated learning objectives organized according to the cognitive levels of Bloom's taxonomy. To overcome this problem, we resorted to building our dataset.

We managed to build a dataset of 2,394 LOs but this size remained limited. The transfer learning technique has demonstrated its performance on small datasets. We proposed a model based on bidirectional encoder representations from Transformers (BERT) for the cognitive classification of MOOCs using various fine-tuning strategies, and we examined the effect of different classifiers upon layers of BERT. Our experiment results showed, on the one hand, that using the pre-trained BERT model and fine-tuning it by adding dense layers outperformed the use of more complex classifiers like long short-term memory (LSTM) or (Bi-LSTM). On the other hand, using dense layers upon BERT in combination with dropout and the rectified linear unit (ReLU) activation function helped avoid overfitting.

The rest of this paper is organized as follows: Section 2 is a literature review. Section 3 describes the methodologies. Section 4 shows the experimental study. Section 5 answers the research questions and Section 6 is a conclusion.

Literature Review

Pedagogical Classification of MOOCs

To improve the quality of the instructional design of MOOCs, several studies have carried out comparisons of pedagogies that are most suited to large-scale learning and teaching, and that highlight the special characteristics and properties of these pedagogies. However, analysis and classification of MOOCs have been a difficult task given the variability of MOOC content, structure, designs, and providers. Educational researchers have adopted several approaches to understand these variations and identify the pedagogical models that exist in the pedagogical design of MOOCs. Kopp and Lackner (2014) studied MOOC models and designs. They structured these elements into a comprehensive checklist in the form of a framework to assist teachers in the design and development of a MOOC. However, this framework was descriptive and did not specify the characteristics associated with either the MOOC dimensions or assessment. Yousef et al. (2014) conducted research that classified MOOC quality criteria in two dimensions and six categories, which were manifested via 74 criteria. However, this study did not go further to evaluate MOOCs based on these criteria.

After a review of classification and description systems of existing MOOCs, Major and Blackmon (2016) proposed a descriptive framework with 11 dimensions including the educational dimension. Nevertheless, they did not go as far as pedagogical assessment. Similarly, to describe MOOCs, Rosselle et al. (2014) mapped eight dimensions to various characteristics of MOOCs. However, no assessment system was proposed. This mapping was an extension of that proposed by Pardos and Schneider (2013) who provided a conceptual mapping of MOOC designs. They categorized five main dimensions, which included four elements of the learning environment that could potentially affect design—instruction, content, assessment, and community.

To provide teachers with the guidance and the assistance they need to make better design decisions, Conole (2014) offered the 7Cs of learning design framework. This framework can be used for both designing and evaluating MOOCs. Moreover, Conole (2016) has also offered a 12-dimensional framework as well as a rating scale (low, medium, or high). These dimensions covered structural, philosophical, and pedagogical aspects, though the organizational system can be confusing. In addition, the assessment of some dimensions was unclear. Margaryan et al. (2015) proposed the course scan assessment system, a 37-item checklist based on existing instruments for quality instructional design. Margaryan et al. evaluated a sample of 76 MOOCs using the three dimensions: (a) course details (7 elements); (b) objectives and organization (6 elements); and (c) pedagogical principles (24 elements). The MOOCs evaluated were a random sample of those available at the end of 2013 on various platforms.

Based on an approach focused on pedagogy Swan et al. (2014) offered the MOOC assessing MOOC pedagogy (AMP) tool for evaluating pedagogy. The AMP generates a specific MOOC profile based on 10 pedagogical dimensions: (a) epistemology, (b) role of the teacher, (c) orientation of activities, (d) structure, (e) approach to content, (f) feedback, (g) cooperative learning, (h) adaptation to individual differences, (i) activities and evaluation, and (j) user's roles. The rating scale ranged from 0 to 5. Quintana and Tan (2019) introduced an extended version of the AMP tool with modified terminology and more sophisticated indicators. After evaluating 20 MOOCs (from the same platform and institution, but different fields), they showed how machine learning with the k-nearest neighbor (k-NN) algorithm helped identify pedagogically similar MOOCs. Xing (2019), using machine learning for the classification of MOOCs, analyzed 205 MOOCs to identify clusters of MOOCs using the k-means algorithm. Their goal was to study the impact of design features on learner engagement. Davis et al. (2018) used hierarchical clustering to group MOOCs according to their structures. They manually collected data from 177 MOOCs and looked only at the MOOCs' structures. An automatic notation was made by calculating the similarities via the two approaches (clustering transition probability and trajectory mining).

Table 1 summarizes the works reviewed here and classifies them according to whether their objective was description or evaluation. Information on the number of MOOCs analyzed is also provided to assess the large-scale character of the classification, while the data gathering column indicates whether the study gathered data automatically or manually. The third point of comparison shows whether the work integrated a MOOC assessment tool and whether it was automatic or manual. The sixth aspect of comparison deals with the use of machine learning for automating the classification, and the last column addresses the use of a theoretical foundation.

Table 1

Comparing Studies That Classify MOOCs

Research paper	Research objective	Number of MOOCs analyzed	Data gathering method	Assessment tool	Assessment method	Machine learning	Theoretical foundation
Conole (2014)	Description	-	-	Yes (low, medium, high)	Manual	-	Good learning
Conole (2016)	Description Evaluation (7Cs of learning design)	-	-	Yes	Manual	-	Good learning principles

Major and Blackmon (2016)	Description	-	-	-	-	-	-
Pardos and Schneider (2013)	Description Evaluation		-	Yes	Manual	-	Online learning
Rosselle et al. (2014)	Description	-	-	-	-	-	-
Kopp and Lackner (2014)	Description	-	-	-	-	-	ADDIE, Molenda (2003)
Yousef et al. (2014)	Description	-	-	-	-	-	-
Swan et al. (2014)	Pedagogical evaluation	17	Manual	AMP tool	Manual	-	-
Margaryan et al. (2015)	Evaluation	76	Manual	Course scan	Manual	-	Merrill
Xing (2019)	Evaluation	205	Manual	Yes	Manual	K-means	Web-based online instruction
Davis et al. (2018)	Evaluation	177	Manual	Calculate similarities	Automatic	Hierarchical clustering	-
Quintana and Tan (2019)	Evaluation (AMP tool extension)	20	Manual	Expanded AMP instrument	Manual	k-NN	-

Pedagogical Classification of E-Learning Content Based on Bloom's Taxonomy

Classification based on Bloom's taxonomy makes use of Bloom's taxonomy action verbs (BTAV). Swart and Daneti (2019) and Nevid and McClelland (2013) used BTAV to manually classify LOs. This time-consuming task required the participation of educational specialists in order to assure accurate outcomes. On the other hand, some verbs in the BTAV are unclear, since determining the necessary cognitive level is challenging. Verbs such as choose, describe, design, explain, show, and use can be found at different levels of cognition.

The use of Bloom's taxonomy to classify e-learning content (e.g., questions, forum texts) has received considerable attention in recent years. For automatic classification, researchers have used a variety of methodologies, ranging from rule-based to traditional machine learning to deep learning approaches. In the 1980s, the rule-based approach was the most popular. Omar (2012) and Haris and Omar (2012) demonstrated the effectiveness of the rule-based approach, although it was not without flaws. Among these disadvantages was the requirement for professionals to manually construct many rules to cover all sorts and domains of inquiries in order to improve the accuracy of the output.

When it comes to e-learning content, researchers have been mainly interested in evaluation questions. Abduljabbar and Omar (2015) combined three classifiers—support vector machine (SVM), nearest neighbor (NB), and k-nearest neighbor (k-NN)—and found that using an only bag of words (BOW) feature extraction improved accuracy. In 2020, Mohammed and Omar (2020) used a combination of Term Frequency–Inverse Document Frequency (TF-IDF), Part Of Speech (POS), and word2vec for feature extraction and tested the classifiers (i.e., k-NN, logistic regression (LR), and SVM). This combination, according to their research, increased F1-measurement performance. Osman and Yahya (2016) combined multiple feature extraction (BOW, POS, and n-grams) techniques with different machine learning algorithms (i.e., NB, SVM, LR, and decision trees) in order to test and compare them. Their research revealed that the feature extraction technique used had an impact on the machine learning classifier's performance.

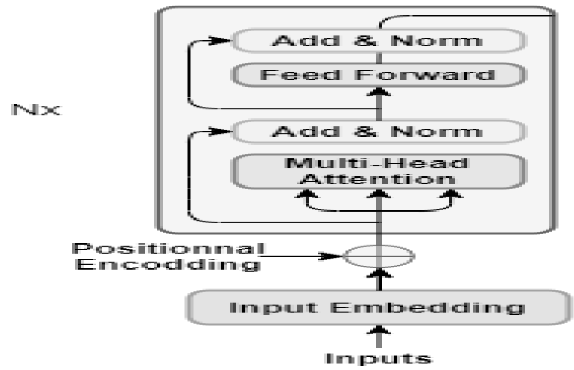
Some studies have used a deep learning-based approach to classify e-learning content according to the six cognitive levels of Bloom's Taxonomy. Ting Fei et al. (2003) examined the application of automated question classification tests in e-learning systems. They presented a text classification model that used a back-propagation learning approach to train a text classifier using an artificial neural network. Their technology outperformed the competition by about 78% in terms of F1 value. Yusof and Hui (2010) used an artificial neural network (ANN) strategy that employed numerous feature reduction strategies to develop a model that categorized question items. Das et al. (2020) proposed two strategies for automatically estimating the cognitive learning challenges of given questions. Their first method used latent Dirichlet allocation (LDA) as a deep learning strategy. For multi-class text classification, the second methodology employed BERT. According to their findings, BERT had an accuracy of 89%, which was higher than LDA's 83%.

BERT for Text Classification

The BERT model is based on two stages: pre-training and fine-tuning (Devlin et al., 2019). During pre-training, the model is trained on a large unlabeled corpus. The model is then fine-tuned, starting with the pre-trained parameters and refining all parameters with task-specific labeled data. BERT uses the transformer that is a new architecture presented in Vaswani et al. (2017). A simple transformer consists of an encoder that reads text input and a decoder that generates a task prediction. BERT requires only the encoder depicted in Figure 1 because its objective is to develop a model of the language representation.

Figure 1

The BERT Encoder



BERT is based on the attention mechanism (Vaswani et al., 2017) that was invented to allow a model to comprehend and remember the contextual relationships between features and text. The attention mechanism maps a set of queries to their corresponding sets of keys and values—vectors that contain information about related or neighboring entities from the text input. The procedure involves the dot product of the input query with the existing keys, and then a softmax to give a scaled dot product attention score, which is defined as follows (Vaswani et al., 2017):

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V$$

where Q is the query matrix, K is the key matrix, V is the value matrix, and d_k is the dimension of the Q and K matrices. This resulting score vector is then multiplied by each value and summed to give the final self-attention result for that particular query. Interpretation of this multi-head attention helps the model determine how much attention it should pay to each word in the text block (Vaswani et al., 2017). Multi-head attention is defined as follows:

$$MultiHead(Q, K, V) = Concat(head_1, \ldots, head_h)W^O$$

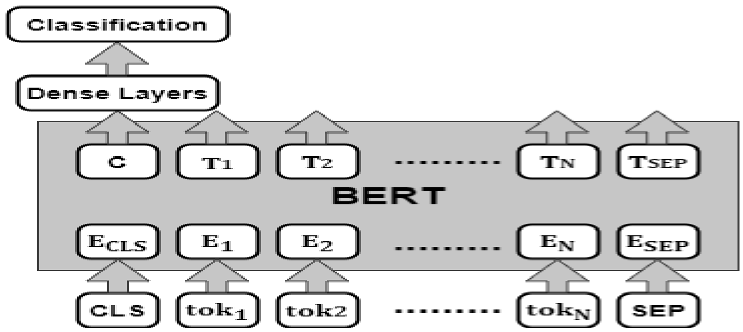
where $head_i$ = Attention (QW^Q_i, $KW^K{}_i$, $V W^V{}_i$) and the projections are parameter matrices $W_i^Q \in \mathbb{R}^{dmodel\times dk}$, $W_i^K \in \mathbb{R}^{dmodel\times dk}$, $W_i^V \in \mathbb{R}^{dmodel\times dv}$, and $W^O \in \mathbb{R}^{hdv\times dmodel}$ (Vaswani et al., 2017).

BERT represents a single sentence or a pair of sentences as a sequence of tokens with the following characteristics:

- The first token in the sequence is [CLS].
- When there is a pair of sentences in the sequence, they are separated by the token [SEP].
- For a given token, its input representation is constructed by summing the corresponding token, position, and segment embeddings (see Figure 2).

Figure 2

BERT Architecture



BERT is a leading model for a variety of Natural Language Processing (NLP) tasks, demonstrating its efficiency and potential. In this study, we explored fine-tuning methods for applying BERT to a cognitive classification task.

Methodology

Theoretical Foundation of Our Proposed Approach

According to our comparative study summarized in Table 1, several theoretical frameworks have been adopted for the evaluation and classification of MOOCs. Conole (2014) associated good learning with quality learning. It was critical, in his opinion, to meet the characteristics of good learning in order to accomplish effective learning. Conole (2016) based the12-dimensional assessment framework, as well as the 7Cs for learning design framework on this principle. On the other hand, according to Merrill (2012), the first principles of instruction he proposed constituted the foundation of all present pedagogical models and theories. Merill suggested five guidelines for the development of learning activities. Margaryan et al. (2015) built on Merill's first principles of instruction and added five more principles related to learning resources. Conole (2014, 2016) and Magrayan et al. (2015) used these two theoretical frameworks to drive their research into the development of evaluation frameworks focused on open-ended questions and necessitating the assistance of an expert. Xing (2019) used a Web-based online instruction approach to drive their evaluation of MOOCs. This approach was more generic and incorporated three global design dimensions: information, instruction, and learning. All the frameworks adopted in these studies focused on open-ended questions and called for expert assistance, while our research objective was to automate the evaluation of LOs.

The ultimate aim of our research was to evaluate and classify the pedagogical dimension of MOOCs based on their learning objectives. The theoretical foundation of Bloom's taxonomy was most appropriate for our context since it covered the different levels of cognitive learning and allowed for classifying learning objectives according to six hierarchical levels. The initial purpose of Bloom's taxonomy was to assist teachers in developing rubrics and measuring the achievement of their learning goals by providing guidelines. We used a modification of Bloom's taxonomy adapted from Krathwohl (2002) who proposed a revision of the original taxonomy. It defined a two-dimensional framework consisting of knowledge and cognitive processes. The first dimension took the subcategories of the first level of the original taxonomy; the second dimension renamed the six levels as verbs—remembering, understanding, applying, analyzing, evaluating, and creating. Our research considered the dimension of cognitive learning processes.

Most studies that have classified e-learning content based on Bloom's taxonomy focused on classifying assessment questions. No research has been done on the automatic classification of LOs. Machine learning has been used most often, followed by the rule-based approach. The majority of research publications have focused on merging multiple feature extraction and feature selection methods to improve the performance of machine learning classifiers. The deep learning approach has been used less often; only the ANN architecture has been tested in this context. BERT was used in a single study for cognitive classification purposes (Das et al., 2020). There has been some research comparing BERT and other machine learning or deep learning models (González-Carvajal & Garrido-Merchán, 2020). Our research, on the other hand, is the first to investigate the cognitive classification of LOs.

A BERT-Based Cognitive Approach for Classifying MOOCs

From our review of the literature, we have deduced that there is no large-scale, automatic classification system for MOOCs based on their pedagogical approaches. As we summarized in Table 1, the existing research has addressed one of the following:

- Frameworks developed for quality assurance that are generalist and lack educational considerations and means to operationalize the review of MOOCs' pedagogical quality.

- Case studies that detailed the design of individual MOOCs to highlight best practices and pedagogical models. However, these studies concerned only a small number of MOOCs and were not based on a well-defined evaluation framework.

- Descriptive frameworks that were intended for designing MOOCs from scratch, which was not our objective.

- Evaluation frameworks that dealt with several dimensions including the pedagogical.

However, not all frameworks of the latter type were focused on pedagogy, and they all suffered from the lack of an automatic and large-scale system for classifying MOOCs according to their pedagogical models. In addition, their dimensions were broad and had to be operationalized via qualitative and quantitative indicators as well as concrete characteristics. This research analyzed course design and pedagogy to understand variations in the two, but much of this analysis relied on a human categorization process based on broad interpretations of the learning designs. In addition, Assessment tools were based on open-ended questions that required the intervention of an expert. Since it is difficult to automate the assessment, this has remained a manual task and automated tools are not yet widely adopted by researchers in the MOOC community. The only study whose assessment was automatic, Davis et al. (2018) was restricted to comparing MOOC structures, similar to Pardos and Schneider (2013). Our objective was to classify MOOCs on a large scale; the number of MOOCs analyzed in the research cited above was not sufficient to deduce the different pedagogies in MOOCs. Both Xing (2019) and Davis et al. (2018) used machine learning for a large analysis of MOOCs. However, the number of MOOCs they examined remained limited, and their data collection methods were manual. Swan et al. (2014) used machine learning for the analysis of about 20 MOOCs. Nevertheless, the result of their clustering cannot be generalized given the limited number of MOOCs they analyzed.

The main challenge of our study was the lack of annotated data. Despite thorough research, we found no annotated learning goal datasets, so we created our learning objectives dataset. Evens so, the size of the dataset remained limited. For its part, BERT is the state-of-the-art technique in NLP (Devlin et al., 2018) and it has demonstrated its performance on small datasets. The contributions of this study are:

- The automatic classification of MOOCs according to their pedagogical approaches based on the cognitive levels of Bloom's taxonomy. The first phase of our automated approach has been implemented in our previous work (Sebbaq & Faddouli, 2021).
- The large scale of our approach was based on the repository of MOOCs already built in our previous work (Sebbaq & Faddouli, 2021).
- The use of transfer learning to resolve the issue of the lack of annotated data.
- Fine-tuning BERT via different strategies: we investigated the impact of choosing different classifiers, from a simple softmax classifier to a more complex classifier like dense layers, LSTM, and Bi-LSTM.

Fine-Tuning Strategies

BERT fine-tuning involves training a classifier with different layers on top of the pre-trained BERT transformer to minimize task-specific parameters. Fine-tuning for a specific task can be done using several approaches, either by fine-tuning the architecture or by fine-tuning different hyper-parameters such as the learning rate or the choice of the best optimization algorithm. Our objective in this research work was the cognitive classification of MOOCs according to their learning objectives. For this type of classification problem, we simply adopted the basic architecture of BERT and then added an output layer for the classification. This layer took as input the final hidden state of the first token [CLS]. We considered this exit to be the ultimate representation of the entire entry sequence. The output layer can be either a simple classifier like softmax or a more complicated network like the bidirectional Bi-LSTM. In our approach, we proposed six different architectures for fine-tuning BERT.

BERT-Based Fine-Tuning

Figure 3 represents the first architecture. In this basic architecture, we mainly relied on the BERT base; we used the output of the token [CLS] provided by BERT only. The output [CLS] was a vector of size 768; we gave it as input to a network that was fully connected with no hidden layer. Since our classification problem was multi-class, the output layer was based on a softmax activation layer. The softmax function formula is:

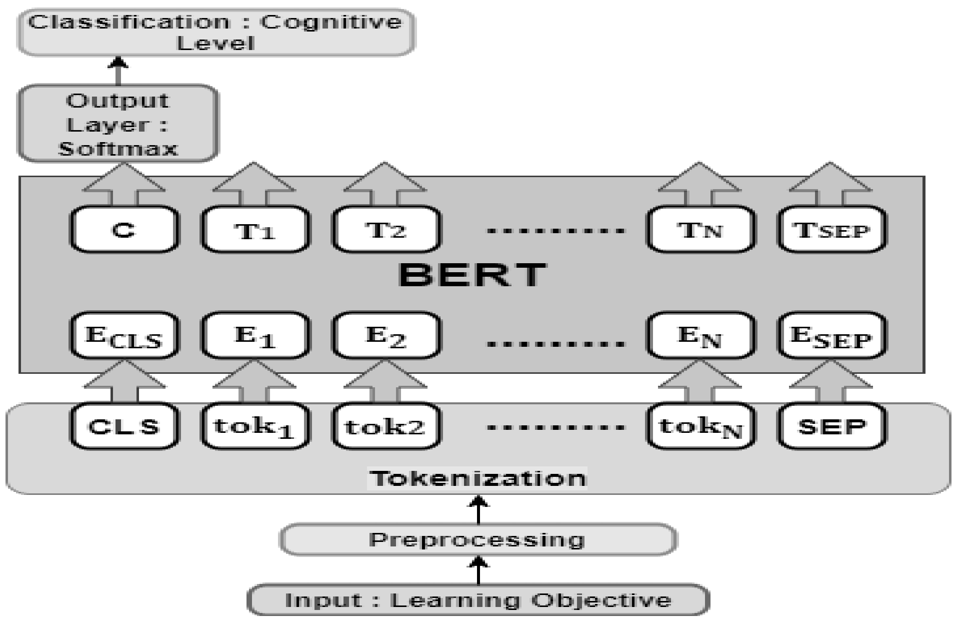
$$\sigma(\vec{z}) = \frac{e^{\mathrm{zi}}}{\sum_{j=1}^{K} e^{zj}}$$

where $\vec{Z} = (z_1; : : : ; z_K)$, z_i values are the elements of the input vector to the softmax function, K is the number of classes in the multi-class classifier. The output node with the highest probability is then chosen as the predicted label for the input.

For preprocessing, we simply cleaned the text of non-alphabetic characters and converted it to lower case.

Figure 3

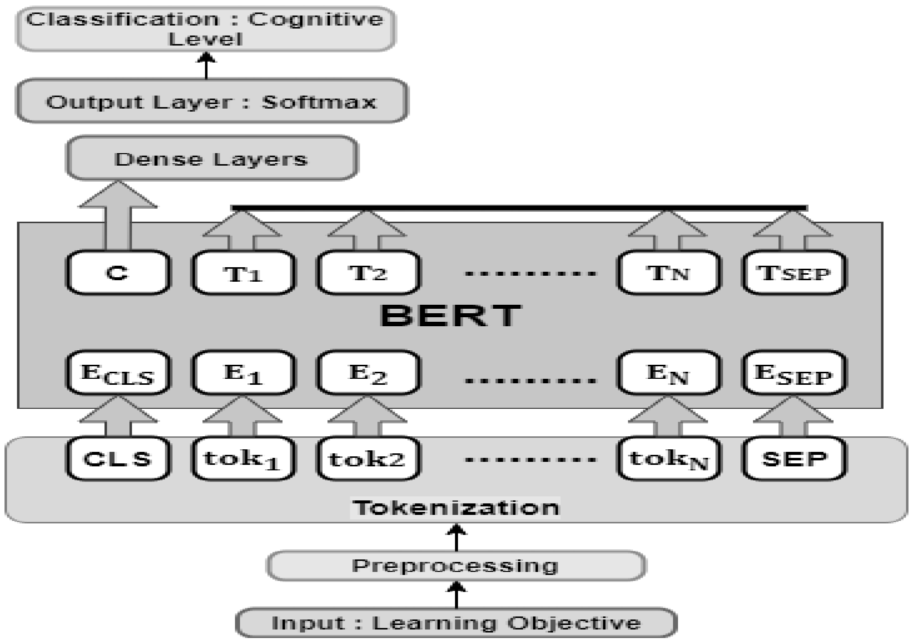
BERT-Based Fine-Tuning Architecture



In this architecture (see Figure 4), we added fully connected layers. The fully connected layer took the output of BERT's 12 layers and transformed it into the final output of six classes that represented the six cognitive levels of Bloom's taxonomy. This layer consists of three dense layers.

Figure 4

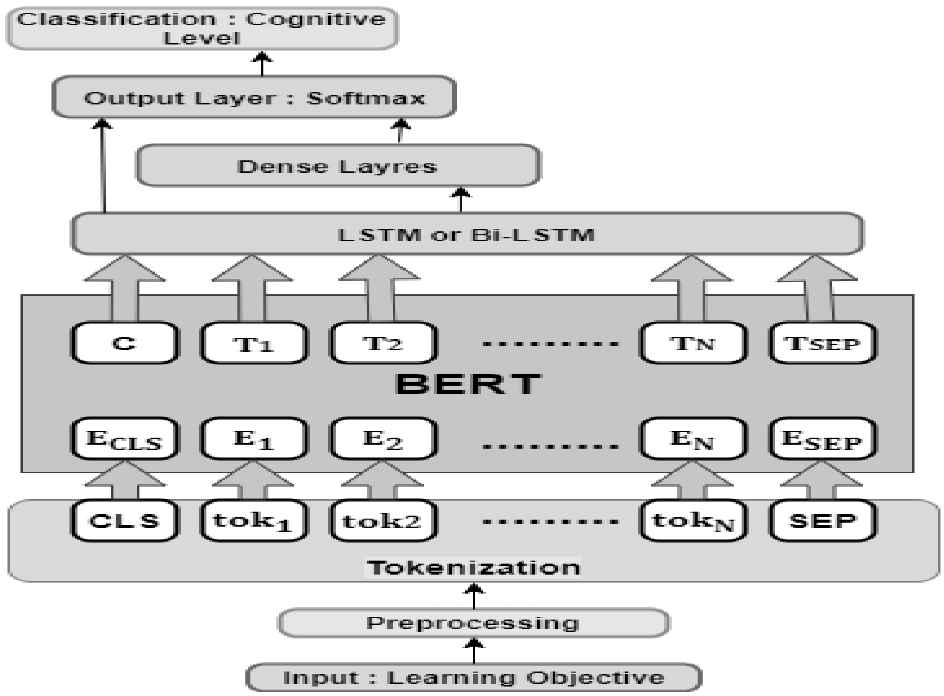
BERT With Fully Connected Layers Architecture



In previous architectures, the [CLS] output was the only one used as input for the classifier. In this architecture, we used all the outputs of the last transformer encoder as inputs to an LSTM or Bi-LSTM recurrent neural network as shown in Figure 5. After the input was processed, the network sent the final hidden state to the output layer, which was a fully connected network, to perform classification using the softmax activation function. We also experimented with architecture that was more complex by inserting dense layers between the deep network layer and the output layer.

Figure 5

BERT With Fully Connected Layers and Deep Network Layers Architecture



Experimental Study

In this section, we provide a representation and a detailed analysis of the dataset as well as a complete presentation of the results that we obtained from experimentation with the different models.

Dataset Description and Analysis

Given the challenge posed by the lack of annotated datasets of LOs according to the cognitive levels of Bloom's taxonomy, our solution was to create our dataset as presented in Table 2. We started by collecting LOs from the MOOCs providers, Coursera, and edX, and then manually annotated them based on Bloom's taxonomy action verbs list. However, some of the action verbs in BTAV overlap at several levels of the hierarchy (Krathwohl, 2002). This leads to ambiguity about the actual meaning of the required cognition (Stanny, 2016) and affects the effectiveness of the BTAV-based classification. Moreover, this method has the drawback of not being able to guarantee the accuracy of our annotations, as well as being time-consuming.

Table 2

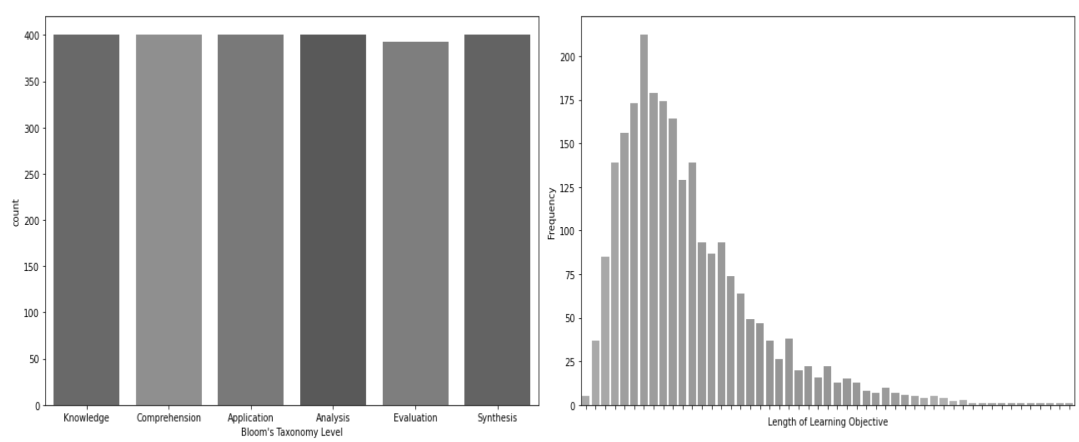
The Distribution of LOs in Our Dataset

Cognitive level	2 394	Example
Knowledge (remembering)	400	Describe the concept of modular programming and the uses of the function in computer programming
Comprehension (understanding)	400	At the end of this module, the learner will be able to classify clustering algorithms based on the data and cluster requirements
Application (applying)	400	Apply a design process to solve object-oriented design problems
Analysis (analyzing)	400	Analyze the appropriate quantization algorithm
Evaluation (evaluating)	394	Compare the semantic and syntactic ways encapsulation
Synthesis (creating)	400	Create a Docker container in which you will implement a Web application by using a flask in a Linux environment

The training dataset consisted of 2,394 training objectives. Figure 6 illustrates the number of words per input data point in the form of a histogram. According to the histogram, the average length of the training objectives was about 225. Regarding the class distribution of the data in the input dataset, analysis of Figure 6 suggests that the classes are balanced, with 400 learning objectives per cognitive level.

Figure 6

Distribution of Data by Level and Item Length



Evaluation Metrics

Several considerations, including class balance and expected outcomes, guided the selection of the best measures to evaluate the performance of a given classifier on a certain dataset. Given a dataset with an approximately balanced number of samples from all classes, we used the accuracy measure to evaluate the performance of our model and compare it with other models (Grandini et al., 2020). Accuracy is the sum of true positive (TP) and true negative (TN) items divided by the sum of all other possibilities, defined as follows:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN)$$

where TP = True Positives, TN = True Negatives, FN = False Negatives, and FP = False Positives.

Environment Setup

We used the Google Colab and Tensorflow environment as well as Keras Tensorflow to build the BERT models. Keras TensorFlow is an open-source mathematical software library used for machine learning applications. It has tools to run on graphic processing units, which can significantly reduce training and inference times on some models. Keras is a high-level API for TensorFlow. It has a modular and easily extensible architecture, and it allows users to create sequential models or a graph of modules that can be easily combined. The library contains many different types of neural layers, cost, and activation functions. We implemented different fine-tuning strategies of BERT on Tensorflow Hub (TFHub). TFHub provides a way to try, test, and reuse machine learning models.

Implementation Details

For our experiments, we used the basic pre-trained model bert_multi_cased_L-12_H-768_A-12/2, which had 12 layers, 768 hidden, 12 self-attention heads, and 110M parameters. We used the Adam optimizer, which is one of the most stable and widely used in the deep learning world (Kingma & Ba, 2014). Adam was used in combination with the warmup steps, which were low learning-rate updates that helped the model converge. After trying many different configurations, and after numerous unsuccessful attempts that ran out of memory, we arrived at a working configuration of hyper-parameters. The base learning rate was 3e-5, and the warm-up proportion was 0.1. We empirically set the maximum number of epochs to 15 with a batch size of 32 and saved the best model on the validation set for testing.

For the implementation of the models adopted, we use the Keras Layer function of Tensorflow Hub to build our BERT layer. Then we tokenized our text based on the variables of this layer. This allowed us to have the first input of our BERT model, which was input_word_ids. Then we built the two other inputs of BERT, which were the embeddings of position input_mask and segments segment_ids. We added a dropout of 0.1 after each layer.

Results Analysis

We conducted experiments to demonstrate the efficiency of our proposed approach in terms of performance. Our main task was to explore the performance of BERT on cognitive text classification and evaluate the impact of different fine-tuning strategies. We used six models: (a) standard BERT-based fine-tuning, (b) BERT with fully connected layers, (c) BERT with LSTM, (d) BERT with Bi-LSTM, (e) BERT with both LSTM and fully connected layers, and (f) BERT with both Bi-LSTM and fully connected layers.

In particular, we aimed to answer the following research questions via our experiments.

- (RQ1): How do different fine-tuning strategies have different impacts on the cognitive classification task?

- (RQ2): How effectively does our BERT-based cognitive approach produce better results than other baseline fine-tuning strategies?

(RQ1) Comparing Performance of Various Architectures With Different Classifiers

In order to answer (RQ1), we investigated the effects of various classifiers on BERT. To use a basic softmax classifier upon the last layer of BERT, we experimented with a cascade of dense layers with the activation function ReLU and the more complex classifiers, LSTM and Bi-LSTM. Table 3 presents the accuracies of the six models ranked from the basic to the more complicated. The results demonstrate that the use of a more complex classifier did not improve performance. Instead, it lowered accuracy on the five classification models, which is understandable given that BERT also has deep networks and advanced training techniques.

(RQ2) BERT With Three Dense Layers Performed Best

As a response to (RQ2), our proposed model performs better than other baseline fine-tuning strategies. Thanks to the addition of dense layers on top of BERT. A dense layer is a regular, deeply connected neural network layer of neurons. Each neuron in the previous layer receives feedback from all the neurons in the layer before it, making it densely connected. To prevent overfitting, we also used the dropout, a regularization technique where randomly selected neurons are ignored during training. At each upgrade of the training process, dropout randomly set the outgoing edges of hidden units to zero at random. We added a dropout of 0.1 after each dense layer. We used the activation function ReLU. The biggest benefit of using the ReLU mechanism over other activation functions was that it did not simultaneously stimulate any of the neurons.

Table 3

Accuracies of the Different Models

BERT Model	Accuracy
BERT-based fine-tuning	88.75%
BERT with three fully connected layers	92.5%
BERT with LSTM	91.25%
BERT with Bi-LSTM	90.83%
BERT with three fully connected layers + LSTM	91.46%
BERT with three fully connected layers + Bi-LSTM	92.08%

Conclusion

The main constraint of our study was the availability of annotated LOs datasets. We managed to build a dataset of 2,394 LOs but this size was still limited. On the other hand, each LO needed to be carefully annotated for the training data to be correct. This makes building a larger dataset a cumbersome and difficult task to handle.

In this study, our goal was to propose a model for the automatic classification of MOOCs according to their pedagogical approaches, and this on a large scale, based on the cognitive levels of Bloom's taxonomy. To this end, we opted for BERT, and then we experimented with different strategies to fine-tune it. In this sense, we investigated the impact of choosing different classifiers upon BERT, from a simple softmax classifier to a more complex classifier such as dense layers, LSTM, and Bi-LSTM. The results demonstrated that using a more complex classifier did not improve performance. Instead, it lowered accuracy on the five classification models, which is understandable given that BERT also has deep networks and advanced training techniques. We also demonstrated that the use of dense layers upon BERT in combination with dropout and the activation function ReLU allowed us to reach the highest accuracy value. Although BERT with dense layers performed well in our experiment, we have not yet explored other fine-tuning strategies. In a future study, we will tackle other techniques such as multitask learning to enhance the performance of our BERT model.

Overall, our proposed approach proved its ability to classify learning objectives in MOOCs. Since our approach was based on learning objectives for the pedagogical classification of MOOCs, potential applications to other learning objects in the context of distance learning are worth exploring in future research and practice.

References

Abduljabbar, D. A., & Omar, N. (2015). Exam questions classification based on Bloom's taxonomy cognitive level using classifiers combination. *Journal of Theoretical and Applied Information Technology, 78*(3), 447. http://www.jatit.org/volumes/Vol78No3/15Vol78No3.pdf

Conole, G. (2014, April). The 7Cs of learning design: A new approach to rethinking design practice. *Proceedings of the Ninth International Conference on Networked Learning* (pp. 502–509). Edinburgh, Scotland. https://www.lancaster.ac.uk/fss/organisations/netlc/past/nlc2014/abstracts/pdf/conole.pdf

Conole, G. (2016). MOOCs as disruptive technologies: Strategies for enhancing the learner experience and quality of MOOCs. *Revista de Educación a Distancia, 50*(2). http://dx.doi.org/10.6018/red/50/2

Das, S., Das Mandal, S. K., & Basu, A. (2020). Identification of cognitive learning complexity of assessment questions using multi-class text classification. *Contemporary Educational Technology, 12*(2), Article ep275. https://doi.org/10.30935/cedtech/8341

Davis, D., Seaton, D., Hauff, C., & Houben, G. J. (2018, June). Toward large-scale learning design: Categorizing course designs in service of supporting learning outcomes. *Proceedings of the Fifth Annual ACM Conference on Learning at Scale* (pp. 1–10). https://doi.org/10.1145/3231644.3231663

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding*. arXiv preprint arXiv:1810.04805. https://arxiv.org/abs/1810.04805

González-Carvajal, S., & Garrido-Merchán, E. C. (2020). *Comparing BERT against traditional machine learning text classification*. arXiv preprint arXiv:2005.13012. https://arxiv.org/abs/2005.13012

Grandini, M., Bagli, E., & Visani, G. (2020). *Metrics for multi-class classification: An overview*. arXiv preprint arXiv:2008.05756. https://arxiv.org/abs/2008.05756

Haris, S. S., & Omar, N. (2012, December). A rule-based approach in Bloom's taxonomy question classification through natural language processing. *Seventh International Conference on Computing and Convergence Technology* (ICCCT; pp. 410–414). Institute of Electrical and Electronics Engineers. https://ieeexplore.ieee.org/abstract/document/6530368

Kingma, D. P., & Ba, J. (2014). *Adam: A method for stochastic optimization*. arXiv preprint arXiv:1412.6980. https://arxiv.org/abs/1412.6980

Kopp, M., & Lackner, E. (2014). Do MOOCs need a special instructional design? *Proceedings of Sixth International Conference on Education and New Learning* (EDULEARN14; pp. 7138–7147). Barcelona, Spain. https://library.iated.org/view/KOPP2014DOM

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Major, C. H., & Blackmon, S. J. (2016). Massive open online courses: Variations on a new instructional form. *New Directions for Institutional Research, 2015*(167), 11–25. https://doi.org/10.1002/ir.20151

Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of massive open online courses (MOOCs). *Computers & Education, 80*, 77–83. https://doi.org/10.1016/j.compedu.2014.08.005

Merrill, M. D. (2012). *First principles of instruction*. John Wiley & Sons. https://digitalcommons.usu.edu/usufaculty_monographs/100/

Mohammed, M., & Omar, N. (2020). Question classification based on Bloom's taxonomy cognitive domain using modified TF-IDF and word2vec. *PLOS ONE 15*, e0230442. https://doi.org/10.1371/journal.pone.0230442

Molenda, M. (2003). In search of the elusive ADDIE model. *Performance improvement*, *42*(5), 34-37. http://www.damiantgordon.com/Courses/DT580/In-Search-of-Elusive-ADDIE.pdf

Nevid, J. S., & McClelland, N. (2013). Using action verbs as learning outcomes: Applying Bloom's taxonomy in measuring instructional objectives in introductory psychology. *Journal of Education and Training Studies, 1*(2), 19–24. https://doi.org/10.11114/jets.v1i2.94

Omar, N., Haris, S. S., Hassan, R., Arshad, H., Rahmat, M., Zainal, N. F. A., & Zulkifli, R. (2012). Automated analysis of exam questions according to Bloom's taxonomy. *Procedia: Social and Behavioral Sciences, 59*, 297–303. https://doi.org/10.1016/j.sbspro.2012.09.278

Osman, A., & Yahya, A. A. (2016). Classifications of exam questions using linguistically-motivated features: A case study based on Bloom's taxonomy. https://www.researchgate.net/publication/298286164_Classifications_of_Exam_Questions_Using_Linguistically-Motivated_Features_A_Case_Study_Based_on_Blooms_Taxonomy

Pardos, Z. A., & Schneider, E. (2013). *AIED 2013 workshops proceedings* (Vol. 1). http://people.cs.pitt.edu/~falakmasir/docs/AIED2013.pdf

Quintana, R. M., & Tan, Y. (2019). Characterizing MOOC pedagogies: Exploring tools and methods for learning designers and researchers. *Online Learning, 23*(4), 62–84. https://doi.org/10.24059/olj.v23i4.2084

Rosselle, M., Caron, P. A., & Heutte, J. (2014, February). A typology and dimensions of a description framework for MOOCs. In *Proceedings of European MOOCs Stakeholders Summit 2014,* (eMOOCs 2014; pp. 130–139). Lausanne, Switzerland. Proceedings document published by Open Education Europa (www. openeducationeuropa. eu). https://hal.archives-ouvertes.fr/hal-00957025/

Sebbaq, H., El Faddouli, N. E., & Bennani, S. (2020, September). Recommender system to support MOOCs teachers: Framework based on ontology and linked data. *Proceedings of the 13th International Conference on Intelligent Systems: Theories and Applications,*Article 18. https://doi.org/10.1145/3419604.3419619

Sebbaq, H., & Faddouli, N. E. E. (2021, January). MOOCs semantic interoperability: Towards unified and pedagogically enriched model for building a linked data repository. *International Conference on Digital Technologies and Applications*. Springer, 621-631.

Stanny, C. J. (2016). Reevaluating Bloom's taxonomy: What measurable verbs can and cannot say about student learning. *Education Sciences, 6*(4), 37. https://doi.org/10.3390/educsci6040037

Swan, K., Day, S., Bogle, L., & van Prooyen, T. (2014). AMP: A tool for characterizing the pedagogical approaches of MOOCs. *E-mentor, 2*(54), 75–85. https://doi.org/10.15219/em54.1098

Swart, A. J., & Daneti, M. (2019, April). Analyzing learning outcomes for electronic fundamentals using Bloom's taxonomy. *2019 IEEE Global Engineering Education Conference* (EDUCON; pp. 39–44). Institute of Electrical and Electronics Engineers. https://ieeexplore.ieee.org/document/8725137

Ting Fei, Wei Jyh Heng, Kim Chuan Toh, & Tian Qi. (2003). Question classification for e-learning by artificial neural network. *Proceedings of the Joint Fourth International Conference on Information, Communications and Signal Processing, and the Fourth Pacific Rim Conference on Multimedia* (pp. 1757–1761). Institute of Electrical and Electronics Engineers. Singapore. https://doi.org/10.1109/ICICS.2003.1292768

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *31st Conference on Neural Information Processing Systems* (NIPS 2017; pp. 5998–6008). Long Beach, USA. https://arxiv.org/abs/1706.03762

Xing, W. (2019). Exploring the influences of MOOC design features on student performance and persistence. *Distance Education, 40*(1), 98–113. https://doi.org/10.1080/01587919.2018.1553560

Yousef, A. M. F., Chatti, M. A., Schroeder, U., & Wosnitza, M. (2014, July). What drives a successful MOOC? An empirical examination of criteria to assure design quality of MOOCs. *14th International Conference on Advanced Learning Technologies* (pp. 44–48). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ICALT.2014.23

Yusof, N., Hui, C. J. (2010). Determination of Bloom's cognitive level of question items using artificial neural network. *10th International Conference on Intelligent Systems Design and Applications* (ISDA; pp. 866–870). Institute of Electrical and Electronics Engineers. Cairo, Egypt. https://doi.org/10.1109/ISDA.2010.5687152

Athabasca University

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

Designing Asynchronous Online Discussion Forum Interface and Interaction Based on the Community of Inquiry Framework

Lintang Matahari Hasani, Harry Budi Santoso, and Kasiyah Junus
Faculty of Computer Science, Universitas Indonesia

Abstract

The community of inquiry (CoI) framework describes a process for creating collaborative learning through three elements or presences: social, cognitive, and teaching. Despite its popularity among researchers and practitioners, use of the CoI model is limited to mapping instructional activities, which are yet to be developed into an interaction design for online collaborative learning intended to support the CoI presences. This study was aimed at developing the interaction design of an asynchronous online discussion forum employing a user-centered design method contextualized to the learning-centered design approach. Seven scenario and user interfaces were created to facilitate one introductory activity and four phases of inquiry. The design was evaluated through contextual interviews with ten students. The interviews revealed that the prototype encouraged and supported (a) introductory activity (social presence), (b) idea exploration (cognitive presence), (c) summarizing the discussion (cognitive presence), and (d) facilitating discussion (teaching presence). Future research could be aimed at improving the proposed design based on recommendations and developing a fully functional working system to be tested in real settings.

Keywords: community of inquiry, interaction design, e-learning, user-centered design

Introduction

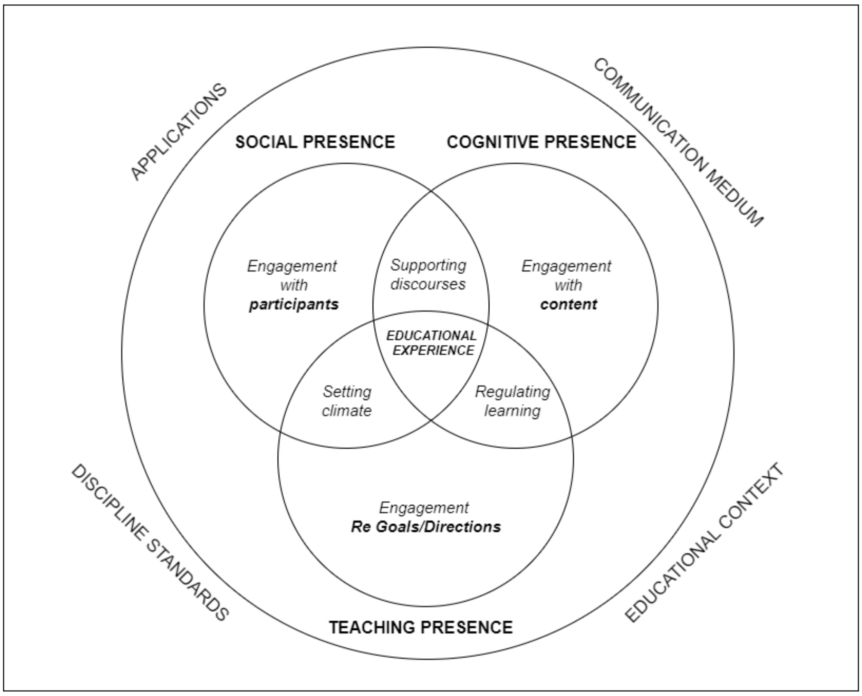
Online collaborative learning (OCL) has been widely discussed and recommended in the discourse of online learning (Laal, Laal, & Kermanshahi, 2012). The notion of collaboration in OCL refers to a group of learners who participate in *idea transactions* in which learners construct their personal knowledge (Garrison, 2016). This concept was derived from constructivist theories that describe learning as an active process in which a person constructs knowledge based on experience (Hendry, Frommer, & Walker, 1999), prior knowledge (Bransford, Brown, & Cocking, 2000), and interactions with other learners as well as environments (Garrison, 2016).

The discourse of OCL has brought forward various theories and frameworks. A systematic review of the trends in online educational research by Valverde-Berrocoso, Garrido-Arroyo, Burgos-Videla, and Morales-Cevallos (2020) found that the community of inquiry (CoI) framework developed by Garrison, Anderson, and Archer (2000) has become the most widely adopted framework in online educational research. Garrison et al. (2000) developed the CoI framework based on the concept of collaborative thinking. The community of inquiry refers to a learning environment in which a group of learners collaborates to construct understanding of a concept (Garrison, 2016).

The CoI framework consists of three interrelated constructs or presences: cognitive presence, teaching presence, and social presence. According to the CoI model, the three constructs make up a meaningful educational experience. Cognitive presence refers to a learner's construction of meaning through interacting with other learners. Teaching presence refers to the organization and management of learning activities to sustain meaningful interaction and knowledge construction. Social presence refers to a learner's expression of authentic individual characteristics, which encompasses social cues, openness, and group cohesion. The CoI model, including its three presences, is depicted in Figure 1.

Figure 1

The Community of Inquiry (CoI) Model



Note. From "About the Framework: An Introduction to the Community of Inquiry," by Centre for Distance Education at Athabasca University, 2021 (http://thecommunityofinquiry.org/coi). CC BY-SA 4.0.

The dynamics of the CoI presences are related to the need to foster and sustain certain presences during each specific phase of a collaborative learning activity. Garrison (2016) recommended that social presence be nurtured during the early phase of learning. Once social presence has created a conducive learning environment for effective collaboration (e.g., strong group cohesion, openness for communication, etc.), cognitive and teaching presence need to be cultivated to ensure the attainment of learning goals. Therefore, effective instructional strategies should be designed to meet a range of needs.

In the context of online learning in higher education institutions, the CoI model has become popular as a framework that provides guidance to deliver computer-assisted collaborative learning. The CoI framework is also viewed as suitable for higher education settings (Vaughan, Cleveland-Innes, & Garrison, 2013) based on the understanding that education at this level is fundamentally a process of inquiry in which learners act as discoverers and not as mere users or followers (Lipman, 2003).

In regard to CoI framework implementation in an online learning environment, an asynchronous online discussion forum (AODF) could serve to facilitate collaboration among learners. Both collaboration and critical thinking could be fostered due to the forum's asynchronous nature, giving ample time to participants to read messages and plan responses well before posting them (Garrison & Anderson, 2003). Moreover, a well-designed environment is crucial to the success of OCL, as shown in Sun, Franklin, and Gao (2015), in the context of informal English language education. Thus, there is a need to design effective online learning environments to ensure the success of OCL.

Regarding the need to design an effective OCL environment, especially one that adopts the CoI framework, there are limitations found in the discourse on CoI. Despite being widely adopted and referred to in a number of studies, use of the CoI framework is still limited to addressing two issues. First, recent studies have focused only on mapping instructional activities to certain CoI presences (Dunlap & Lowenthal, 2018; Fiock, 2020; Stephen & Roberts, 2017; Stewart, 2017). Second, research into design that encompasses both instructional and interaction design aspects has yet to evaluate the effectiveness of any proposed design to foster and sustain the three CoI presences, as shown in Fiock (2020), or is limited to only a single feature supporting a single CoI presence, as demonstrated in Faisal, Junus, and Santoso (2020). Therefore, in order to address these research gaps, the following research questions (RQ) were posed in this study.

- RQ1: How is the interaction design of an AODF that is intended to foster and sustain CoI presences developed?
- RQ2: What are learners' reactions when interacting with a prototype of an AODF that is intended to foster and sustain CoI presences?

In order to answer the research questions, the interaction (i.e., scenario of use) and user interface (i.e., graphical elements) design of an asynchronous online forum were developed in this study. The prototype designs were evaluated, applying user insights to further improve the designs.

Relevant Theories

The Community of Inquiry (CoI) Framework

The CoI is defined by Garrison (2016) as a community in which participants engage in free inquiries to experience meaningful and complete learning of a concept and the inquiry process itself. Furthermore, the notion of inquiry refers to a collaborative approach to problem solving through reflective and interactive discussions among learners. There are aspects that are brought up by participants in the community. These aspects are shown in Figure 1 as the CoI presences that interact to form educational experiences in the context of a collaborative and constructive learning experience (Garrison & Anderson, 2003.

Each of the three presences, (social, cognitive, and teaching,) has specific categories that define it. Each category consists of several indicators which represent the operationalization of the presence. Garrison and Arbaugh (2007) outlined the categories of the CoI presences and gave examples of their indicators. Social presence has three categories, namely open communication, group cohesion, and affective. Cognitive presence has four categories, namely triggering events, exploration, integration, and resolution. Teaching presence has three categories, specifically learning activity design, direct instruction, and discourse facilitation. Each category has indicators. For example, risk free expression and the use of emoticons are indicators for open communication and affective category respectively.

The categories and indicators could be used as a reference to develop a coding scheme for analyzing the contents of discussion transcripts in order to diagnose what CoI presences are being shown by participants in a discussion forum (Garrison & Arbaugh, 2007). The indicators are only some examples that represent CoI presences.

The cognitive presence is operationalized in the practical inquiry model (Garrison, 2016). This model includes several phases that could be undertaken by learners in the process of constructing and negotiating meaning through both private world reflection and a shared world discourse. The phases include initial triggering events, exploration, integration, and resolution. Triggering events which take place in a shared world discourse could foster participants' awareness of the problem. This triggers exploration and integration in a private reflection to produce conception (i.e., possible solutions). The conception was then discussed to make a resolution. In this study, we used the practical inquiry model as a reference for defining the context of the use of the AODF.

The operationalization of the CoI presences is closely related to the medium in which the education experience is taking place. The next subsection describes the evolving concept of e-learning and the use of an AODF to facilitate OCL based on the CoI framework.

E-learning and the Asynchronous Online Discussion Forum

The concept of e-learning has gone through different phases of development that coincide with the development of information technology in education. As a result, the term e-learning has evolved along with its development phases. According to a study on its history by Aparicio, Bacao, and Oliveira (2016), the development of e-learning started from the advent of computer-assisted instruction.

In its subsequent development, e-learning referred to the use of Internet technology to facilitate distance learning (Garrison & Anderson, 2003). At present, a comprehensive definition of e-learning has been proposed by Sangrà, Vlachopoulos, and Cabrera (2012), who defined e-learning as an approach in learning that represents all or some aspects of a model that uses media and electronic devices to enhance accessibility, communication, and interaction, or facilitate the adoption of new understandings of learning developments. Similar to the definition by Sangrà et al. (2012), Aparicio et al. (2016) further expanded the scope of e-learning. More than just the use of computer technology to facilitate learning, it encompasses learning strategies, methods, content diffusion, and connections.

In regard to the application of the CoI framework to an e-learning system, the AODF plays an important part. The AODF is one feature of an e-learning system that could facilitate collaborative thinking activities, which are the main activities in online collaborative learning based on the CoI framework (Garrison, 2016). The AODF is a text-based communication medium with certain limitations in comparison to synchronous media, e.g., the inability to facilitate the expression of rich emotional cues (intonations, live facial expressions, etc.). However, an AODF can provide ample time to participants to reflect on how to respond well to ideas expressed in the forum. Thus, the AODF may foster critical thinking and facilitate higher-order learning (Garrison & Anderson, 2003).

The important role of an AODF in facilitating OCL underscores the need to provide effective AODF design to foster and sustain the CoI presences to ensure the attainment of learning goals. Designing interaction in an AODF is a crucial activity to ensure every participant has a meaningful and satisfying learning experience. The following subsection describes the relevant principles of interaction design for creating an effective AODF.

Principles of Interaction Design

Interaction design is an important aspect of computerized systems, including e-learning systems. Interaction design is defined as the development of interactive products to support human communication and interaction in everyday life (Sharp, Preece, & Rogers, 2019).

There are some types of interaction that need to be catered to in the development of an e-learning system, namely learner-interface interaction, learner-content interaction, learner-support system interaction, and learner-context interaction (Anderson, 2008). Moreover, in the context of applying the CoI framework to an e-learning system, a meaningful learning experience is supported by three interacting entities that include learners, facilitators, and contents (Garrison & Anderson, 2003).

Effective interface design is crucial in developing an e-learning system. It is related to the usability of the system. Usability is defined as a set of quality attributes that determines the ease of use and the ease of learning how to use a product (Nielsen, 2012). Both the ease of use and the learnability are two of the six usability goals outlined by Sharp et al. (2019).

There are some best practices that serve as a guideline in developing interface design to attain the usability goals. One set of best practices is Shneiderman's eight golden rules of interface design (Shneiderman & Plaisant, 2005). Another is Nielsen's ten usability heuristics (Nielsen, 2020). Apart from referring to best practices in user interface design, this study adopted a user-centered design (UCD) method that has been adapted to the context of e-learning based on learning-centered design (LCD). Details of the UCD and LCD are further discussed in the next section.

Methodology

This study adopted a research by design approach using an exploratory case study, which puts qualitative data into focus (Lazar, Feng, & Hochheiser, 2017) to explore the possible design solutions for a specific context. In this study, the specific context is the application of the CoI framework to the design of an AODF.

This study adopted the UCD method (Sharp et al., 2019; U.S. Department of Health and Human Services, 2006). UCD is an iterative method of designing interaction that consists of five phases: (a) specifying the context of use; (b) defining requirements; (c) designing solutions; (d) prototyping; and (e) evaluation. In this study, the UCD was also adapted to the context of e-learning system development with the involvement of experts as subject-matter and e-learning specialists. This adaptation reflects the LCD concept proposed by Dhar and Yammiyavar (2012), which focuses on providing an online environment that supports the learning process rather than just good usability.

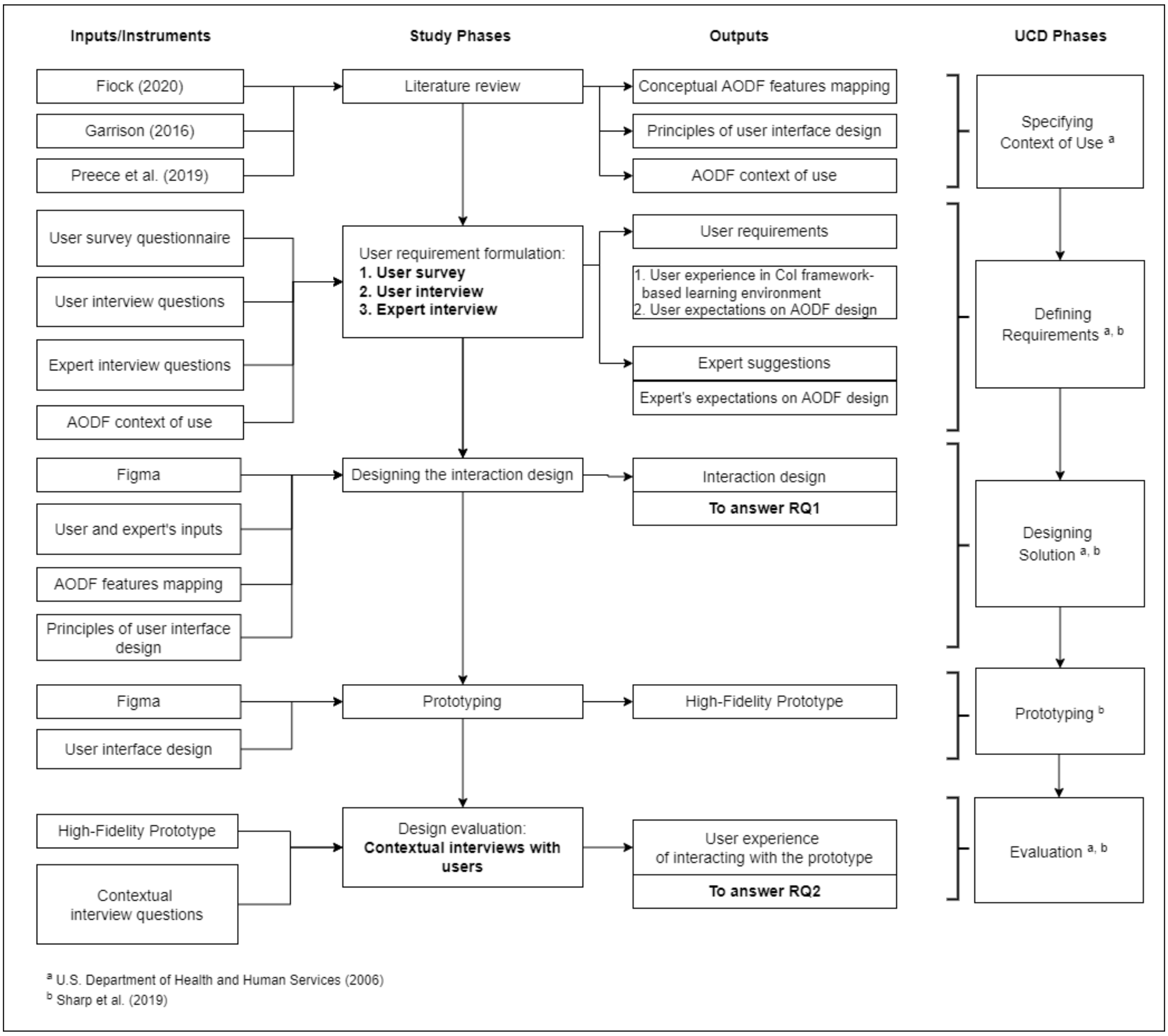
There were five stages in this study, namely: (a) literature study; (b) identifying user requirements; (c) designing the interaction design; (d) prototyping; and (e) design evaluation. These stages, including the outputs of each stage, are illustrated in Figure 2.

The *literature study* stage is related to a phase in UCD in which users and their context of use are identified (U.S. Department of Health and Human Services, 2006). In this study, specifying the context of use was carried out through a literature review of research in the domain of the CoI framework. The outputs of this stage were the user characteristics and the context in which the AODF was used.

The *identifying user requirements* stage is related to another phase in UCD: defining user requirements. During this phase, user needs and goals as criteria for product success are identified (U.S. Department of Health and Human Services, 2006). The output of this stage is a set of design requirements. In this study, both students and lecturers who used the AODF were involved through online surveys and in-depth interviews. Online surveys involved 37 respondents (26 students and 11 lecturers), in-depth user interviews involved ten students, and expert interviews involved four specialists in OCL.

Figure 2

The Stages of This Study



The *designing the interaction design* stage is related to the designing solution phase in UCD. During this stage, high-fidelity user interfaces and contextual scenarios were developed as the interaction design of the AODF. The high-fidelity user interfaces are meant to closely resemble the final product (Sharp et al., 2019). The tool used to design the high-fidelity interfaces was Figma (www.figma.com). Figma was chosen because of its free prototyping features and unlimited design files. The design solutions developed in this stage consisted of the AODF user interface design and its contextual scenarios that describe how the AODF is used in certain situations (e.g., during the early phase of OCL, etc.). Both outputs were used to answer RQ2.

The *prototyping* stage corresponds with the prototyping phase in UCD. In this stage, an AODF mock-up was developed based on the high-fidelity user interfaces from the previous design stage. The mock-up was a high-fidelity clickable prototype made using Figma's prototyping feature.

The *design evaluation* stage is related to the evaluation phase in UCD. During this stage, design solutions are tested by real users (U.S. Department of Health and Human Services, 2006). In order to answer RQ2, the user experience and users' perceptions regarding the prototype's ability to foster and sustain the CoI presences were captured in this stage. The evaluation was conducted qualitatively through contextual semi-structured interviews guided by contextual scenarios. The contextual interviews involved ten student participants.

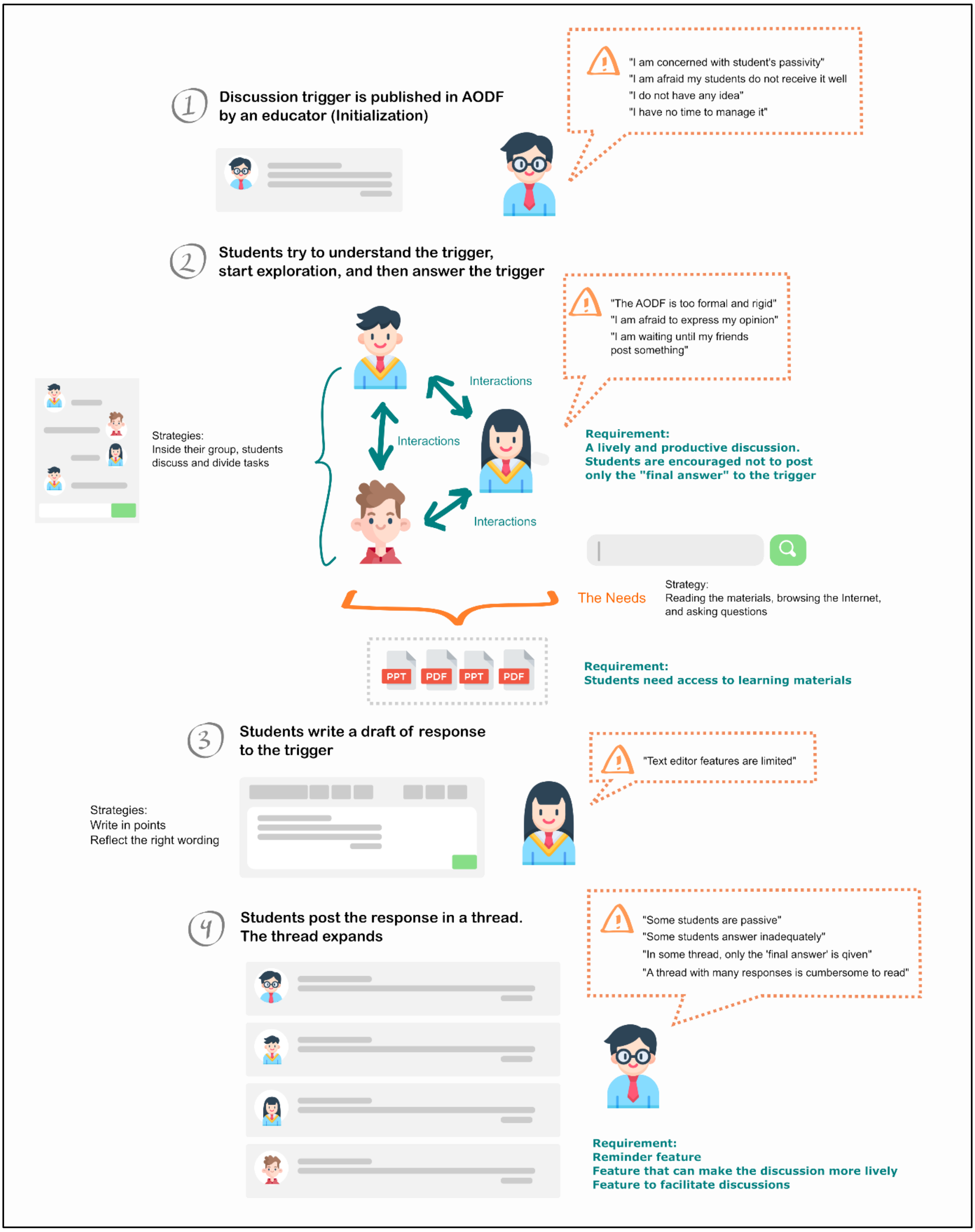
User requirements, user perceptions, and user experience when interacting with the prototype were obtained using the methods illustrated in Figure 2. The results of in-depth interviews and online surveys in the early phase of this study which were used to formulate the design requirements are presented in the next section. The results of the design evaluation are discussed in the Evaluation Results section.

Design Requirements and Solutions

Students, lecturers, and experts expressed various views and reported diverse issues related to fostering and sustaining the CoI presences in the AODF. A summary of the findings from the in-depth interviews and surveys is illustrated in the rich picture shown in Figure 3.

Figure 3

Rich Picture of Findings from In-Depth Interviews and Surveys



Note. This rich picture has been designed using resources from Flaticon.com.

As shown in Figure 3, some students expressed the notion that they considered an AODF as a place for final answers where they are to give the one response they consider the best and most correct to the questions or problems presented by the lecturer. These phenomena were also observed in Junus et al. (2019) as some of the challenges faced by students in CoI-based AODF activities. Dynamic discussions in which exploration and integration activities exist often occur in internal groups outside the AODF. This could lead to unnatural responses written by students in the AODF and hinder efforts to diagnose and correct misconceptions during the early phase of discussions.

As shown in Figure 3, in any discussion thread with many participants, the large and growing number of responses hinder students' ability to monitor the discussion. Some students said that reading replies one by one was very tiring. One strategy adopted by some students to solve this issue was making notes when planning a response.

The in-depth interviews and surveys also revealed some AODF design requirements related to CoI presences. Table 1 presents a summary of the AODF design requirements that were gathered from the views of students, lecturers, and experts.

Table 1

Summary of AODF Design Requirements to Facilitate the CoI Presences

CoI presence	AODF design requirements		
	Students	Lecturers	Experts
Social	Provides emotional cues Facilitates mentioning other participants Supports an informal atmosphere Facilitates rich interactions that resemble face-to-face interactions (e.g., showing who is online, showing who is writing, etc.) Provides posting features for greeting other participants	Provides posting features for initiating discussions Provides emotional cues Facilitates mentioning other participants	Provides profile pictures Provides emoticons Provides digital stickers Provides quick response features (e.g., like, love, & clap) Provides a grouping feature Provides mention with notification feature Provides quoting feature Provides discussion analytics Provides badges for the participants
Cognitive	Provides access to relevant learning materials Enables students to take notes before responding Enables the students to comprehend the main	Provides posting features for giving questions and feedbacks Enables lecturers to score students' participation	Provides smaller breakout forums Provides activities reminders Provides posting attributes (e.g., questions, asking for help, etc.)

CoI presence	AODF design requirements		
	Students	Lecturers	Experts
	points of the discussion as the discussion progresses Provides a text editor that supports easy font formatting, formula writing, image resizing, and indentation formatting		Provides activity guideline
Teaching	Enables the students to monitor the progress of the discussion Provides posting features for reminding, triggering responses, informing the mechanism of the discussion, and dividing tasks.	Provides posting features for giving questions and feedbacks Enables the lecturers to monitor the progress of the discussion Enables the lecturers to give information on the discussion mechanism Enables the lecturers to know if a posting has been read by other participants	Provides discussion analytics Provides activities reminders Provides dashboards depicting deadlines for assignments Provides identifiers for lecturers Provides a word clouds feature

Note. AODF = asynchronous online discussion forum.

The interaction design of the AODF was developed based on the design requirements presented in Table 1 and stages included in the practical inquiry model as the contexts of use, which reflect how a discussion progresses from triggers presentation to resolution. Table 2 presents AODF feature mapping as one of the deliverables of AODF interaction design.

Table 2

AODF Feature Mapping

Context of use	Related indicators of the CoI presences/practical inquiry phases	AODF features
Ice-breaking activities	Social presence: Open communication, affection, and group cohesion	Thread & reply, mention, text editor, discussion onboarding, discussion guide, and pop-ups (with instructions for ice-breaking activities), emotional cues (emoticons, stickers, & GIFs), quick reactions (like, love, & claps), and profile picture

Context of use	Related indicators of the CoI presences/practical inquiry phases	AODF features
Comprehending and expressing initial understandings of the trigger	Cognitive presence: Triggering events (initial phase)	Thread & reply, mention, text editor, discussion references (learning materials attachment card), discussion onboarding, discussion guide, and pop-ups *(with instructions for ice breaking-initial triggering events-related activities)*
Exploring and sharing relevant learning materials	Cognitive presence: Transition from the triggering events (late phase) to the exploration (initial phase)	Thread & reply (multimedia attachments & tags), mention, text editor, discussion references (learning materials attachment card), summarizing tool, discussion analytics, discussion onboarding, discussion guide, and pop-ups (*with instructions for exploration-related activities*)
Assisting participants in need of assistance by facilitating the discussion through giving direct instruction	Teaching presence: Direct instruction (in exploration phase context)	Thread & reply, mention, text editor, tags, discussion analytics, word clouds, discussion onboarding, and pop-ups (*with triggers to encourage helping participants in need of assistance*)
Selecting and elaborating ideas that have been discussed during the exploration	Cognitive presence: Transition from the exploration (late phase) to the integration (initial phase)	Thread & reply, mention, text editor, discussion references (learning materials attachment card), summarizing tool, discussion onboarding, discussion guide, and pop-ups (*with instructions for integration-related activities*)
Discussing conclusion	Cognitive presence: integration (late phase)	Thread & reply, mention, text editor, discussion references (learning materials attachment card), summarizing tool, discussion analytics, word clouds, discussion onboarding, discussion guide, and pop-ups (*with instructions for discussing the conclusion of the discussion*)
Formulating discussion resolution	Cognitive presence: Transition from the integration (late phase) to the resolution (initial phase)	Thread & reply, mention, text editor, discussion references (learning materials attachment card), summarizing tool, discussion onboarding, discussion guide, and pop-ups (*with instructions for discussing the resolution of the discussion*)

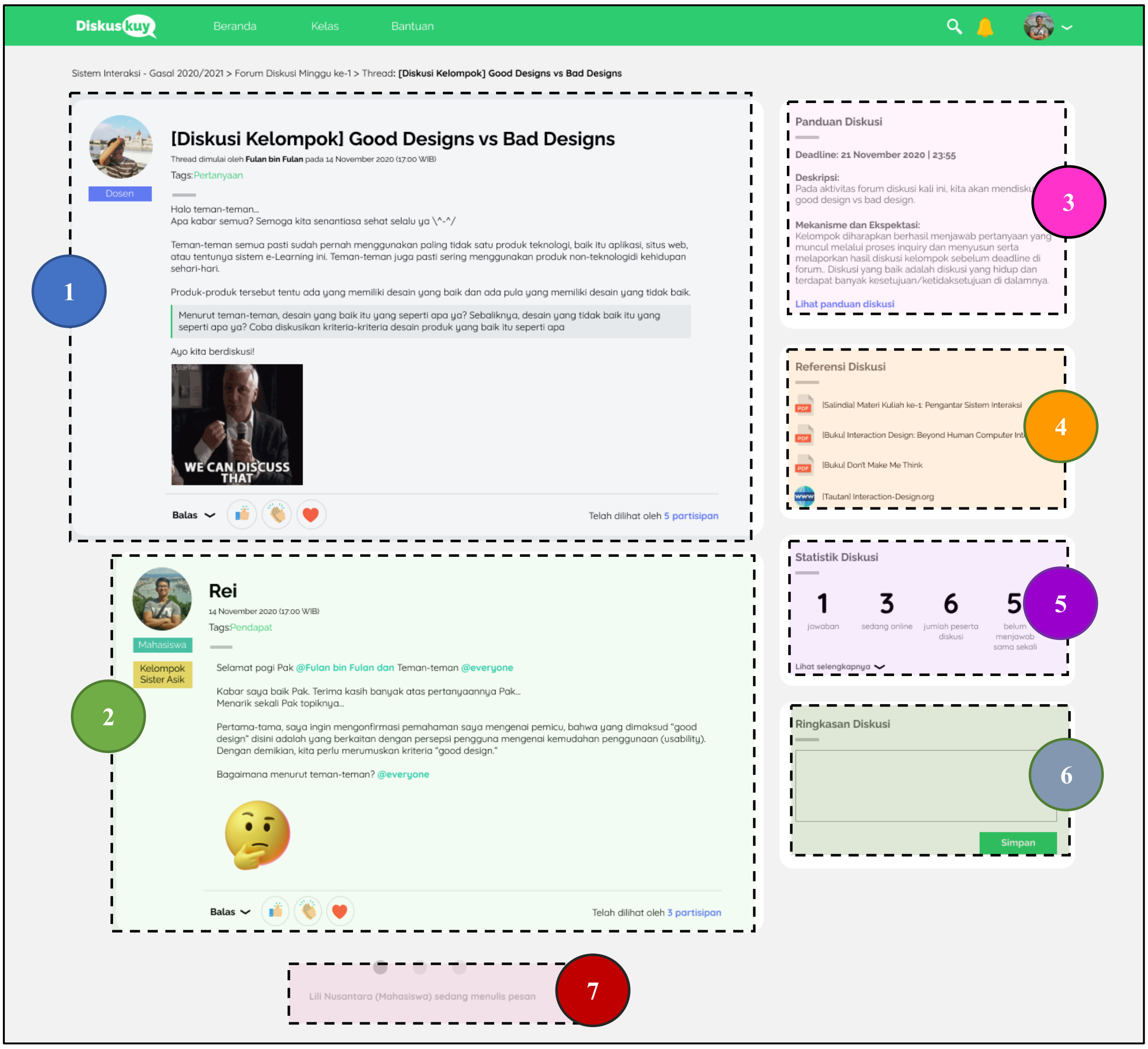
The AODF interaction design developed in this study consisted of a prototype containing 14 features mapped to the CoI presences and seven scenarios that described seven contexts of use as presented in Table 2. The AODF features included: thread and reply; text editor; profile picture; emotional cues (emoticons, stickers, and GIFs); quick reactions (like, love, and claps); mentions; quotes (as a text highlighting feature included in the text editor); discussion guide; discussion references (learning materials attachment card); discussion analytics; word clouds; summarizing tool; discussion onboarding; and pop-ups. Some of these features are displayed on the user interface of the AODF as shown in Figure 4.

In Figure 4, number 1 indicates a thread message to initialize a discussion. Below the thread message, number 2 indicates a reply posted by a participant. On the right side of the thread page, there are some features such as a discussion guide (number 3), discussion reference (number 4), discussion analytics (number 5), and summarizing tool (number 6). Profile picture, emotional cues (emoticons, stickers, and GIFs), quick reactions (like, love, and claps), mentions, and quotes (text highlighting) are located inside the reply section. In addition, the word clouds feature is located inside the discussion analytics section and can be shown by clicking on the *see more* button. Other features include pop-ups and onboarding, which are shown during specific phases of inquiry (e.g., the triggering event). Additionally, a notification that says *someone is writing* (number 7) is shown when a participant is writing a reply in the thread.

The discussion analytics feature helps users to understand and monitor the progress of a discussion through simple numerical data (e.g., number of answers, number of online participants, total number of participants, and number of participants who have not responded to the discussion thread) and word clouds. Similar to the discussion analytics feature and word clouds, the summarizing tool is intended to facilitate participants in understanding the progress of the discussion. This feature arises from student suggestions (see Table 1) and from the results of Faisal et al. (2020).

Figure 4

Interface Design of a Page in the AODF



Note. 1 = initiated discussion; 2 = reply; 3 = discussion guide; 4 = discussion reference; 5 = discussion analytics; 6 = summarizing tool; 7 = "someone is writing" notification.

Apart from these features, the designed prototype also included onboarding and pop-ups. The onboarding and pop-ups were dynamic features whose appearance and content varied in different inquiry phases. The appearance and content of the onboarding and pop-ups features were adjusted to the indicators of each inquiry phase, according to Shea et al. (2010). The onboarding feature was intended to provide an overview of the activities that needed to be carried out at a certain inquiry stage, information about discussion mechanisms, expectations, and discussion deadlines. Meanwhile, the pop-ups were intended to inform discussion participants about the relevant features in each phase of inquiry and guide discussion activities.

In addition to the user interface design and the contextual feature mapping based on CoI presence indicators, the interaction design developed in this study included contextual scenarios that describe the AODF scenario for each context of use as presented in Table 2. An example of the user scenario for ice-breaking activities is shown here.

The AODF is used for self-introductory and discussion orientation activities. In this study, the facilitator opens the self-introductory session by starting a thread with a post expressing greetings and sharing experiences that are relevant to the discussion topic. Emotional cues such as emoticons, stickers, GIF images, or multimedia attachments could be used to create a lively and informal discussion environment. After the self-introductory session, the following steps are undertaken by the participants.

In the first step, the participants reply to the thread. The participants are then shown onboarding cards that give them some information about the purpose of the discussion, the mechanism, the expectation of the facilitator, and the deadline for discussion outputs.

In the second step, the participants enter the thread page and are shown a pop-up that encourages them to upload their profile pictures to enhance their social presence (if participants have yet to upload their profile pictures).

In the third step, after uploading profile pictures, the participants are shown a pop-up that instructs them to create a lively discussion environment by using informal language, having a sense of respect towards other participants, using emotional cues (e.g., emoticons, etc.) features in their reply, and using quick reply features (e.g., claps, like, etc.).

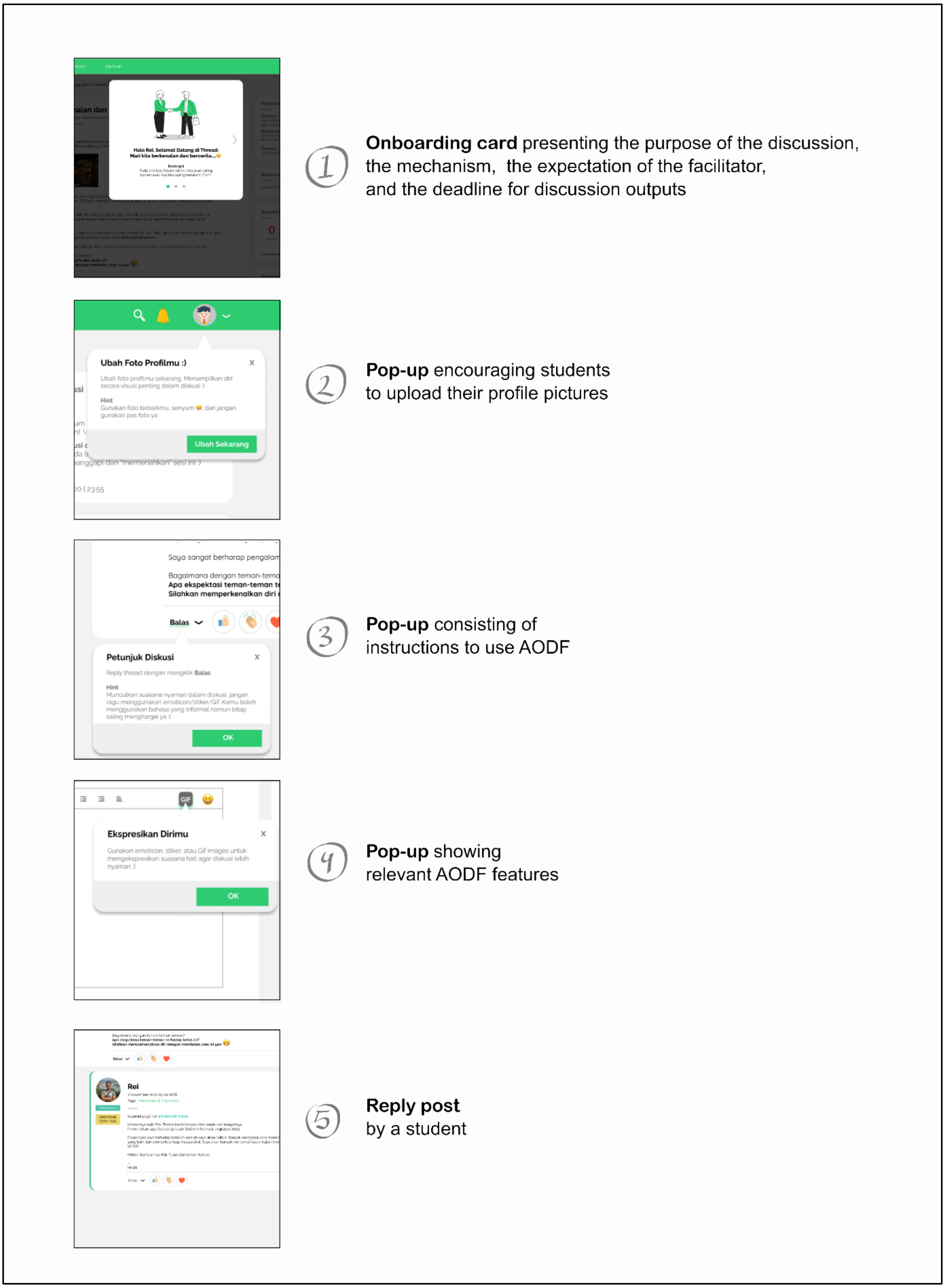
In the fourth step, when participants open the text editor to write a reply in the thread, a pop-up shows some relevant AODF features located inside the reply section (e.g., stickers, emoticons, etc.), which can be used in self-introduction.

In the fifth step, the participants post their replies in which the content includes self-introduction.

In this scenario, the AODF features used were discussion onboarding, discussion guide, and pop-ups (with instructions for ice-breaking activities), emotional cues (emoticons, stickers, and GIFs), quick reactions (like, love, and claps), and profile picture. The scenario is illustrated in a storyboard shown in Figure 5.

Figure 5

Storyboard for Ice-Breaking Activities in the AODF



The interaction design was then tested with ten students. A summary of the results is presented in the next section.

Evaluation Results

The contextual testing conducted on the prototype provided insights into how participants viewed the prototype. Ten students who had prior experience in using an AODF for OCL activities were involved. The context was an introductory discussion activity and a general orientation discussion on good versus bad design through inquiry stages. The results of this contextual testing were intended to answer RQ2.

The students were asked to do various tasks on the prototype and then describe their experiences when interacting with the prototype. The testing was conducted using scenarios that were adapted for evaluation purposes (e.g., giving participants the context of use without aiding them in completing given tasks). Participants' perceptions within the context of the introductory activity (task 1) and the specific stages of inquiry (task 2–task 7) are summarized in Table 3.

Table 3

Summary of the Contextual Testing Results

Task	Awareness of inquiry stages	Ability to encourage inquiry stages-related activities	Ability to facilitate inquiry stages-related activities	Helpful AODF features according to participants
Task 1: Pre-discussion introduction (ice-breaking activity)	All participants (n = 10) were *aware* that they needed to introduce themselves due to explicit instructions given in the prototype.	Some participants (n = 5) said that the prototype encouraged them to introduce themselves, but others (n = 4) said that direct instruction from lecturers would be more encouraging.	Most participants (n = 8) said that the prototype *helped* them to introduce themselves in the forum.	Emoticons, GIFs, quick reactions (like, love, & claps), tags, mention, pop-ups, participant identifier, list of participants who had read a thread/post, and profile picture.
Task 2: Expressing initial understanding of the trigger of the discussion	The majority of participants (n = 7) had a tendency to directly post a final answer or a solution that indicated exploration-related activities to the problems presented in the trigger.	Some participants (n = 4) felt compelled to understand and respond to the trigger despite having different expectations on how the trigger	Some participants (n = 5) said that the prototype helped them to express their answers.	Discussion references [a], pop-ups [a], discussion guide [a], tags [a], mention [a], notification [a], quotes (text highlighting) [a], and a live list of

Task	Awareness of inquiry stages	Ability to encourage inquiry stages-related activities	Ability to facilitate inquiry stages-related activities	Helpful AODF features according to participants
		should be responded to.		participants writing a reply.
Task 3: Sharing information acquired after exploring relevant sources to answer the trigger	The participants understood that they needed to explore relevant sources (n = 4), respond to the trigger (n = 3), read relevant references (n = 1), and continue the discussion (n = 1).	Some participants (n = 5) said that the prototype encouraged them to explore relevant sources. Others disagreed (n = 3).	The majority of participants (n = 8) said that the prototype helped them to share relevant information.	Multimedia attachments, discussion references, summarizing tool, pop-ups, links, mention, and discussion guide.
Task 4: Helping others understand the problems by diagnosing and correcting misconceptions	The majority of participants (n = 8) were aware that they needed to help others in need (i.e., confused by the problems, etc.).	Most participants (n = 6) felt encouraged to help others. However, a participant stated that it would be more encouraging if there was a notification for a direct request for help.	Almost all participants (n = 9) felt that the prototype was helpful in facilitating participants to help others in need.	Tags (a tag with the indication "I'm confused"), a sign on a reply post showing the participant in need of help, pop-ups, and word clouds.
Task 5: Integrating ideas	Some participants (n = 5) were aware of the integration phase, but some were confused by the term (n = 4).	Some participants (n = 3) said the prototype encouraged them to integrate ideas. However, others disagreed (n = 2).	Some participants (n = 5) said the prototype helped them integrate ideas. However, one participant disagreed.	Summarizing tool [b], pop-ups [b], word clouds [b], and the positioning of the text editor for a reply in the thread page [b].
Task 6: Making a conclusion	The majority of participants said they *understand* the context of use, which was formulating a conclusion to the discussion (n = 7).	Almost all participants stated that the prototype *encouraged* them to make a conclusion to	Almost all participants stated that the prototype *helped* them in making a	Summarizing tool, pop-ups, discussion analytics, and

Task	Awareness of inquiry stages	Ability to encourage inquiry stages-related activities	Ability to facilitate inquiry stages-related activities	Helpful AODF features according to participants
		the discussion (n = 9).	conclusion (n = 9).	discussion guide.
Task 7: Understanding the achievements of the discussion and formulating the resolution	Some participants were *aware* that they were expected to make a discussion resolution (n = 5), but the majority had no idea what a discussion resolution is (n = 6).	The majority of participants said that the prototype *encouraged* them to make a resolution (n = 6). However, a participant disagreed.	The majority of participants said the prototype was unhelpful in forming a resolution (n = 6). However, some stated that the prototype was helpful (n = 4).	Summarizing tool [c], pop-ups [c], and discussion references [c].

Note. [a] AODF features that were expressed in tasks in which there were participants who intended to reply in the AODF with messages indicating exploration. [b] AODF features that were expressed in tasks in which there were participants who were confused about what integration is or what messages were expected to indicate integration. [c] AODF features that were expressed in tasks in which there were participants who had confusion on what resolution is or what messages were expected to indicate resolution.

The summary of participants' responses to the prototype as shown in Table 3 is helpful in answering RQ2. The prototype was found to be helping most participants in the pre-discussion introduction (ice-breaking activities; task 1), sharing information acquired from exploring relevant sources to answer the trigger (task 3), helping others in understanding the problems by diagnosing and correcting misconceptions (task 4), and formulating conclusions (task 6). However, there were issues in expressing initial understanding of the trigger of the discussion (task 2), integrating ideas (task 5), and understanding the achievements of the discussion and formulating a resolution (task 7). These issues are related to cognitive presence at the inquiry phases of triggering events, initial integration, and resolution.

Nevertheless, in general, all participants responded positively to the AODF prototype (n = 10). Most also stated that the prototype provided a different and positive experience compared to other discussion forums. Moreover, some participants (n = 8) stated that they were aware of the inquiry phases despite having difficulties in understanding what to do in certain phases.

Conclusion

This study developed an AODF interaction design based on the CoI framework using the UCD method, which was contextualized for the development of an e-learning system. The design requirements were defined by students, lecturers, and OCL experts.

To answer RQ1 (How is the interaction design of an AODF that is intended to foster and sustain CoI presences developed?), we created an AODF interaction design consisting of a clickable high-fidelity prototype containing 14 features mapped against the presences in CoI and seven scenarios describing the use of various features during different inquiry phases based on the practical inquiry model. The 14 features were mapped to foster and facilitate the CoI presences in the context of an initial self-introduction activity and a general orientation discussion that was carried out through four stages of inquiry.

To answer RQ2 (What are learners' reactions when interacting with a prototype of an AODF that is intended to foster and sustain CoI presences?), all contextual testing participants (n = 10) responded positively to the AODF prototype. Most participants also considered that the prototype provided a different and positive experience when compared to other online discussion forums. Participants revealed that the prototype helped them to create a social presence in the context of the initial self-introduction activity, cognitive presence in the context of drawing conclusions, and teaching presence when there were participants in need of help during the exploration phase.

There were several suggestions for improvement made by participants. Issues that arose were related to the prototype's user interface and user perceptions regarding the effectiveness of AODF interaction design in facilitating activities to foster cognitive presence at the inquiry phases of triggering events, initial integration, and resolution.

This study has practical implications, providing descriptions of how the user interface and interaction of an AODF could be designed to foster CoI presences. Further research could be aimed at evaluating the proposed AODF design quantitatively in a classroom setting and identifying AODF features that significantly nurture CoI presences. Moreover, the UCD and LCD methods used in this study could be adopted when designing other e-learning systems in various contexts.

References

Anderson, T. (Ed.). (2008). *The theory and practice of online learning*. Athabasca University Press. https://read.aupress.ca/projects/the-theory-and-practice-of-online-learning

Aparicio, M., Bacao, F., & Oliveira, T. (2016). An e-learning theoretical framework. *Educational Technology & Society, 19*(1), 292–307. http://www.jstor.org/stable/jeductechsoci.19.1.292

Athabasca University, Centre for Distance Education. (2021). *About the framework: An introduction to the community of inquiry*. http://thecommunityofinquiry.org/coi

Bransford, J. D., Brown A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school: Expanded edition*. National Academy Press.

Dhar, D., & Yammiyavar, P. (2012). *Design approach for e-learning systems: Should it be user centered or learner centered*. In S. Murthy & D. Sampson (Eds.), *Proceedings: 2012 IEEE fourth international conference on technology for education* (pp. 239–240). IEEE. https://doi.org/10.1109/T4E.2012.57

Dunlap, J. C., & Lowenthal, P. R. (2018). Online educators' recommendations for teaching online: Crowdsourcing in action. *Open Praxis, 10*(1), 79–89. https://openpraxis.org/articles/10.5944/openpraxis.10.1.721/

Faisal, T. A., Junus, K., & Santoso, H. B. (2020). Development of the online collaborative summarizing feature on student-centered e-learning environment. In S. Yazid (Chair), *Proceedings: ICACSIS 2019: 11th international conference on advanced computer science and information systems,* (pp. 435–440). IEEE. https://doi.org/10.1109/ICACSIS47736.2019.8979883

Fiock, H. S. (2020). Designing a community of inquiry in online courses. *The International Review of Research in Open and Distributed Learning, 21*(1), 135–153. https://doi.org/10.19173/irrodl.v20i5.3985

Garrison, D. R. (2016). *Thinking collaboratively: Learning in a community of inquiry*. Routledge.

Garrison, D. R., & Anderson, T. (2003). *E-learning in the 21st century: A framework for research and practice*. Routledge.

Garrison, D. R., & Arbaugh, J. B. (2007). Researching the community of inquiry framework: Review, issues, and future directions. *Internet and Higher Education 10*(3), 157–172. https://doi.org/10.1016/j.iheduc.2007.04.001

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), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Hendry, G. D., Frommer, M., & Walker, R. A. (1999). Constructivism and problem-based learning. *Journal of Further and Higher Education, 23*(3), 369–371. https://doi.org/10.1080/0309877990230306

Junus, K., Suhartanto, H., Suradijono, S. H. R., Santoso, H. B., & Sadita, L. (2019). The community of inquiry model training using the cognitive apprenticeship approach to improve students' learning strategy in the asynchronous discussion forum. *Journal of Educators Online, 16*(1), 1–17. https://files.eric.ed.gov/fulltext/EJ1204388.pdf

Laal, M., Laal, M., & Kermanshahi, Z. (2012). 21st century learning: Learning in collaboration. *Procedia - Social and Behavioral Sciences, 47*, 1696–1701. https://doi.org/10.1016/j.sbspro.2012.06.885

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction (2nd Edition)*. Morgan Kaufmann.

Lipman, M. (2003). *Thinking in education* (2nd ed.). Cambridge University Press.

Nielsen, J. (2020, November 15). *10 Usability heuristics for user interface design*. NN/g Nielsen Norman Group. https://www.nngroup.com/articles/ten-usability-heuristics/

Nielsen, J. (2012, January 3). *Usability 101: Introduction to usability*. NN/g Nielsen Norman Group. https://www.nngroup.com/articles/usability-101-introduction-to-usability/

Sangrà, A., Vlachopoulos, D., & Cabrera, N. (2012). Building an inclusive definition of e-learning: An approach to the conceptual framework. *The International Review of Research in Open and Distributed Learning, 13*(2), 145–159. https://doi.org/10.19173/irrodl.v13i2.1161

Sharp, H., Preece, J., & Rogers, Y. (2019). *Interaction design: Beyond human-computer interaction* (5th ed.). Wiley.

Shea, P., Hayes, S., Vickers, J., Gozza-Cohen, M., Uzuner, S., Mehta, R., & 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. http://dx.doi.org/10.1016/j.iheduc.2009.11.002

Shneiderman, B., & Plaisant, C. (2005). *Designing the user interface: Strategies for effective human-computer interaction* (4th ed.). Pearson Education.

Stephens, G. E., & Roberts, K. L. (2017). Facilitating collaboration in online groups. *Journal of Educators Online, 14*(1), 1–16. https://files.eric.ed.gov/fulltext/EJ1133614.pdf

Stewart, M. K. (2017). Communities of inquiry: A heuristic for designing and assessing interactive learning activities in technology-mediated FYC. *Computers and Composition, 45,* 67–84. https://doi.org/10.1016/j.compcom.2017.06.004

Sun, Y., Franklin, T., & Gao, F. (2015). Learning outside of classroom: Exploring the active part of an informal online English learning community in China. *Visual Communication and Technology Education Faculty Publications, 38*. https://scholarworks.bgsu.edu/vcte_pub/38

U.S. Department of Health and Human Services. (2006). User-centered design basics. In *The research-based web design and usability guidelines, enlarged/expanded edition*. https://www.usability.gov/what-and-why/user-centered-design.html

Valverde-Berrocoso, J., Garrido-Arroyo, M. d. C., Burgos-Videla, C., & Morales-Cevallos, M. B. (2020). Trends in educational research about e-learning: A systematic literature review (2009–2018). *Sustainability, 12*(12), 5153. https://doi.org/10.3390/su12125153

Vaughan, N. D., Cleveland-Innes, M., & Garrison, R. D. (2013). *Teaching in blended learning environments: Creating and sustaining communities of inquiry*. Athabasca University Press. https://read.aupress.ca/projects/teaching-in-blended-learning-environments

Athabasca University

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

Are K–12 Teachers Ready for E-learning?

Elif Polat[1], Sinan Hopcan[1], and Ömer Yahşi[2]
[1]Department of Computer Education and Instructional Technology; Hasan Ali Yucel Education Faculty, Istanbul University-Cerrahpasa, Istanbul, Turkey; [2]Ministry of National Education, Ankara, Turkey

Abstract

Readiness is important for the success of the e-learning process. The purpose of this study was twofold: to develop a scale to measure K–12 teachers' e-learning readiness, and to examine their readiness to teach online. The participants were 3,295 K–12 teachers working in Izmir, Turkey. First EFA, then CFA-SEM was performed. Additionally, teachers' e-readiness in terms of gender, years of service, school level, and daily device usage time were examined. Teachers are ready for e-learning considering their overall scores. A significant difference was found in favor of males in the "technical competence" factor and in favor of females in the "colleague, content, and pedagogical and ethical competence factors". The readiness of younger teachers is generally higher. On a factor basis, there is only a significant difference in the factors of computer self-efficacy and student readiness according to educational level. As the use of devices increases, technology-related readiness increases. The readiness of teachers plays an important role in determining future strategies, measures, and interventions that need to be taken to advance e-learning.

Keywords: e-learning, readiness factor, teachers' e-learning readiness, scale development, K–12 teachers

Introduction

Readiness is important for the success of the e-learning process. For effective e-learning, it is essential to understand each individual's readiness. Lopes (2007) demonstrated that technology, content, culture, human resources, and financial resources affect e-learning readiness. Hong and Gardner (2018) stated that e-readiness includes self-efficacy, self-regulation, social competence, and digital competence. Aydin and Tasci (2005) identified four areas that determine the overall readiness of an institution to adopt e-learning, namely, technology, innovation, people, and self-development. As teachers deal directly with both students and course content, they are key to adapting and applying e-learning platforms to their learning environment and are expected to use e-learning to pursue the globalization of knowledge and provide technologically enhanced classroom interaction (Obara & Abulokwe, 2012). Teachers play a critical role in the implementation of online education (Mercado, 2008) and their readiness is dependent on factors such as the design of learning content and ensuring students are successful (Eslaminejad et al., 2010).

The success of technology in teaching and learning does not only depend on the availability and usability of technical tools such as a strong network infrastructure and fast, modern computers and applications. Where teachers are not trained to adopt and use e-learning and the technologies that facilitate it, implementation will generally be unsuccessful (Ziphorah, 2014). To ensure to the extent possible the successful introduction of an e-learning program, some form of assessment of teachers' preparation is needed as any reluctance may impact implementation (Summak et al., 2010).

Some of the variables affecting structure and interaction are gender, strategy and approaches, skills, and readiness of technology. Teachers have an important role as they take on tasks such as preparing online content and motivating students. For this reason, teachers must be well prepared if e-learning is to be a success. The e-learning environment is very different from traditional learning environments, and it is essential to ensure that teachers are able to adapt (Phan & Dang, 2017).

Teachers not only need the technical competence and ability to develop content but also knowledge of online teaching methods (Phan & Dang, 2017). The e-learning environment is not about providing a set of documents. It involves basic features such as enabling interaction with and among students, designing content, and using appropriate teaching methods (Eslaminejad et al., 2010).

Most of the existing models of e-learning readiness were designed and tested in commercial organizations and higher education institutions (Koloseni & Mandari, 2017) rather than primary and secondary schools (Summak et al., 2010). Any measurement tool developed for e-learning readiness needs to be carefully considered before applying it to a particular context. Demir and Yurdugül (2015) suggested that the selection of a developed model and measurement tools should be done according to the needs of each context and target audience, and that any deficiencies identified should be eliminated (Demir & Yurdugül, 2015).

There are a number of studies that measure the e-learning readiness of teachers (Al-Furaydi, 2013; Amalia et al., 2021; Çınar et al., 2021; Howard et al., 2021; Hu et al., 2020; Ouma et al., 2013; Pusparini et al., 2018; Setati & Paledi, 2019; So & Swatman, 2006; Trayek et al., 2016; Yun & Murad, 2006). Çınar et al. (2021), for example, investigated the readiness of in-service Turkish teachers. Their study, with 555 teachers from pre-school to high school, revealed that teachers had a medium-level of e-learning readiness. Another study with 222 secondary education teachers from different countries that aimed to determine their readiness revealed the importance of institutional support (Howard et al., 2021).

Al-Furaydi (2013) provided a descriptive analysis in his study with 71 English teachers and found they were ready to adopt e-learning and had a high level of computer literacy. In a similar study, Hu et al. (2020) examined the reliability of an e-learning readiness survey in secondary schools in Kenya. The authors conducted the study using the descriptive survey design model with 72 teachers, principals, and students. They revealed that teachers were ready to embrace e-learning technology, but their technical capacity required improvement through training for successful e-learning adoption (Ouma et al., 2013). A study by Amalia et al. (2021), conducted with 15 teachers using qualitative methods, found that teachers were ready for e-learning. According to So and Swatman (2006), teachers in Hong Kong were not fully ready to use e-learning technologies for teaching and learning. They conducted their research with 131 teachers from primary and secondary schools. Setati and Paledi (2019) assessed the e-learning readiness of 120 primary and secondary school teachers in Africa. Trayek et al. (2016) revealed the e-learning readiness of 475 secondary school teachers in Palestine. Similarly, Yun and Murad (2006) measured 412 secondary school teachers' e-learning readiness. Pusparini et al. (2018) conducted a study with a small sample of 20 people using the explanatory sequential design model to investigate the e-learning readiness of high school English teachers.

Currently the available literature has focused specifically on either only one or two primary/secondary/high school levels or on specific disciplines such as English teaching. As far as we are aware, there are no studies directly addressing K–12 teachers. In addition, the studies generally use the descriptive method, the sample sizes are not particularly large, and few studies aim to develop a scale. Despite K–12 teachers' e-readiness being crucial to the success of e-learning, there is room for much more work to be done in this area. It is vital for teachers to be prepared for online teaching. Where e-learning readiness levels are insufficient, the chance of success in e-learning is low (Moftakhari, 2013). Understanding these factors and planning for them can increase the success of K–12 institutions in applying e-learning. There appears to be a gap in the literature for both developing a valid scale to measure K–12 teachers' e-learning readiness and conducting research with robust methods using this tool with large samples. To this end, the study aimed to develop a scale to measure K–12 teachers' e-learning readiness, and to examine their readiness to teach online.

HOT-Fit Model

This study posits factors that can influence e-learning readiness using the HOT-fit model. Human organization and technology-fit (HOT-fit) is a framework developed by Yusof et al. (2008) based on a combination of DeLone and McLean's information system (IS) success model and IT organization fit model. Human factors consist of computer self-efficacy and subjective norms (Çiğdem & Topcu, 2015; Oketch et al., 2014; Zheng et al., 2018) and organizational factors comprise IS/IT knowledge and management support (Oketch et al., 2014). Technological factors are relative advantage, compatibility, and complexity (Oketch et al., 2014; Zheng et al., 2018).

Research Questions

1. What are the reliability and validity evidence of the developed scale to measure teachers' e-learning readiness?

2. What factors (gender, years of teaching, level taught, and daily device usage time) are related to teachers' e-learning readiness?

Method

A 70-question scale consisting of 13 factors was developed by the researchers. The developed scale was distributed to 3,525 people. Separate datasets were used for exploratory factor analysis (EFA) (n = 1,081), confirmatory factor analysis (CFA) (n = 1,086), and implementation (n = 1,128).

Item Generation

The item development process began after consulting the literature and studies that had examined e-learning readiness. Theoretical frameworks, models, and previous scales were carefully scrutinized (Martin, Wang, et al., 2019; So & Swatman, 2006; Texas A&M University, 2021; University of Toledo, 2021). Following the literature review on indicators of online teaching readiness, the first phase of the study generated 75 items measuring 13 constructs: (a) technical competence, (b) attitude, (c) communication skills, (d) course design/pedagogical competence, (e) time management, (f) computer self-efficacy, (g) infrastructure, (h) management support, (i) colleagues, (j) student readiness, (k) content, (l) complexity of technology, and (m) relative usefulness. These items were measured using a five-point Likert scale (-2 = *strongly disagree*, -1 = *disagree*, 0 = *neutral*, 1 = *agree*, 2 = *strongly agree*). The study used the HOT-fit model to structure the categories and factors. This model has the potential to evaluate the appropriateness of online teaching readiness (Mirabolghasemi et al., 2019).

As a key step to ensure potential respondents would be able to understand the items, one-on-one interviews were conducted with five teachers with knowledge and experience of online education. Necessary revisions were made.

In order to provide face validity, content validity, and clarity of scale items, one-on-one interviews were conducted with four researchers in the field of teacher education. Based on suggestions and recommendations of the experts, items found to have a double-meaning, or to be ambiguous, complex, or redundant were revised or removed. At this stage, the number of scale items decreased from 75 to 70.

Participants

A total of three unique sets of samples were reached in the study. Two sample sets of participants were employed for K–12 teachers online teaching readiness scale development. The first sample included 1,081 K–12 teachers working in Izmir, Turkey. The second sample consisted of 1,086 K–12 teachers working in Izmir, Turkey. The literature suggests at least 300 participants are sufficient for EFA (Field, 2009; Tabachnick & Fidell, 2007). The demographic characteristics of these participants is presented in Table 1. In addition, a third set of participants was used in the relational study. This consisted of 1,128 K–12 teachers working in Izmir, Turkey. The demographic information regarding these participants is also presented in Table 1.

A questionnaire was sent to the K–12 teachers in the six districts of Izmir. Teachers were informed about voluntary participation, given a brief explanation of the purpose of the study, and told that they could withdraw from the study at any time. It took approximately 10–15 minutes for participants to complete the scale.

Table 1

Demographic Characteristics of Participants

Characteristics	Sample 1		Sample 2		Sample 3		Full Sample	
	n	%	*n*	%	*n*	%	*n*	%
Gender								
Female	807	74,7	801	73,8	820	72,7	2428	73,7
Male	274	25,3	285	26,2	308	27,3	867	26,3
Years of service								
1-5	80	7,4	73	6,7	65	5,8	218	6,6
6-10	128	11,8	163	15,0	147	13	438	13,3
11-15	188	17,4	202	18,6	219	19,4	609	18,5
16-20	186	17,2	201	18,5	193	17,1	580	17,6
21 and above	499	46,2	447	41,2	504	44,7	1450	44,0
Level of service								
Pre-school	67	6,2	67	6,2	78	6,9	212	6,4
Primary school	351	32,5	347	32,0	387	34,3	1085	32,9
Secondary School	371	34,3	403	37,1	394	34,9	1168	35,4
High school	292	27,0	269	24,8	269	23,8	830	25,2
Being Technology Literate [a]	913	84,5	926	85,3	962	85,3	2801	85,0
Daily Technological Devices Usage time								
Under 1 hr	19	1,8	21	1,9	27	2,4	67	2,0
1-3 hr	150	13,9	182	16,8	188	16,7	520	15,8
3-5 hr	336	31,1	320	29,5	337	29,9	993	30,1
More than 5 hr	576	53,3	563	51,8	576	51,1	1715	52,0

Note. N = 3295 (n = 1081 for Sample 1; n = 1086 for Sample 2; n = 1128 for Sample 3); Sample 1 = for EFA analysis; Sample 2 = for CFA analysis; Sample 3 = for Implementation analysis; [a] Reflects the number and percentage of participants answering "yes" to this question.

Data Analysis

In order to test the psychometric properties of the K–12 teachers' online teaching readiness scale, first EFA then CFA-SEM was performed. Next, analyses were conducted to evaluate the validity and reliability of the scale. With the help of the first sample, EFA was conducted. With the help of sample two, confirmatory factor analysis was performed to verify the factors occurring in the EFA. SPSS 18.0 was used to perform the EFA and reliability analysis, and AMOS 21.0 was used for the CFA and SEM.

Results

The Process of Determining the Number of Items and Factors

Stage One

EFA and CFA assumptions were provided. Principle components of extraction method in EFA analysis and Direct Oblimin for rotation method were used. After the first EFA, item 24 was removed from the scale as it did not fall under any factor and the EFA was repeated. Item 44 was then removed from the scale as it did not fall under any factor and the EFA was repeated again. Several cross-loading problems were observed. Item 25 was removed from the scale as it was included in two factors with close values (.324 and -.352) and the EFA was repeated. Item 67 was then removed from the scale as this question was included in two factors with close values (.41 and -.341), and the EFA was repeated. Next, item 23 was removed from the scale as it was included in two factors with close values (-.373 and -.327) and the EFA was repeated. Item 43 was removed from the scale as the question did not fall under any factor and the EFA was again repeated. The load of item 26 was low (.33). Expert opinion for item 26 was that the item was not suitable for the relevant factor so it was removed from the scale and the EFA was repeated. At the end of this first round, 12 factors including 63 items were formed, and the CFA was made.

After the first CFA, item 19 was removed from the scale because its load (.38) was below .50 and the EFA was repeated. Item 21 was then removed from the scale as it was included in two factors with close values (.403 and -.303), and the EFA was repeated. Item 2 was removed because it passed to another factor with a low value (-.331), and it was not deemed appropriate by the researchers for it to be in that factor. EFA and CFA were then repeated. Covariance was created between the error terms of the items suitable to improve the CFA values. The covariances generated were as follows: 1-3, 13-14, 55-56, 48-49, 33-34, 36-37, 5-6, 29-30, 27-29, and 32-34. Then, reliability analyses were performed. AVE, CR, and CA values were checked for reliability. In light of Table 2, item 35 was removed from the scale, and the CFA was repeated.

Table 2

AVE Reliability for Factor 10

Item no.	AVE	AVE if item removed
66		.27
19	.37	.38
35		.45[a]

Note. AVE = average variance extracted; [a] It's considerably high value.

Item 19 was removed from the scale because its estimate value (.46) was below .50, and the EFA was repeated. Item 66 passed to factor 9, and factor 10 was removed. Then, CFA was performed again. In light of Table 3, item 36 was removed from the scale, and the CFA was repeated. After all these processes, 10 factors and 55 items were included in the scale.

Table 3

AVE Reliability for Factor 12

Item no.	AVE	AVE if item removed
30		.42
31		.40
32		.40
33		.41
34	.40	.40
27		.41
28		.38
29		.40
36		.43[a]
37		.40

Note. AVE = average variance extracted; [a] It's considerably high value.

Stage Two

In our EFA, in deciding on the number of factors, analytical techniques such as parallel analysis, scree plot, and contributions to variance were used (Field, 2009; Pallant, 2007). Since the EFA eigenvalue of the 10th factor (1.02) was lower than the parallel analysis eigenvalue (1.24), the factor number was reduced to nine and the EFA was repeated. Item 33 was removed because it passed under another factor with a low value (.368) and it was not deemed appropriate by the researchers. Then, CFA was repeated. Item 66 was removed from the scale because its load (.43) was below 0.50, and the EFA was repeated. As a result of the EFA, no change was required and CFA was made. Item 70 was removed from the scale because its load (.44) was below 0.50, and the EFA was repeated. As a result of the EFA, no change was required and CFA was made. No change was deemed necessary in the CFA results, and reliability analyses were performed.

Final Results

EFA Results

EFA was carried out to determine the factor structure. The number of people required for EFA according to the literature is at least 300, and the first version of the scale was applied to 1,081 people in our study (Field, 2009). The Kaiser-Meyer-Olkin (KMO) value was checked in order to determine whether the sample size was suitable for performing EFA. The KMO value was calculated as .94, which is higher than .50 and therefore appropriate for EFA (Kaiser, 1974).

Principal component analysis was employed as the extraction method to determine the factor structure of the scale, and the Oblimin with Kaiser Normalization method was applied. The cut-off point of the items' factor loadings was accepted as .30 (Izquierdo et al., 2014). EFA eigenvalues, parallel analysis eigenvalues, and explained variances can be seen in Table 4.

Table 4

Total Variance Explained

Component	EFA (PCA) eigenvalues	Parallel analysis eigenvalues	% of variance	Cumulative %
One	15.43	1.46	28.05	28.05
Two	5.61	1.42	10.20	38.25
Three	4.63	1.39	8.41	46.66
Four	2.38	1.37	4.33	50.99
Five	1.99	1.34	3.62	54.61
Six	1.83	1.32	3.33	57.94
Seven	1.62	1.30	2.94	60.88
Eight	1.46	1.28	2.66	63.53
Nine	1.31	1.26	2.39	65.92

Note. EFA = exploratory factor analysis; PCA = principal components analysis.

CFA Results

Since the data set has a normal distribution, the maximum likelihood method as parameter estimation method and covariance matrix method as the data matrix were employed. All of the t values of the items were higher than +1.96, and the t values of the indicators should differ from +-1.96, according to the literature (Kline, 2011). In addition, the error variance was less than .90, which is not high. All values were significant ($p < .05$). The path diagram is presented in Figure 1.

The *p* value of the $\chi 2$ value was examined considering the fit indices of the model. As this value is .00 ($p < .05$), it was accepted as a good fit. Since this value is likely to be meaningful in large sample sizes, a ratio of $\chi 2/df$ and other indices should be evaluated (Tabachnick & Fidell, 2007). The $\chi 2$ value was 4112.46 (df = 1231). In this context, the ratio of $\chi 2/df$ (4112.46/1231) was calculated as 3.34. Since this value was less than five, it is acceptable (Wheaton et al., 1977). Other fit indices are presented in Table 5 and examined in terms of the literature. All indices were found to be either a perfect or good fit, and only two were acceptable. In this way, the model was verified to have nine factors.

Table 5

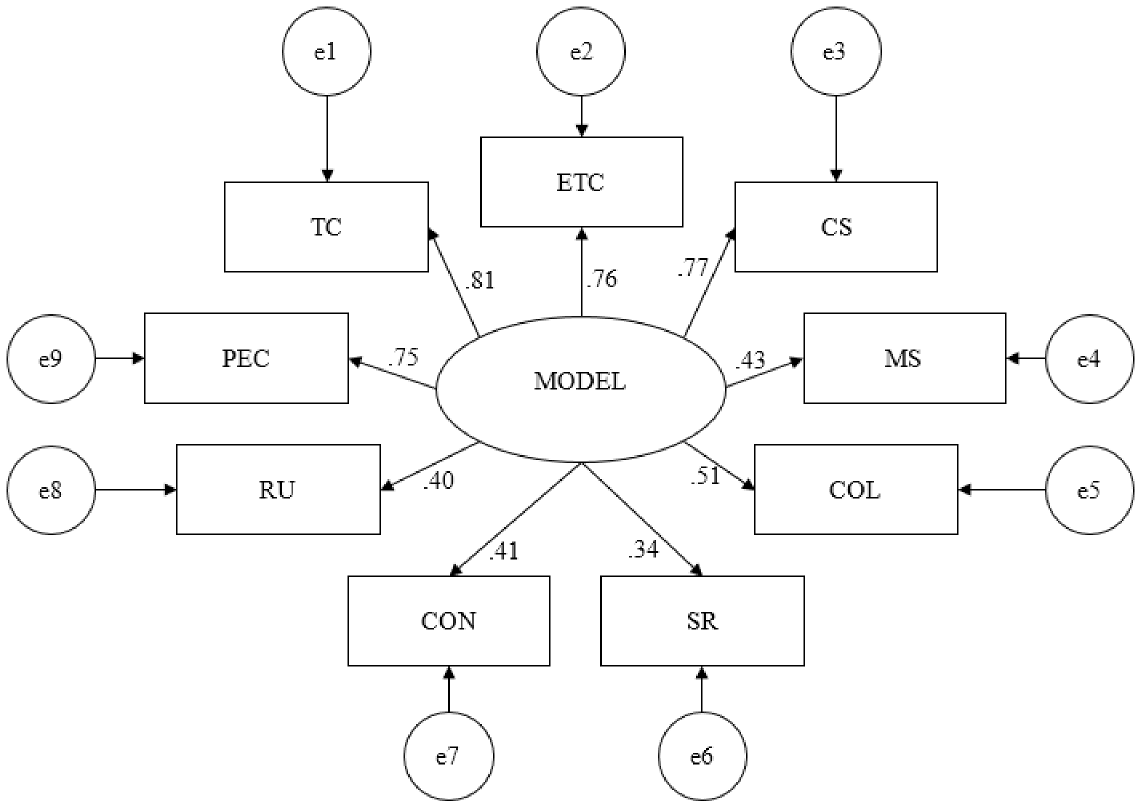
Model Fit Measurements

Model Fit Statistics / Indices	Model	Criteria	Decision	Rationale
$\chi 2$	4112.46			
df	1231			
$\chi 2/df$	3.34	< 5	Acceptable	Wheaton et al. (1977)
TLI	.92	≥ .92	Good fit	Tabachnick and Fidell (2007)
NFI	.90	≥ .90	Acceptable	Tabachnick and Fidell (2007)
CFI	.93	≥ .90	Good fit	Tabachnick and Fidell (2007)
RMSEA	.05	< .05	Perfect fit	Hooper et al. (2008)
SRMR	.06	≤ .08	Good fit	Hu and Bentler (1999)
RMR	.05	≤ .05	Perfect fit	Hu and Bentler (1999)
AGFI	.84	≥ .85	Acceptable	Schermelleh-Engel et al. (2003)
IFI	.93	≥ .90	Good fit	Baumgartner and Homburg (1996)
PNFI	.83	≥ .50	Good fit	Hu and Bentler (1999)

Note. TLI = Turker-Lewis index; NFI = normed fit index; CFI = the comparative fit index; RMSEA = the root mean square error of approximation; SRMR = standardized root mean square residual; RMR = root mean square residuals; AGFI = adjusted goodness-of-fit index; IFI = the incremental fit index; PNFI = parsimony normed fit index.

Figure 1

Confirmatory Factor Analysis Path Diagram



Note. This figure shows the factor loadings. TC = technical competence; ETC = educational technology competence; CS = computer self-efficacy; MS = management support; COL = colleague; SR = student readiness; CON = content; RU = relative usefulness; PEC = pedagogical and ethical competence.

Reliability Results

The Cronbach alpha internal consistency coefficient ranged from .82 to .92. The AVE value of pedagogical and ethical competence factor was below .50, but it is acceptable because CR and CA were high (Table 6). After all these processes, nine factors and 52 items were included in the scale.

Table 6

AVE, CR, and CA Scores of the Factors

Factor	AVE	CR	CA
Technical competence	.61	.93	.92
Educational technology competence	.59	.90	.90
Computer self-efficacy	.60	.88	.89

Management support	.61	.90	.90
Colleague	.75	.92	.91
Student readiness	.52	.81	.83
Content	.60	.88	.88
Relative usefulness	.51	.82	.84
Pedagogical and ethical competence	.44	.86	.86

Note. AVE = average variance extracted; CR = composite reliability; CA = Cronbach alpha.

Implementation

As can be seen in Table 7, the total score average was 64.12 and showed a normal distribution. Since there was no normal distribution in the preschool group in the level variable, it was not included in the analysis of variance (ANOVA). Since the scores of those who answered "no" to the question "Are you technology literate?" also did not show a normal distribution, a comparison analysis was not conducted according to this variable. Comparison and correlation analyses were performed for other variables. *T*-tests for gender, correlation for years of service and daily device usage, and ANOVA analyses were performed for the level.

Table 7

Descriptive Statistics (Standardized Scores Between 0–100)

Characteristic	*n*	*M*	*SD*	Skewness	Kurtosis
TOTAL	1,128	64.12	12.11	-0.30	1.17
Gender					
Female	820	64.35	11.25	-0.13	0.93
Male	308	63.49	14.15	-0.46	0.98
Years of service					
1–5	65	68.78	11.12	-0.31	-0.40
6–10	147	65.58	10.98	-0.05	0.57
11–15	219	65.55	10.68	0.21	0.07
16–20	193	65.30	11.25	-0.22	0.26
21 and above	504	62.01	13.07	-0.35	1.53
Level					
Pre-school	78	63.82	12.89	-1.29	5.32
Primary school	387	63.87	11.99	-0.17	0.67
Secondary school	394	64.76	11.65	0.13	0.04
High school	269	63.61	12.72	-0.61	1.62
Technology literate					

No	166	54.99	12.74	-0.72	2.34
Yes	962	65.69	11.28	-0.05	0.44
Daily duration					
Under 1 hr	27	53.24	16.83	-0.80	1.76
1–3 hr	188	61.20	11.71	-0.17	0.05
3–5 hr	337	64.43	11.63	0.02	0.48
More than 5 hr	576	65.40	11.87	-0.30	1.38

Gender Comparisons

Descriptive statistics are presented in Table 8. There was no significant difference in total scores. There was a significant difference in favor of females in the colleague, content, and pedagogical and ethical competence factors and in favor of males in the technical competence factor. There was no significant difference in other factors (Table 9).

Table 8

Gender Differences in Scores

Factor and Gender	n	*M*	*SD*
TOTAL			
Female	820	64.35	11.25
Male	308	63.49	14.15
Technical competence			
Female	820	13.07	2.82
Male	308	13.52	3.30
Educational technology competence			
Female	820	5.84	2.41
Male	308	6.02	2.79
Computer self-efficacy			
Female	820	5.86	1.82
Male	308	6.10	2.14
Management support			
Female	820	8.45	1.91
Male	308	8.20	2.26
Colleague			
Female	820	6.19	1.17
Male	308	5.98	1.40
Student readiness			
Female	820	3.36	1.61
Male	308	3.35	1.76
Content			
Female	820	5.23	1.94

Male	308	4.92	2.15
Relative usefulness			
Female	820	4.32	1.84
Male	308	4.06	2.05
Pedagogical and ethical competence			
Female	820	12.02	1.80
Male	308	11.35	2.48

Table 9

T Tests by Gender

Factor	*t*	*df*	*d*
TOTAL	0.96	460.24	
Technical competence	-2.12*	485.78	0.15
Educational technology competence	-0.97	488.51	
Computer self-efficacy	-1.73	481.69	
Management support	1.77	480.21	
Colleague	2.36*	477.43	0.16
Student readiness	0.12	1126.00	
Content	2.33*	1126.00	0.15
Relative usefulness	1.98	503.80	
Pedagogical and ethical competence	4.30**	434.72	0.31

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

Correlation Between Scores and Years of Service

The correlation between years of service and scores has been examined. Apart from the relative usefulness (RU) factor, a significant and negative correlation was found between the other factor scores and the total score (Table 10). This suggests that younger teachers are more ready to teach online.

Table 10

Correlations Between Scores and Years of Service

	TOTAL	TC	ETC	CS	MS	COL	SR	CON	RU	PEC
Years of service	-.15**	-.20**	-.18**	-.07*	-.11**	-.09**	.06*	-.09**	-.05	-.10**

Note. TC = technical competence; ETC = educational technology competence; CS = computer self-efficacy; MS = management support; COL = colleague; SR = student readiness; CON = content; RU = relative usefulness; PEC = pedagogical and ethical competence.

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

School Level Comparisons

Preschool level is not included in ANOVA because it did not show normal distribution (Table 7). Homogeneity of variances assumption was met except management support factor (Table 12), therefore ANOVA couldn't be conducted for it. According to the Welch test results F=2.58 (p>0.05), there was no significant difference among the groups for this factor.

Table 11

School Level Differences in Scores

Factor and Level	n	*M*	*SD*
TOTAL			
PS	387	63.87	12.00
SS	394	64.76	11.65
HS	269	63.61	12.72
Total	1050	64.14	12.06
Technical competence			
PS	387	12.99	2.95
SS	394	13.34	2.74
HS	269	13.48	3.18
Educational technology competence			
PS	387	5.76	2.44
SS	394	6.06	2.52
HS	269	5.92	2.64
Computer self-efficacy			
PS	387	5.77	1.88
SS	394	5.91	1.90
HS	269	6.19	1.95
Management support			
PS	387	8.32	2.06
SS	394	8.58	1.83
HS	269	8.16	2.21
Colleague			
PS	387	6.17	1.15
SS	394	6.14	1.21
HS	269	6.02	1.36
Student readiness			
PS	387	3.62	1.66
SS	394	3.30	1.66
HS	269	2.97	1.55
Content			
PS	387	5.05	2.06
SS	394	5.19	2.04
HS	269	5.12	1.91

Relative usefulness			
PS	387	4.32	1.86
SS	394	4.32	1.94
HS	269	4.10	1.95
Pedagogical and ethical competence			
PS	387	11.89	1.97
SS	394	11.92	1.88
HS	269	11.65	2.21

Note. PS = primary school; SS = secondary school; HS = high school.

Table 12

Tests of Homogeneity of Variances

Factor	Levene Statistic	*df*1	*df*2
TOTAL	0.30	2	1047
Technical competence	2.29	2	1047
Educational technology competence	0.75	2	1047
Computer self-efficacy	0.34	2	1047
Management support	5.16*	2	1047
Colleague	1.06	2	1047
Student readiness	0.57	2	1047
Content	0.88	2	1047
Relative usefulness	1.20	2	1047
Pedagogical and ethical competence	1.56	2	1047

*$p < .05$.

Considering the levels, there was no significant difference between the total score averages. Considering the factor score averages, significant differences were found only for the computer self-efficacy and student readiness factors (Table 13). Post-hoc tests were conducted for these factors.

Table 13

Analysis of Variance Statistics by Teaching Level

Factor and groups	*SS*	*df*	*MS*	*F*	η2
TOTAL					
Between groups	256.93	2	128.46	0.88	
Within groups	152217.16	1047	145.38		
Total	152474.09	1049			
Technical competence					

Between groups	44.93	2	22.46	2.60	
Within groups	9037.75	1047	8.63		
Total	9082.68	1049			
Educational technology competence					
Between groups	18.20	2	9.10	1.43	
Within groups	6646.12	1047	6.35		
Total	6664.32	1049			
Computer self-efficacy					
Between groups	28.28	2	14.14	3.88*	.01
Within groups	3815.32	1047	3.64		
Total	3843.61	1049			
Colleague					
Between groups	3.85	2	1.92	1.28	
Within groups	1578.66	1047	1.51		
Total	1582.51	1049			
Student readiness					
Between groups	66.35	2	33.18	12.43**	.02
Within groups	2794.80	1047	2.67		
Total	2861.16	1049			
Content					
Between groups	3.74	2	1.87	0.46	
Within groups	4249.10	1047	4.06		
Total	4252.84	1049			
Relative usefulness					
Between groups	9.70	2	4.85	1.32	
Within groups	3840.33	1047	3.67		
Total	3850.04	1049			
Pedagogical and ethical competence					
Between groups	13.33	2	6.67	1.66	
Within groups	4208.11	1047	4.02		
Total	4221.44	1049			

*$p < .05$. **$p < .01$.

Considering the post-hoc tests, there was a significant difference between primary school and high school for the computer self-efficacy factor in favor of high school. There was a significant difference between all groups for the student readiness factor. The score for primary school is highest, the second highest score is for secondary school, and the lowest score is for high school (Table 14 and Table 15).

Table 14

Post-Hoc (Bonferroni) Results

Dependent variable	(I) level	(J) level	MD
Computer self-efficacy	PS	SS	-0.14
		HS	-0.42*
	SS	PS	0.14
		HS	-0.28
	HS	PS	0.42*
		SS	0.28
Student readiness	PS	SS	0.32*
		HS	0.64**
	SS	PS	-0.32*
		HS	0.32*
	HS	PS	-0.64**
		SS	-0.32*

Note. PS = primary school; SS = secondary school; HS = high school.

* $p < .05$. ** $p < .01$.

Correlation Between Scores and Daily Device Usage Time

As can be seen from Table 15, a low-level and positive relationship was found between the total score and the factors of technical competence, educational technology competence , computer self-efficacy, relative usefulness, and pedagogical and ethical competence according to the duration of daily technological device use. As the usage time increased, the rate of readiness increased.

Table 15

Correlation Between Scores and Daily Device Usage Time

	TOTAL	TC	ETC	CS	MS	COL	SR	CON	RU	PEC
Daily device usage time	.17**	.20**	.17**	.16**	0.04	0.03	0.03	0.02	0.07*	.16**

Note. TC = technical competence; ETC = educational technology competence; CS = computer self-efficacy; MS = management support; COL = colleague; SR = student readiness; CON = content; RU = relative usefulness; PEC = pedagogical and ethical competence.

*$p < .05$. ** $p < .01$.

Discussion

Scale Development

The scales for e-learning readiness show the necessity of examining the e-learning readiness of teachers within a more comprehensive structure. There are some studies on e-learning readiness in the literature. For example, Aydın and Tascı (2005) stated that human resource readiness is an important factor in e-learning effectiveness. Other studies have focused on technological readiness and organizational readiness.

According to the technical competence factor, the nine items here are largely related to readiness. The factor load was .81 (Figure 1). It was observed that items on computer and Internet use skills in general are particularly important in terms of readiness (items three, five, six, and seven). One of the most vital factors affecting e-learning outcomes is technical competence (Eslaminejad et al., 2010; Gay, 2016; Keramati et al., 2011; Ouma, 2013). Yun and Murad (2006) also stated that one of the most important barriers preventing the readiness of secondary school teachers is the lack of technical skills.

The educational technology competence factor, which consisted of six items, had a load of .76 (Figure 1), and this factor is important in terms of readiness. The items of online collaboration and using online exam/quiz tools were particularly important in terms of readiness (items 12 and 16). Educators' technical skills are an important factor; if the e-learning participants do not have the necessary skills to use the technology and learn the content, the e-learning process will not be successful (Berge et al., 2000). When it comes to designing e-learning environments and using tools for this environment, teachers need to have educational technology competence (Eslaminejad et al., 2010).

Computer self-efficacy, which had five items, had a load of .77 (Figure 1), and this factor is important in terms of readiness. It has been observed that providing technical support, giving sufficient time, and providing practical training are especially important in terms of readiness (items 17, 18, and 19). Agboola (2006) pointed out that education was an important factor in supporting the e-learning readiness of staff in a positive way. It is therefore important to determine the computer self-efficacy of educators (Hung et al., 2010). Having basic computer skills and a high computer self-efficacy perception is one of the factors affecting students' success in online learning environments (Çelen et al., 2011). Gay (2016) revealed that instructors need an online help desk for technical support. Giving sufficient time is also mentioned in the literature. Martin, Budhrani, et al. (2019) stated that less experienced lecturers in particular may need extra time not only to prepare for online learning but also to acquire the skills necessary for online learning, and therefore may experience time constraints. Online education is more time consuming and teachers should be particularly prepared in terms of time management. The subject of time management is of particular concern to female teachers (Martin, Budhrani, et al., 2019).

The load of the colleagues factor was .51 (Figure 1), and it is moderately important for readiness. Cooperation and help between colleagues were important in terms of readiness (items 28 and 29). Yun and Murad (2006) revealed that colleagues sharing knowledge and technical skills regarding e-learning has a positive effect on teachers' readiness.

It can be said that management support, student readiness, content, and relative usefulness factors were all lower than .50 (Figure 1), and readiness was less related here than in other factors. In the management support factor, preparation of the curriculum with consideration of teachers' and students' needs was an item that had a greater impact on readiness when compared to other factors (item 23). Barefoot (2004) found that institutional support is a crucial element in learning persistence. According to research, institutional elements such as technological support, pedagogical assistance, and the school vision for the adoption of online or blended learning can all have an impact on the effectiveness of online teaching (Almpanis, 2015; Bao, 2020). According to Howard et al. (2021), strong leadership and unambiguous support for incorporating new technology and practices in teaching and learning can inspire teachers to change, but a lack of organizational commitment to change can demotivate teachers and impede change.

In the student readiness factor, both high interest in online learning and a high use of their time are the most important items (items 33 and 34). In the content factor, adequate technical support of the Ministry of National Education in Turkey and in-service training for teachers were found to be more important than other items (items 37 and 38). So and Swatman (2006) found that student readiness, teacher readiness, technological support, managerial support, and school culture are factors that affect the e-learning process.

Finally, in the relative usefulness factor, teachers' readiness was higher if they thought that online education is efficient and effective (items 42 and 43). According to Engholm and McLean (2002), organizational support and training and development are key factors in a successful e-learning process. Akaslan and Law (2011) also emphasized managerial support.

The load of the pedagogical (professional) and ethical competency factor was .75 (Figure 1), and this factor is important for readiness. In particular, the ability to communicate with colleagues online has been found to be important in terms of readiness (item 50). Pedagogical readiness is also related to computer skills (Eslaminejad et al., 2010) in the case of, for example, designing online material. Teaching in e-learning environments requires skills, and replicating the methods and materials of face-to-face classroom settings is not an adequate substitute (Mercado, 2008).

Implementation

Gender Comparisons

According to the mean scores by gender, there is no significant difference in the total score. This result contrasts with the results of instructors at the university level found in previous research. For example, the average score of females in e-learning readiness has been shown to be lower than that of males (Akaslan & Law, 2011).

A significant difference was found in favor of males in the technical competence factor and in favor of females in the colleague, content, and pedagogical and ethical competence factors. The higher the technical competence factor score of males can be attributed to their high interest in using technology, computers, and Internet tools. This finding is supported by research by Sáinz and López-Sáez (2010). Female teachers feel less comfortable and less secure when it comes to technical matters (Correa, 2010). Similarly, So and Swatman (2006) found that male teachers have high confidence in IT proficiency and feel ready to learn, despite receiving the same IT training. According to a study by Çınar et al. (2021), males have a higher level of e-learning readiness in K–12 schools. Considering the factors in which females ranked higher, these are related to professional-ethical competencies and sociality. In parallel with this, So and Swatman (2006) revealed a significant difference in favor of female teachers in aspects such as teamwork and sharing. Supporting this finding, female instructors have perceived communication and technical competencies as more important than have male instructors (Martin, Wang, et al., 2019).

Martin, Budhrani, et al. (2019) revealed that female instructors' perceptions of course design, communication, and time management were significantly higher than male instructors'. Differences in male and female communication styles affect the way faculties communicate online. Time management is also a more crucial concern for women, especially among those with family.

Correlation Between Scores and Years of Service

Considering the relationship between years of service and readiness, there was a negative and significant relationship with all factors except relative usefulness. The readiness of younger teachers was generally higher. This situation mostly affects the technical competence and educational technology competence factors. This may have been caused by the interest and knowledge of teachers in technology and educational technology, especially during their initial years in the profession. Le et al. (2014) stated that older teachers may take time to get used to technologies such as the LMS when compared to younger teachers who are more familiar with technology.

School Level Comparisons

There was no significant difference in the total score according to educational level which were primary, secondary, and high schools. However, on a factor basis, there was only a significant difference in the factors of computer self-efficacy and student readiness according to educational level. In the computer self-efficacy factor, high school teachers had a significantly higher level of readiness than primary school teachers. The reason for this may be that high school teachers are required to use technology more in their lessons/administrative activities. Similarly, So and Swatman (2006) revealed that although the opportunities and training offered to primary and secondary school teachers are the same, primary school teachers think they know less about e-learning than secondary school teachers.

A surprising result is that primary school teachers have higher student readiness than secondary and high school teachers. The reason for this may be that primary school students use technology less in daily life and the innovation effect that technology has. In contrast, So and Swatman (2006) found that primary school teachers think their students are not ready.

Correlation Between Scores and Daily Device Usage Time

There was a low correlation between readiness and daily technological device use. In particular, technological competence, educational technology competence, computer self-efficacy, relative usefulness, and pedagogical competence and readiness were relevant. As the use of devices increased, technology-related readiness in particular increased. As teachers use devices, their familiarity increases so they feel more prepared to use different technologies. In line with this, Phan and Dang (2017) highlighted that in order for teachers to be facilitators in the online learning process, teachers should use learning management systems (LMS), live conference systems, etc., which have to be made available by administrations.

Conclusion

It is vital for teachers to be prepared for online teaching. This study examined teacher preparedness. The purpose of this study was twofold: to develop a scale to measure K–12 teachers' e-learning readiness, and to examine their readiness to teach online. The e-learning readiness scale developed in this study and analyzed for validity and reliability is expected to guide and support future studies on these issues. The measurement of readiness of teachers for the e-learning process plays an important role in terms of determining future strategies, measures, and interventions that need to be taken. Every variable that increases the quality of e-learning is of great importance scientifically for increasing the quality of the process and student achievement.

Limitations and Suggestions

The study included a large sample size. Also, teachers from different districts were included, exhibiting different characteristics and cultures. However, it is a limitation that teachers from different geographical regions were not included. Future research should examine the readiness of teachers in various geographic regions. Studies can be carried out by collecting data from different schools, and the findings can be compared with the findings of this research. In future studies, mobile learning readiness can also be investigated, examining the increase in teachers' use of smart phones. Based on the results of the research, qualitative studies can also be carried out in order to obtain deeper data. Research could look at the effect of teachers' e-learning readiness level on teachers' online teaching performance. Studies could also be conducted on how administrators can encourage teachers in the online learning process, involve teachers in the online learning process, and motivate them.

It is important to evaluate the readiness of teachers and design training accordingly. Teachers take on tasks such as preparing online content, applying methods for online learning, and motivating students through interaction. In order to ensure readiness, teachers should be equipped with these competencies. Training on how to use applications and application tools such as LMSs and other tools suitable for their courses will be useful. Courses should not focus solely on technical skills but should also include topics such as educational technology competencies, effective use of time, how to engage and motivate students, and methodologies to be used in e-learning environments. Continuous training is recommended rather than one-off training. Furthermore, it is not only in-service training that is important. Pre-service training in the content of courses given in education faculties should be enriched, and the e-learning readiness of pre-service teachers should be increased by giving practical lessons.

Since the role of school administrators as technology leaders is also very important, they should also be given theoretical and practical training. Administrators should be guided on how to support and encourage teachers in the e-learning process.

Support from technical personnel is required for better implementation of e-learning. In addition, educational technologists in schools must play a role in the execution of e-learning. Again, infrastructure problems such as Internet speed should be tackled as much as possible by policy makers.

In order to increase the collaboration of teachers with their colleagues, online platforms can be established where teachers can help one another with e-learning. Teachers with more experience can be assigned as e-mentors.

The study has implications for teachers who teach online, for instructional designers who design online learning environments, and for administrators and policy makers who support online learning at K–12. The study can also guide policy makers and educational institutions by shedding light on the dimensions of e-learning readiness that can contribute to the success of e-learning.

References

Agboola, A. K. (2006). Assessing the awareness and perceptions of academic staff in using e-learning tools for instructional delivery in a post-secondary institution: A case study. *The Innovation Journal: The Public Sector Innovation Journal*, *11*(3), 1–12. https://www.innovation.cc/scholarly-style/2006_11_3_4_agboola_islamic-university.pdf

Akaslan, D., & Law, E. L. C. (2011). Measuring teachers' readiness for e-learning in higher education institutions associated with the subject of electricity in Turkey. In *Proceedings of 2011 IEEE Global Engineering Education Conference: Learning environments and ecosystems in engineering education* (pp. 481–490). IEEE. https://doi.org/10.1109/EDUCON.2011.5773180

Al-Furaydi, A. A. (2013). Measuring e-learning readiness among EFL teachers in intermediate public schools in Saudi Arabia. *English Language Teaching*, *6*(7), 110–121. http://doi.org/10.5539/elt.v6n7p110

Almpanis, T. (2015). Staff development and institutional support for technology enhanced learning in UK universities. *The Electronic Journal of E-Learning*, *13*(5), 381–389. https://academic-publishing.org/index.php/ejel/article/view/1940/1903

Amalia, R. R., Anggoro, S., & Eka, K. I. (2021). Identification of teachers and students' readiness to e-learning implementation. *Journal of Teaching and Learning in Elementary Education (JTLEE)*, *4*(2), 170–179. https://jtlee.ejournal.unri.ac.id/index.php/JTLEE/article/download/7884/6822

Aydin, C., & Tasci, D. (2005). Measuring readiness for e-learning: Reflections from an emerging country. *Educational Technology & Society*, *8*(4), 244–257. https://www.jstor.org/stable/jeductechsoci.8.4.244

Bao, W. (2020). COVID-19 and online teaching in higher education: A case study of Peking University. *Human Behavior and Emerging Technologies*, *2*(2), 113–115. https://doi.org/10.1002/hbe2.191

Barefoot, B. O. (2004). Higher education's revolving door: Confronting the problem of student drop out in U.S. colleges and universities. *Open Learning*, *19*(1), 9–18. https://doi.org/10.1080/0268051042000177818

Baumgartner, H., & Homburg, C. (1996). Applications of structural equation modeling in marketing and consumer research: A review. *International Journal of Research in Marketing*, *13*(2), 139–161. https://doi.org/10.1016/0167-8116(95)00038-0

Berge, Z. L., Collins, M., & Dougherty, K. (2000). Design guidelines for web-based courses. In B. Abbey (Ed.), *Instructional and Cognitive Impacts of Web-Based Education* (pp. 32–40). IGI Global. http://doi.org/10.4018/978-1-878289-59-9.ch002

Correa, T. (2010). The participation divide among "online experts": Experience, skills, and psychological factors as predictors of college students' web content creation. *Journal of Computer-Mediated Communication*, *16*(1), 71–92. https://doi.org/10.1111/j.1083-6101.2010.01532.x

Çelen, F. K., Çelik, A., & Seferoğlu, S. S. (2011). Yükseköğretimde çevrimiçi öğrenme: sistemde yaşanan sorunlar ve çözüm önerileri [Education in higher education: Problems experienced and solution suggestions]. *Journal of European Education*, *1*(1), 25–34. http://www.ajindex.com/dosyalar/makale/acarindex-1423906196.pdf

Çiğdem, H., & Topcu, A. (2015). Predictors of instructors' behavioral intention to use learning management system: A Turkish vocational college example. *Computers in Human Behavior*, *52*, 22–28. https://doi.org/10.1016/j.chb.2015.05.049

Çınar, M., Ekici, M., & Demir, Ö. (2021). A snapshot of the readiness for e-learning among in-service teachers prior to the pandemic-related transition to e-learning in Turkey. *Teaching and Teacher Education*, *107*(2021), 1–10. https://doi.org/10.1016/j.tate.2021.103478

Demir, Ö., & Yurdugül, H. (2015). The exploration of models regarding e-learning readiness: Reference model suggestions. *International Journal of Progressive Education*, *11*(1), 173–194. https://ijpe.inased.org/makale_indir/2420

Engholm, P., & McLean, J. J. (2002). What determines an organisation's readiness for e-learning? In A. Green (Ed.), *E-business in Australia: Concepts and cases* (pp. 200 - 212). Pearson Australia Group.

Eslaminejad, T., Masood, M., & Ngah, N. A. (2010). Assessment of instructors' readiness for implementing e-learning in continuing medical education in Iran. *Medical Teacher*, *32*(10), 407–412. https://doi.org/10.3109/0142159X.2010.496006

Field, A. P. (2009). *Discovering statistics using SPSS: And sex and drugs and rock 'n' roll*. Sage.

Gay, G. H. E. (2016). An assessment of online instructor e-learning readiness before, during, and after course delivery. *Journal of Computing in Higher Education*, *28*(2), 199–220. https://doi.org/10.1007/s12528-016-9115-z

Gurley, L. E. (2018). Educators' preparation to teach, perceived teaching presence, and perceived teaching presence behaviors in blended and online learning environments. *Online Learning*, *22*(2), 179–220. https://doi.org/10.24059/olj.v22i2.1255

Hong, Y. & Gardner, L. (2018). An exploration of undergraduates' preparedness and experiences in blended courses. In M. Tanabu & D. Senoo (Eds.), *PACIS 2018 proceedings*. Article 156. Association for Information Systems (AIS). https://aisel.aisnet.org/pacis2018/156

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*(1), 53–60. https://academic-publishing.org/index.php/ejbrm/article/view/1224/1187

Howard, S. K., Tondeur, J., Siddiq, F., & Scherer, S. (2021). Ready, set, go! Profiling teachers' readiness for online teaching in secondary education. *Technology, Pedagogy and Education*, *30*(1), 141–158. https://doi.org/10.1080/1475939X.2020.1839543

Hu, K., AlSaqqaf, A., & Swanto, S. (2020). E-learning readiness among English teachers in Malaysia: A pilot study. *Journal of English Language Teaching Innovations and Materials*, *2*(2), 120–127. http://dx.doi.org/10.26418/jeltim.v2i2.42155

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hung, M.-L., Chou, C., Chen, C.-H., & Own, Z.-Y. (2010). Learner readiness for online learning: Scale development and student perceptions. *Computers & Education*, *55*(3), 1080–1090. https://doi.org/10.1016/j.compedu.2010.05.004

Izquierdo, I., Olea, J., & Abad, F. J. (2014). Exploratory factor analysis in validation studies: Uses and recommendations. *Psicothema*, *26*(3), 395–400. https://doi.org/10.7334/psicothema2013.349

Johnson, S. G., & Berge, Z. (2012). Online education in the community college. *Community College Journal of Research and Practice*, *36*(11), 897–902. https://doi.org/10.1080/10668920903323948

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*, 31–36. https://doi.org/10.1007/BF02291575

Keramati, A., Afshari-Mofrad, M., & Kamrani, A. (2011). The role of readiness factors in e-learning outcomes: An empirical study. *Computers & Education*, *57*(3), 1919–1929. https://doi.org/10.1016/j.compedu.2011.04.005

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Publications.

Koloseni, D., & Mandari, H. (2017). The role of personal traits and learner's perceptions on the adoption of e-learning systems in higher learning institutions. *African Journal of Finance and Management, 26*(1), 61–75. https://www.researchgate.net/publication/319980804_The_Role_of_Personal_Traits_and_Learner%27s_Perceptions_on_the_Adoption_of_E-learning_Systems_in_Higher_Learning_Institutions

Le, H. K., Sriratanaviriyakul, N., Nkhoma, M., & Vo, H. X. (2014). Ho Chi Minh City Open University: IT governance in Vietnamese higher education. *Journal of Information Technology Education: Discussion Cases*, *3*(1), 1–11. https://doi.org/10.28945/2055

Lopes, C. (2007). *Evaluating E-learning readiness in a health sciences higher education institution*. Paper presented at the proceedings of IADIS international conference of E-learning, Portugal.

Martin, F., Budhrani, K., & Wang, C. (2019). Examining faculty perceptions of their readiness to teach online. *Online Learning*, *23*(3), 97–119. https://doi.org/10.24059/olj.v23i3.1555

Martin, F., Wang, C., Jokiaho, A., May, B., & Grübmeyer, S. (2019). Examining faculty readiness to teach online: A comparison of US and German educators. *European Journal of Open, Distance and E-learning*, *22*(1), 53–69. https://doi.org/10.2478/eurodl-2019-0004

Mercado, C. (2008, December 11–12). *Readiness assessment tool for an e-learning environment implementation* [Conference session]. Fifth International Conference on E-Learning for Knowledge-Based Society, Bangkok, Thailand. http://ijcim.th.org/SpecialEditions/v16nSP3/18_fullpaper_cecilia%20Mercado_Revised.pdf

Mirabolghasemi, M., Choshaly, S. H., & Iahad, N. A. (2019). Using the HOT-fit model to predict the determinants of e-learning readiness in higher education: A developing country's perspective. *Education and Information Technologies*, *24*(6), 3555–3576. https://doi.org/10.1007/s10639-019-09945-9

Moftakhari, M. M. (2013). *Evaluating e-learning readiness of faculty of letters of Hacettepe* [Unpublished master's thesis]. Hacettepe University.

Obara, J. K., & Abulokwe, A. C. (2012). Utilization of e-learning for effective teaching of vocational education courses in Nigeria. *International Journal of Research Development (Global Academic Group)*, *7*(1), 1–8. https://www.globalacademicgroup.com/journals/approaches/Utilization%20of%20E-Learning%20for%20Effective%20Teaching%20of%20Vocation.pdf

Oketch, H., Njihia, J., & Wausi, A. (2014). E-learning readiness assessment model in Kenya's higher education institutions: A case study of University of Nairobi. *International Journal of Scientific Knowledge, 5*(6), 29–41. https://profiles.uonbi.ac.ke/njihia/files/elearning_readiness__assessment_model_in_kenya.pdf

Ouma, G. O. (2013). E-learning readiness in public secondary schools in Kenya. *European Journal of Open, Distance and E-learning*, *16*(2), 97–110. https://files.eric.ed.gov/fulltext/EJ1017522.pdf

Pallant, J. (2007). *SPSS survival manual: A step by step guide to data analysis using SPSS for Windows*. Open University Press.

Paturusi, S., Chisaki, Y., & Usagawa, T. (2015). Assessing lecturers' and students' readiness for e-learning: A preliminary study at national university in North Sulawesi Indonesia. *GSTF Journal on Education (JEd), 2*(2), 1–8. https://doi.org/10.5176/2345-7163_2.2.50

Phan, T. T. N., & Dang, L. T. T. (2017). Teacher readiness for online teaching: A critical review. *International Journal Open Distance E-Learning*, *3*(1), 1–16. https://ijodel.com/wp-content/uploads/2017/12/001_Phan_Dang.pdf

Pusparini, K. S., Made Hery Santosa, S. P., & Myartawan, I. P. N. W. (2018). An investigation on teachers' e-learning readiness toward e-learning implementation in public senior high schools. *Journal Pendidikan Bahasa Inggris Undiksha*, *5*(2), 1–10. https://ejournal.undiksha.ac.id/index.php/JPBI/article/download/15160/9263

Sáinz, M., & López-Sáez, M. (2010). Gender differences in computer attitudes and the choice of technology-related occupations in a sample of secondary students in Spain. *Computers & Education*, *54*(2), 578–587. https://doi.org/10.1016/j.compedu.2009.09.007

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online*, *8*(2), 23–74. https://www.researchgate.net/publication/251060246_Evaluating_the_Fit_of_Structural_Equation_Models_Tests_of_Significance_and_Descriptive_Goodness-of-Fit_Measures

Setati, P. P., & Paledi, V. N. (2019). Conceptualised framework for assessing teachers' e-learning readiness in South African rural schools. In N. M. Ochara & J. N. Odhiambo (Eds.), *2019 Open Innovations Conference* (pp. 320–330). IEEE. https://www.doi.org/10.1109/OI.2019.8908204

So, T., & Swatman, P. M. (2006). *E-learning readiness of Hong Kong teachers*. University of South Australia. https://nurhadiw.files.wordpress.com/2010/08/swatman.pdf

Summak, M. S., Bağlıbel, M., & Samancıoğlu, M. (2010). Technology readiness of primary school teachers: A case study in Turkey. *Procedia - Social Behavior Science*, *2*(2), 2671–2675. https://doi.org/10.1016/j.sbspro.2010.03.393

Tabachnick, G. G., & Fidell, L. S. (2007). *Experimental designs using ANOVA*. Duxbury Press.

Texas A&M University. (2021). *Readiness for Online Teaching Survey*. https://www.tamut.edu/Training/downloads/Readiness_Survey.xlsx

Trayek, F. A., Ahmad, T. B. T., Nordin, M. S., Dwikat, M. A., Abulibdeh, E. S. A., Asmar, M., & Sawari, S. S. M. (2016). Underlying structure of e-learning readiness among Palestinian secondary school teachers. *MATEC Web of Conferences: 2016 8th International Conference on Computer and Automation Engineering*, *56*, Article 01011. https://doi.org/10.1051/matecconf/20165601011

University of Toledo. (2021). *Faculty Online Teaching Readiness Survey*. https://www.utdl.edu/lv/assessments/faculty_readiness.html

Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. *Sociological Methodology*, *8*(1), 84–136. https://doi.org/10.2307/270754

Yun, G. P., & Murad, M. W. (2006). Factors influencing psychology and skills of the secondary school teachers' e-learning readiness: A case study in Malacca, Malaysia. In *Current developments in technology-assisted education* (pp. 2135–2140). Formatex. http://library.oum.edu.my/oumlib/sites/default/files/file_attachments/odl-resources/4304/factors-influencing.pdf

Yusof, M. M., Kuljis, J., Papazafeiropoulou, A., & Stergioulas, L. K. (2008). An evaluation framework for health information systems: Human, organization and technology-fit factors (HOT-fit). *International Journal of Medical Informatics*, *77*(6), 386–398. https://doi.org/10.1016/j.ijmedinf.2007.08.011

Zheng, Y., Wang, J., Doll, W., Deng, X., & Williams, M. (2018). The impact of organisational support, technical support, and self-efficacy on faculty perceived benefits of using learning management system. *Behaviour & Information Technology*, *37*(4), 311–319. https://doi.org/10.1080/0144929X.2018.1436590

Ziphorah, R. M. (2014). Information and communication technology integration: Where to start, infrastructure or capacity building? *Procedia—Social and Behavioral Sciences*, *116*, 3649–3658. https://doi.org/10.1016/j.sbspro.2014.01.818

Athabasca University

CC BY

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

Cross Validating a Rubric for Automatic Classification of Cognitive Presence in MOOC Discussions

Yuanyuan Hu, Claire Donald, and Nasser Giacaman
Faculty of Engineering, The University of Auckland

Abstract

As large-scale, sophisticated open and distance learning environments expand in higher education globally, so does the need to support learning at scale in real time. Valid, reliable rubrics of critical discourse are an essential foundation for developing artificial intelligence tools that automatically analyse learning in educator-student dialogue. This article reports on a validation study where discussion transcripts from a target massive open online course (MOOC) were categorised into phases of cognitive presence to cross validate the use of an adapted rubric with a larger dataset and with more coders involved. Our results indicate that the adapted rubric remains stable for categorising the target MOOC discussion transcripts to some extent. However, the proportion of disagreements between the coders increased compared to the previous experimental study with fewer data and coders. The informal writing styles in MOOC discussions, which are not as prevalent in for-credit courses, caused ambiguities for the coders. We also found most of the disagreements appeared at adjacent phases of cognitive presence, especially in the middle phases. The results suggest additional phases may exist adjacent to current categories of cognitive presence when the educational context changes from traditional, smaller-scale courses to MOOCs. Other researchers can use these findings to build automatic analysis applications to support online teaching and learning for broader educational contexts in open and distance learning. We propose refinements to methods of cognitive presence and suggest adaptations to certain elements of the Community of Inquiry (CoI) framework when it is used in the context of MOOCs.

Keywords: cognitive presence, MOOC, text classification, online discussion, content analysis

Introduction

The Problem

In this paper, we argue that the existing empirical inquiries and theoretical frameworks for analysing learning engagement in conventional for-credit university courses are limited when analysing learning engagement in massive open online courses (MOOCs), due to the differences in their educational contexts. MOOCs differ from traditional, smaller-scale online courses in terms of the course design, with shorter course durations, limited direct educator involvement (Kovanović et al., 2018), a wide range of learner profiles, and diverse learner motivations (Alario-Hoyos et al., 2017). During the COVID-19 pandemic, MOOCs gained much more attention in addressing the limitations of remote learning, as they provide learners with a diverse range of educational experiences (Buchem et al., 2020; Cha & So, 2020). MOOC educators require support to monitor and moderate learner progress in these massive audiences, and MOOC learners require responsive, high-quality feedback and remediation from educators. Automatic classification of cognitive presence (to supplement automated feedback which is currently the norm) can provide such pedagogical support. A vital foundation for such a classification system is a reliable theoretical basis.

The asynchronous discussion forum in MOOCs offers a virtual zone for participants to interact mutually through written dialogue. These written conversations, or messages, provide educators and researchers with meaningful insights into learners' critical discourse (i.e., critical thinking, higher-order thinking, and cognitive presence). However, experimental studies are required to validate and improve the methods of critical discourse analysis in online discussion forums for broader educational contexts, such as MOOCs (Amemado & Manca, 2017; Kaul et al., 2018). Garrison et al. (1999, 2001) proposed the community of inquiry (CoI) framework and its coding rubrics to evaluate cognitive presence and two other dimensions in online transcripts by content and textual analysis methods. Over two decades, this framework has been broadly used to assess students' learning and guide learning designs in traditional, online, for-credit, smaller-scale courses (Liu & Yang, 2014; Sadaf & Olesova, 2017). There are still shortcomings for the CoI coding rubrics to reliably assess critical discourse in MOOCs. For example, clear instances of the cognitive presence phases have not been elucidated by Garrison et al. (2001); therefore, researchers have had to revise the rubric each time they used it (Rourke & Kanuka, 2009). Also, the coding rubric was initially developed as a descriptive, qualitative analysis method in smaller-scale courses rather than as a quantitative, inferential procedure (Garrison, 2007). Moreover, online discussion consists of an informal and conversational flow, which is relatively chaotic and does not fit into the coherent patterns in the CoI framework (Xin, 2012).

The Significance

The validation of a cognitive presence rubric in MOOCs is a crucial foundation for developing automatic approaches for real time learning support and remediation. Preparing reliable and valid machine learning data sets is an essential prerequisite for training automatic classifiers (Ullmann, 2019). Analysing the common language patterns in the data sets is beneficial for selecting appropriate machine learning algorithms and predictive features (Mladenić, 2010). Automatic evaluation of learners' cognitive presence in MOOCs can help educators monitor the learners' progress in real time and provide personalised feedback at scale. From a learner's perspective, effective and efficient feedback from both educators and peers encourages high participation in MOOCs, assisting students to achieve their learning goals (Phan et al., 2016).

Purpose and the Research Questions

This study aims to cross validate the use of an adapted coding rubric (Hu et al., 2020) to categorise online discussion transcripts from a target MOOC into phases of cognitive presence. A larger dataset and more coders were involved in examining whether the inter-rater reliability could still reach excellent agreement, as reported in the previous study using less data and fewer coders (Hu et al., 2020). The disagreements between coders were also deeply analysed to gain insights about the feasibility of the cognitive presence phases in the CoI framework. Our main research question was: Is the adapted coding rubric a reliable tool to classify cognitive presence in MOOC discussions? The following sub questions were included to guide the main research question:

> SQ1. What are the inter-rater reliability values when we classify the discussion messages from a target MOOC with more coders and a larger dataset than in the previous study (Hu et al., 2020)?

> SQ2. What is the proportion of disagreements across all cognitive phases between coders, and what are possible causes of the disagreements?

Review of Previous Studies

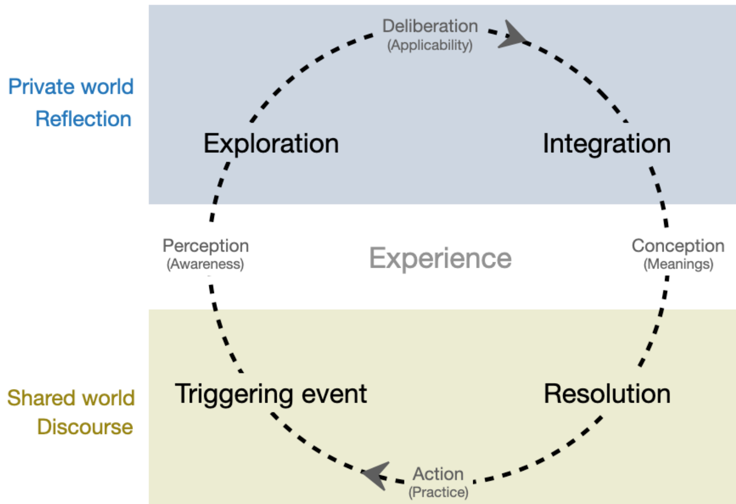
Cognitive Presence

The Community of Inquiry (CoI) framework proposed by Garrison et al. (1999) has been most widely used to analyse learning in online discussions for over two decades. It describes the educational experience that occurs in a learning community, in which "a group of individuals who collaboratively engage in purposeful critical discourse and reflection construct personal meaning and confirm mutual understanding" (Garrison & Anderson, 2011). The CoI framework has three dimensions, called presences: cognitive, social, and teaching presence. Cognitive presence, a primary dimension of the CoI, represents the critical reflection of knowledge (re)construction and problem-solving processes in the learning community (Garrison et al., 2001). Social presence manifests as social communication and emotional interaction between participants, which enriches learning outcomes. Teaching presence, the third dimension, describes the purposeful activities that direct and intervene with the learner's knowledge construction.

This study focused on analysing cognitive presence in MOOC discussion transcripts since the cognitive presence is the "primary issue" of students' learning evidence to be explored before other dimensions (Rourke & Kanuka, 2009). The other two presences of the CoI will be investigated in further research. We adopted the associated four phases of cognitive presence, (a) triggering event, (b) exploration, (c) integration, and (d) resolution from Garrison et al. (1999, 2001). The four phases (Figure 1), called the practical inquiry model, were borrowed from Dewey (1933) who originally described the steps of a complete thought. The definition of the four cognitive phases corresponding to the analysis of discussions in our target MOOC is explained below.

Figure 1

The Practical Inquiry Model Showing the Four Phases of Cognitive Presence in a Learning Community



Note. Adapted from "Critical thinking, cognitive presence, and computer conferencing in distance education," by D. R. Garrison, T. Anderson, and W. Archer, 2001, *American Journal of Distance Education*, *15*, p. 9. Copyright 2001 by University of Chicago Press.

Analysis of Cognitive Presence Phases

Scholars have applied the classification rubric developed by Garrison et al. (2001) and the revised version (Park, 2009) to assess the quality of critical discourse in multidisciplinary online courses. More recent studies have used transcripts data from for-credit, online university courses, but few so far have used transcripts data to analyse cognitive phases in MOOCs. Table 1 summarises the prior research that we have reviewed.

Table 1

Summary of Prior Research Reviewed

Study	Year	Course context	Discipline	Messages *n*	Coders *n*	Agreement %	κ
Garrison et al.	2001	Graduate	Health	24	2	83.33	0.74
Kanuka et al.	2007	Undergraduate	Education	1014	2	-	0.57
Park	2009	Graduate	Nursing	Thematic segments	2	82.34 76.48[a]	-
Liu & Yang	2014	Undergraduate	Information ethics	200	3	90	-

Corich et al.	2006	Undergraduate	Computing systems	484 sentences	2	87	-
McKlin	2004	Undergraduate	History and political science	1300	6	74	0.60
Kovanović et al.	2014	Graduate	Software engineering	1747	2	98.1	0.974
Neto et al.	2018	Undergraduate	Biology	1500	2	91.4	0.86
Kaul et al.	2018	MOOC	Education	78	2	86	0.82
Hu et al.	2020	MOOC	Philosophy	1002	2	95.4	0.93

Note. [a] The Park study included two trials.

Garrison et al. (2001) first proposed their CoI model and reported their manual classification rubric with data from two online graduate-level courses. One course was on workplace learning and the other on health promotion. The former dataset (21 messages) was used to fine tune the measurement rubric, and the latter (24 messages) was used to report the inter-rater reliability result. They reached an agreement of 83.33% and Cohen's κ coefficient (Cohen, 1960) of 0.74 between two coders. It indicated that this classification rubric could be used to evaluate the quality of cognitive presence. However, the sample size was too small (i.e., under 100 messages) and would need further verification with larger, more diverse learner datasets to be sufficiently generalisable.

Two studies made their measurement unit at the message level following Garrison et al.'s (2001) method. Kanuka et al. (2007) used Garrison et al.'s (2001) rubric to analyse online discussion transcripts in an undergraduate education course. They achieved a very low Cohen's κ of 0.57 (1,014 messages) with two coders. The second study (Liu & Yang, 2014) used the rubric in an undergraduate information ethics course, with a percentage agreement of 90% (200 selected from 1,058 messages) reached by three coders. However, Park (2009) deemed that the classification of cognitive phases should be based on the unit of meaning or thematic level rather than the message level. Park (2009) also revised Garrison et al.'s (2001) rubric to use in an online graduate nursing course, reaching agreements of 82.34% and 76.48% in two trials with an unknown number of messages posted by 12 students.

Some studies analysed cognitive phases manually in preparation for developing automated classifiers. Corich et al. (2006) labelled cognitive phases in discussion transcripts from an online undergraduate computing systems course. The correlation between the two coders was 87% (104 messages of 484 sentences) using Garrison et al.'s (2001) rubric. This study was done at the sentence level instead of the message level. McKlin (2004) collected 1,300 learner messages from online undergraduate courses in history and political science. Researchers used 100 of 1,300 messages to report the measure of inter-rater reliability, and the remaining 1,200 messages were coded by six coders (200 messages each). The study reported 74% agreement and an average Cohen's κ of 0.60 (100 messages). Kovanović et al. (2014) classified 1,747 messages from an online graduate course in software engineering research, reaching a very high agreement of 98.1% and Cohen's κ of 0.974 between two coders. This dataset has been reused in four further studies to develop automated classifiers of cognitive presence. To analyse cognitive presence in transcripts written in other languages, Neto et al. (2018) used a Portuguese dataset from an undergraduate biology course, with an agreement of 91.4% (1,500 messages) and Cohen's κ of 0.86 between two coders. This dataset was also used to evaluate the performance of transferring automated classifiers across languages (Barbosa et al., 2020).

Two studies with manual coding approaches of cognitive presence in the CoI focused on MOOCs. Kaul et al. (2018) applied Garrison et al.'s (2010) classification rubric to 78 messages from an education MOOC in the subcategories of all three CoI presences. Agreement between the two coders was 46% and 86% before and after coders' negotiations. An adapted coding rubric (Hu et al., 2020) of cognitive presence was proposed when Garrison et al.'s (2001) classification rubric could not encompass all the cognitive phases in online discussions from a philosophy MOOC. Two expert coders reached a percentage agreement of 95.4% (1,002 messages) and Cohen's κ of 0.93 using the adapted classification rubric. Both studies indicated that most of the coders' disagreements occurred between the exploration and integration phases when classifying cognitive presence phases.

Differences Between the Prior Work and This Study

The cognitive presence analysed in most of the studies we viewed was in the context of smaller-scale, for-credit university courses. Garrison et al.'s (2001) classification rubric preceded MOOCs by seven years. The first MOOC was developed in 2008 (Siemens, 2013). The wide range of learner demographics and diverse learner motivations cause the differences between MOOCs and traditional university courses. The typical MOOC audiences are mature adult learners who are employed and have tertiary qualifications (Dillahunt et al., 2014). Their motivations for learning are updating knowledge, personal curiosity, and upskilling themselves professionally (Alario-Hoyos et al., 2017). These differences may also impact the language they use in online discussions. For instance, students tend to write formally when participating in discussion forums in smaller-scale courses and in professional development MOOCs that are credit-bearing. In contrast, many MOOC learners tend to use a more conversational style of writing when they engage with MOOCs for less formal purposes than professional development or accreditation. We wondered whether the differences in educational contexts would impact the analysis of cognitive presence in discission transcripts.

The reliability reported in most of the reviewed studies was based on two coders. Although six coders were employed in McKlin's (2004) study, they only labelled 100 same messages. Similarly, in Neto et al.'s (2018) study, a third coder was only responsible for resolving disagreements (129 messages) between the other two coders. Although the previous study (Hu et al., 2020) reported excellent inter-rater reliability between two coders, the agreement outcome was reported after the coders' negotiations. We wondered whether the coding rubric of cognitive presence could be reliably applied to analyse MOOC discussions, when we enlarged the dataset, invited more coders to become involved, and reported the outcomes before coders' negotiations.

Methods

Data Description

The MOOC discussion data used in our study was from an archived run of the Logical and Critical Thinking (LCT) MOOC on the FutureLearn platform (University of Auckland, n.d.). It was an introductory undergraduate philosophy course designed and taught by a course design team at our university. This course taught basic concepts in logical and critical thinking (e.g., premises, arguments, etc.), linking those concepts with life experiences. The average number of enrolled users was approximately 11,000, and the discussion transcripts (comprising posts and their replies) included approximately 12,000 messages per course run. There were eight weekly topics with learning tasks in each course run. Firstly, sixteen tasks (two for each week) were evenly and randomly selected. Then, a sample of approximately 100 to 200 messages was randomly selected from each of the 16 tasks. We kept the entire sequential structure of each selected conversation instead of segmenting them to achieve an exact number. Totally, 1,917 messages were selected for this study.

The three coders were postgraduate students from the Philosophy Department, who were also the teaching assistants for the LCT MOOC. They were trained round by round (50 new messages for each round) before reaching agreements over 80% independently without negotiations. They reached an 81% agreement in the third round, so they were allocated to classify the 1,917 messages manually and independently based on the adapted rubric (Hu et al., 2020) The overall study proposal received ethical approval from the University Human Participants Ethics Committee.

Definition of Cognitive Presence Phases in This Research

The five categories of cognitive presence, including four processing phases and the "other" phase, are listed in Table 2. We provide a brief definition and a message example from the LCT MOOC for each category. These definitions are derived from Garrison et al. (2001), Hu et al. (2020) study, and learner messages in the LCT MOOC, and therefore influenced by the disciplinary context of this course, which is philosophy. More details about the definitions can be found in Hu et al. (2020) study.

Table 2

Definitions with Message Instances of the Five Categories in Cognitive Presence

Phase ID	Cognitive phase	Brief definition	Message example
1	Triggering event	Messages state users' confusion.	*"I do find it difficult to override over 30 years of the normalisation of poorly constructed sentences."*
2	Exploration	Messages provide information about the cause of the confusion but without a coherent conclusion.	*"Both overthinking and underthinking leads you to live in low levels of consciousness. I think that [one of the users] explains very well how to find the spot between the two approaches."*

3	Integration	Messages propose coherent conclusions to improve the situation with sufficient substantiation.	*"I think this counter argument works colloquially but not technically it doesn't follow from the premises that having a job will stop you wanting an iPhone unless you add an implied premise to that effect."*
4	Resolution	Messages apply, test, or argue the previous conclusions, usually as new constructs.	*"Another way to test it would be to see if similar positions eg heads of industry are also held by more left-handed people than statistics would suggest. It would be incredibly difficult to iron out other possible factors..."*
0	Other	Messages that do not fall into any category.	*"Thanks. I start that Mooc in May."*

Classification Process

After they were trained, the three coders used our rubric of cognitive phases to classify the sample data (1,917 messages) independently. The unit of analysis was on the message level since the classification on the theme and sentence level may have ignored the contextual information before and after the segment (i.e., theme or sentence) within an entire message. Multiple phases of cognitive presence sometimes existed in one message simultaneously, for example, when a learner stated, diagnosed, and resolved a question in a single post. Our coders labelled each message with the highest cognitive phase in that message, even when lower phases were also represented. An example message of this is:

> Initially, I too felt the conclusion might be that a revolt was required. However, the letter states that a revolt would be a way to get the council to listen. It's the same as cider vinegar would be a way to get rid of your dogs fleas. Therefore, it's a statement rather than an argument.

The first sentence indicates the learner's difficulty, which can be categorised as triggering event. The second sentence illustrates that the learner provided more information for diagnosing the difficulty, which can be the exploration phase. Then, the last two sentences made an analogy and drew a conclusion supported by the reasons stated previously, which can be labelled integration. In this case, the message was classified into the highest phase, integration, rather than the other two phases. More details about the coding rubric and message examples can be found in the adapted rubric of cognitive presence in MOOCs (Hu et al., 2020).

Results

The Overall Analysis of Cognitive Phases

The overall percentage agreement was 77.15%, where all three coders independently agreed on the labels for 1,479 of the 1,917 messages. The average Fleiss' κ (a statistical measure for categorical ratings between more than two coders) was 0.763, shown in Table 3. Across five categories of cognitive presence, the triggering event phase accounted for the highest agreement (κ = 0.828), followed by the phases "other" and exploration. There was less agreement among the coders on the higher cognitive phases of integration and resolution. The average agreement between coder 1 and coder 2 reached Cohen's κ of 0.842. This result was higher than the other two combinations, where Cohen's κ was 0.704 between coder 2 and coder 3, and 0.744 between coder 1 and coder 3.

Table 3

Inter-Rater Reliability Among Three Coders Across Five Categories

Phase ID	Cognitive phases	Fleiss' κ
0	Other	0.792
1	Triggering event	0.828
2	Exploration	0.776
3	Integration	0.689
4	Resolution	0.667
Average		0.763

Note. N = 1,917 messages.

Table 4 illustrates the proportion of the five cognitive phases in the messages of agreement between coders. The phase of exploration accounted for most of the messages (55.46%), which far surpassed triggering event and integration. The highest phase, resolution (2.43%), and the lowest "other" phase (5.75%) had the smallest proportion of messages.

Table 4

Messages of Agreements Between Three Coders by Cognitive Phases

Phase ID	Cognitive phase	Messages	
		n	%
0	Other	85	5.75
1	Triggering event	279	18.86
2	Exploration	835	56.46
3	Integration	244	16.50
4	Resolution	36	2.43%

Note. N = 1,479 messages.

Disagreements Between Coders

In addition to agreements between coders, the distribution of disagreements (i.e., messages that were labelled differently by the coders) is worth considering. Table 5 describes the proportion of the disagreements between the three coders across different combinations of cognitive phases. Most of the disagreements had two labels rather than three (the latter was less than 1.5% in total). The proportion of inter-rater disagreements (96.13%) between adjacent cognitive phases far surpassed the non-adjacent combinations (i.e., exploration and resolution, and "other" and exploration).

Table 5

Distribution of Disagreements

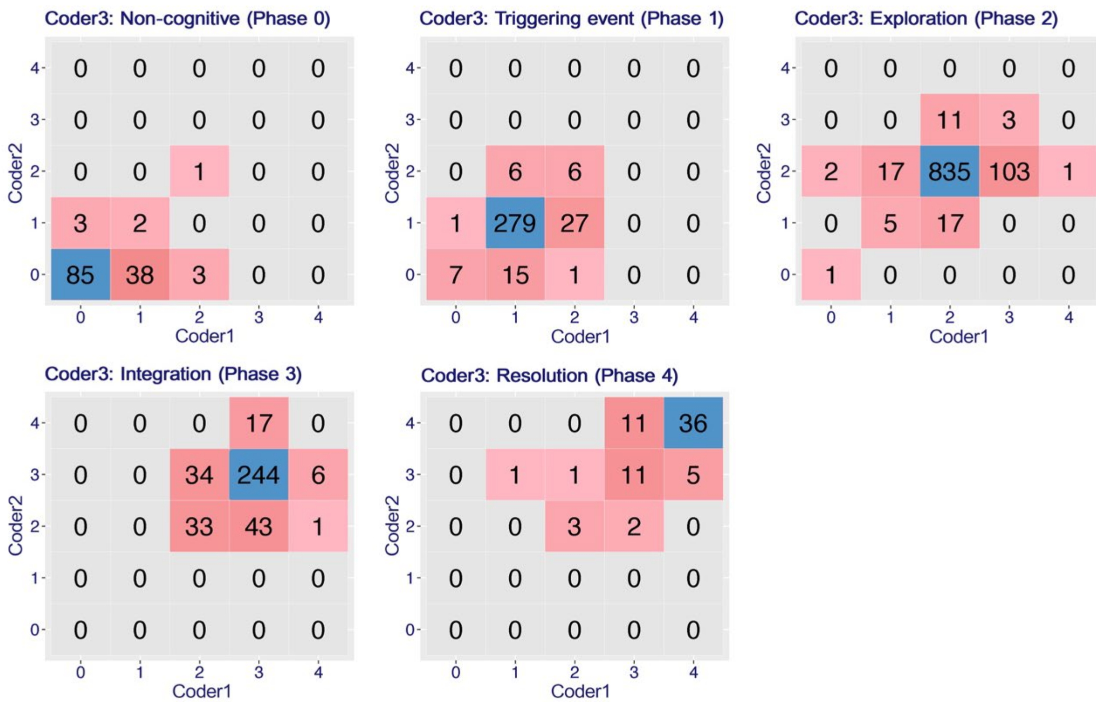
Combination of phases	Messages	
	n	%
Other (0) & triggering event (1)	66	15.07
Triggering event (1) & exploration (2)	78	17.81
Exploration (2) & integration (3)	227	51.83
Integration (3) & resolution (4)	50	11.42
Other (0) & exploration (2)	7	1.60
Exploration (2) & resolution (4)	4	0.91
Three labels	6	1.37

Note. *N* = 438 messages placed in multiple categories. Numbers in parentheses indicate the study's phase ID.

Another way to understand the distribution of the agreements and disagreements data is by using the five contingency tables in Figure 2, which align with Table 4 and Table 5. Figure 2 describes the distribution of messages classified by coder 1 and coder 2 when coder 3's labels ranged from the "other" phase to resolution, respectively. The blue cell on the diagonal of each table (e.g., 85 on the first table) represents the number of agreements between the three coders in each cognitive phase. The red cells (e.g., 38, 3, and 2 on the first table) demonstrate the number of disagreements.

Figure 2

Distribution of Five Cognitive Phases Between Three Coders



Note. Blue indicates the number of agreements among the 3 coders. Red indicates the number of disagreements. The red-colour scale is used to represent the disagreement cells. The larger the number, the darker the cell.

Discussion

Validation of the Adapted Coding Rubric—*SQ1*

The results (agreement of 77.29% and Fleiss' κ of 0.763) answer our first sub-question and show that a reliable inter-rater agreement was reached between the three coders in this research, as over 0.75 represents excellent agreement (Fleiss et al., 2003). The exploration phase accounts for most of the messages, and the resolution phase accounts for the least. This distribution rate is similar to the proportional results of cognitive phases from the reviewed studies (Kaul et al., 2018; Kovanović et al., 2014; Neto et al., 2018; Park, 2009). These discussion messages were reliably classified through the manual categorisation process, providing us with a clean training data set to develop automatic classifiers in our future work.

Disagreements Between Coders and What Caused the Disagreement—*SQ2*

Analysing the disagreements between coders (438 messages) can help to answer our second sub-question. Most of the three coders' disagreements appeared on adjacent phases of cognitive presence (Table 5 and Figure 2), in line with the findings in the previous study (Hu et al., 2020) with two coders. We analysed the common language patterns which may have caused disagreements between the coders.

Common Language Patterns of Messages Between Adjacent Phases

After reviewing the 66 messages (Table 5) which the coders labelled as "other" and triggering event, we came up with two possible reasons that may have caused these disagreements. First, messages with incomplete segments or concise sentences may have made it difficult to grasp the writer's purpose and meaning. There were ambiguous interpretations from different coders, even when they checked the previous and subsequent messages. The message instances were (a) "That works for me," and (b) "I am guessing…". Second, sentences using the structure "I like…" may have caused confusion. The verb *like* could be defined as either "I agree with you" or "appreciate". The instances were (a) "I like your worded comment. Nice!" and (b) "I like the questions you mentioned". The former could have been an indicator, "simple agreement," of a triggering event, whereas the latter could have been a predictor of social expression, which is part of the "other" phase.

The most common pattern reflected in the messages with both triggering event and exploration labels (78 messages as shown in Table 5) was the use of questions to deliver outside information or make personal claims. These language patterns confused our coders. "Ask questions" is a core indicator of the triggering event phase; however, learners can also propose ideas using sentences ending with question marks, such as in the case of rhetorical questions. These messages can be interpreted as an indicator of "suggestion for consideration" which is part of the phase of exploration. Two examples were: (a) "A good example of the strawman fallacy on me?" and (b) "At what point? Early on it is said 'will come to some philosophy department meeting'. Is that when?" Unlike writing essays or research reports as assignments in for-credit university courses, many online conversations in MOOCs contain informal writing. Coders cannot acquire sufficient information from the informal language patterns to verify the writer's actual intention. This vital difference in MOOC conversations compared to formal coursework may be associated with certain learner motivations, such as updating knowledge voluntarily and personal curiosity. This finding has a significant implication for building automatic classifiers of cognitive phases in future studies. For example, the number of question marks in a message in a MOOC would not be a positive indicator of the triggering event phase. However, in online discussions in smaller-scale, for-credit courses, the opposite can be true, and the number of question marks in a message may be reliably used as a predictor when building automatic classifiers (Kovanović et al., 2016).

The central debates in the 227 messages (Table 5) labelled both exploration and integration had two aspects. First, messages that contained conclusions with reasons raised a dispute about whether the supporting ideas were sufficient. A significant criterion to differentiate integration from exploration is that the message should reach "a coherent conclusion" by offering "sufficient substantiation" in the classification rubrics (Garrison et al., 2001; Park, 2009). This criterion is subjective and domain specific. Messages that provide solutions and implicit conclusions ending with a tentative phrase imply more of a "suggestion for consideration" (which should be labelled exploration), rather than sufficiently supported integration. For example, in our study, we saw this message: "I do not think it is an argument because a rates revolt is only a suggestion. The writer states that it is one way to make the councillors listen but does not say this strategy should be adopted". Two coders thought this message firstly disagreed with the previous message, and that the writer then proposed his/her opinion ("not an argument" with the supporting reasons [rest of the message]), and therefore, it should be labelled integration. However, the third coder thought it was just a personal opinion as suggestion for consideration without sufficient support, and labelled it exploration.

Second, messages with misleading language patterns could have impacted coders' decisions. Messages with language patterns, such as "consequently" or "both sides of" might indicate a conclusion or a "convergence" denoting the integration phase. However, such patterns could also be interpreted as a "leap to a conclusion" or a claim without supporting ideas, meaning that the messages should be classified in the phase of exploration. An example of such a message is, "Consequently, both sides of the arguments are equally compelling but have their share of fallacies. It depends on each person's confirmation bias to weigh a particular argument heavier than the other." These misleading language patterns tell us that some phrases and expressions can only be a possible predictor but not absolute evidence for classifying cognitive phases.

Most of the messages that were labelled as part of both the integration and resolution phases disputed whether the supporting ideas of new constructs were sufficient enough. This debate is very similar to the debates on distinguishing integration from exploration as discussed in this section.

Understanding the Reasons for the Disagreements

We found the bulk of the disagreements occurred between the exploration and integration phases. It may be because: (a) the proportion of messages in these two categories was much larger than in the other categories; (b) exploration and integration appear during the middle of a critical thinking activity, which tends to greater uncertainty, rather than at the beginning (awareness of a question) or the conclusion (outcomes after evaluation) stage; or (c) the criteria and instances of these two categories are ambiguous in the CoI framework, which is consistent with Rourke and Kanuka's (2009) critique about the lack of clear instances in Garrison et al.'s cognitive presence rubric. These reasons can also be connected with other critiques of the CoI. Garrison et al. (2001) borrowed from Dewey's (1933) five steps in reflective thinking to propose the four phases of cognitive presence. Still, they did not develop and elaborate on the theoretical foundations of Dewey's model (Jézégou, 2010). Garrison et al. (2001) merged the second ("diagnosis of a question") and third step ("suggestion of possible solution") from Dewey's (1933) model into the exploration phase. They renamed the fourth step ("elaboration of an idea by reasoning") as integration and the last step ("corroboration to form a concluding belief") as resolution. With respect to Dewey's model, the ambiguity of disagreements between exploration and integration occurred mainly when trying to distinguish "a suggestion of possible solution" (assigned to exploration) from the "elaboration of an idea by reasoning" (assigned to integration). In contrast, messages in the "diagnosis of a question step" (assigned to exploration) were easier for the coders to identify. Thus, we question whether the exploration phase should be separated back into the diagnosis step and a suggestion of possible solution step as defined in Dewey's (1933) model.

Henri and Lundgren-Cayrol (2005) proposed three phases of a collaborative learning approach for knowledge construction (e.g., exploration, elaboration, and evaluation), which intersect with the cognitive presence phases (Jézégou, 2010). The elaboration phase is positioned between exploration and evaluation, which is similar to the integration phase in cognitive presence. Henri and Lundgren-Cayrol (2005) also proposed two subcategories in the elaboration phase: *negotiation* and *validation*. The negotiation sub-phase refers to the learning processes that consider and collect other people's ideas to form diverse proposals of knowledge, and the validation sub-phase denotes consensus on the knowledge, reflecting multiple views (Henri & Lundgren-Cayrol, 2005). In this regard, the validation sub-phase is equivalent to the integration phase in cognitive presence. Interestingly, most of the ambiguous messages between exploration and integration in our sample data could be assigned into a negotiation sub-phase. For example, one learner compared two opinions from previous comments and generated her/his own statements in a message, but the statements had not been supported by sufficient reasoning, which meant the message was more exploration and not yet integration. This example fits well into negotiation. Therefore, there may be a negotiation sub-phase between a "considerable solution" step (assigned to exploration) and a "consensus idea by reasoning" step (assigned to integration). This would be an additional phase in the cognitive presence schema.

Apart from ambiguities of language patterns in MOOC discussions and insufficiencies of the CoI framework, another significant factor that caused the disagreements between all the adjacent phases was the increase in the data sample size from a small scale to a vast magnitude. Using the taxonomies to categorise cognitive processes works well on a smaller scale. In comparison, the likelihood of outliers increases when researchers apply the taxonomies developed from a smaller-scale dataset to classify vastly larger samples (Mayer-Schönberger & Cukier, 2013). Also, online discussions have an informal, conversational flow that is relatively messy, and does not fit into the ordered phases in the CoI (Xin, 2012). We assume that investigating the general trend of cognitive processes within the messiness of communication in the myriad MOOC transcripts is more valuable than using rigid classification methods.

Categorising learners' discussion transcripts into single-label cognitive phases tends to be subjective and inaccurate. One possible solution is to label the MOOC discussion messages into multiple cognitive presence phases with confidence levels. Another solution would be to label the messages by multiple models of learners' critical discourse simultaneously. For example, Farrow et al.'s (2021) study applied both cognitive presence and the ICAP framework (Chi & Wylie, 2014). These methods could provide a richer portrait for the interpretation of learners' dialogue by different coders using different frameworks and would reflect the diverse variation in the discourse more authentically.

In response to our main research question (*Is our adapted coding rubric of cognitive presence a reliable tool to classify MOOC discussions?*), we conclude that although the adapted rubric of cognitive presence is a statistically reliable tool to classify the discussion messages from the LCT MOOC by three coders, an additional phase (negotiation) could be included to improve the rubric to accommodate the predominant disagreements between coders.

Limitations

We acknowledge the limitation that a classification rubric of cognitive presence developed for one discipline might not be generalisable to other domains. There are disciplinary differences in the expression of critical reflection and its assessment in the pedagogical designs of different courses. The evaluation of cognitive presence that is mainly based on textual information can be highly domain specific. We are aware that the research findings might only be reliable and valid for the classification of cognitive presence in our target MOOC, done by three coders.

Conclusion and Implications

This research offers several theoretical and practical implications. We have reported on a process of classifying cognitive presence using more coders and a larger dataset to cross validate an adapted coding rubric in a target MOOC. The overall result reveals good inter-rater reliability, indicating that the adapted rubric remains stable for classifying cognitive phases in MOOC discussions by more coders and with larger datasets. We have then dug deeper into the messages where coders disagreed between adjacent cognitive phases. The possible causes of the ambiguous categorisation between adjacent cognitive phases could be theoretical insufficiencies of the CoI, MOOC learners' informal writing styles, and the changes of data size in MOOCs. We envisage that negotiation may be the additional phase between exploration and integration where most disagreements occurred. Our findings can inform the ongoing refinement of the CoI framework and provide a foundation for an approach to developing automatic analysis of educator-learner dialogue at scale.

This study also has practical implications. For preparing the machine learning datasets, we suggest using multiple-label instead of single-label classification to analyse learners' cognitive presence in MOOC discussions. This takes into account learners' informal language usage in MOOC discussions. It provides learning analytics researchers with some hints for choosing algorithms and predictive features in the study of automatic cognitive analysis. For example, better prediction performance may be achieved using corpora that include both informal speech and formal writing texts to train and generate the numeric representations of discussion messages that are fed into machine learning algorithms (e.g., neural networks). Also, the computational linguistic tools (e.g., Coh-metrix), which were created to assess formal essay writing (McNamara et al., 2012), may not be appropriate for analysing MOOC discussion messages. The application of a reliable and smart automatic classifier for analysing the processes of critical discourse in open and distance learning at scale can potentially (a) enable learners to self-evaluate their learning, to complement the automatic learner grading systems, (b) be used to inform the design and adaption of course content, and (c) assist the assessment of educator-learner online dialogue efficiently in real time.

References

Alario-Hoyos, C., Estévez-Ayres, I., Pérez-Sanagustín, M., Kloos, C. D., & Fernández-Panadero, C. (2017). Understanding learners' motivation and learning strategies in MOOCs. *The International Review of Research in Open and Distributed Learning*, *18*(3), 119–137. https://doi.org/10.19173/irrodl.v18i3.2996

Amemado, D., & Manca, S. (2017). Learning from decades of online distance education: MOOCs and the community of inquiry framework. *Journal of E-Learning and Knowledge Society*, *13*(2), 21–32. http://www.je-lks.org/ojs/index.php/Je-LKS_EN/article/view/137/75

Barbosa, G., Camelo, R., Cavalcanti, A. P., Miranda, P., Mello, R. F., Kovanovic, V., & Gaševic, D. (2020). Towards automatic cross-language classification of cognitive presence in online discussions. In C. Rensing & H. Drachsler (Chairs), *LAK '20: Proceedings of the tenth international conference on learning analytics and knowledge* (pp. 605–614). ACM. https://doi.org/10.1145/3375462.3375496

Buchem, I., Amenduni, F., Poce, A., Michaescu, V., Andone, D., Tur, G., Urbina, S., & Šmitek, B. (2020). Integrating mini-moocs into study programs in higher education during COVID-19. Five pilot case studies in context of the open virtual mobility project. In *EDEN 2020 annual conference: Human and artificial intelligence for the society of the future* (pp. 299–310). https://doi.org/10.38069/edenconf-2020-ac0028

Cha, H., & So, H.-J. (2020). Integration of formal, non-formal and informal learning through MOOCs. In D. Burgos (Ed.), *Radical solutions and open science. Lecture notes in educational technology* (pp. 135–158). Springer Singapore. https://doi.org/10.1007/978-981-15-4276-3_9

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Corich, S., Kinshuk, & Hunt, L. (2006). Measuring critical thinking within discussion forums using a computerised content analysis tool. In S. Banks, V. Hodgson, C. Jones, B. Kemp, D. McConnell, & C. Smith (Eds.), *Proceedings of the 5th international conference on networked learning*, *2*(1), 1–8. http://www.lancaster.ac.uk/fss/organisations/netlc/past/nlc2006/abstracts/pdfs/P07 Corich.pdf

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process* (revised ed.). D. C. Heath.

Dillahunt, T., Wang, Z., & Teasley, S. D. (2014). Democratizing higher education: Exploring MOOC use among those who cannot afford a formal education. *The International Review of Research in*

Open and Distributed Learning, *15*(5), 177–196. https://doi.org/10.19173/irrodl.v15i5.1841

Farrow, E., Moore, J., & Gašević, D. (2021). A network analytic approach to integrating multiple quality measures for asynchronous online discussions. In M. Scheffel, N. Dowell, S. Joksimovic, & G. Siemens (Chairs), *LAK '21: Proceedings of the 11th international conference on learning analytics and knowledge* (pp. 248–258). ACM. https://doi.org/10.1145/3448139.3448163

Fleiss, J., Levin, B., & Paik, M. (2003). *Statistical methods for rates and proportions* (3rd ed.). John Wiley & Sons. https://onlinelibrary.wiley.com/doi/book/10.1002/0471445428

Garrison, D. R. (2007). Online community of inquiry review: Social, cognitive, and teaching presence issues. *Journal of Asynchronous Learning Networks*, *11*(1), 61–72.

Garrison, D. R., & Anderson, T. (2011). *E-learning in the 21st century: A framework for research and practice* (2nd ed.). Routledge.

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), 87–105. 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., Anderson, T., & Archer, W. (2010). The first decade of the community of inquiry framework: A retrospective. *The Internet and Higher Education*, *13*(1–2), 5–9. https://doi.org/10.1016/j.iheduc.2009.10.003

Henri, F., & Lundgren-Cayrol, K. (2005). *Apprentissage collaboratif à distance: Pour comprendre et concevoir les environnements d'apprentissage virtuels* [Collaborative distance learning: Understanding and conceptualizing virtual learning environments]. Presses de l'Université du Québec.

Hu, Y., Donald, C., Giacaman, N., & Zhu, Z. (2020). Towards automated analysis of cognitive presence in MOOC discussions: a manual classification study. In C. Rensing & H. Drachsler (Chairs), *LAK '20*: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge (pp.135–140). ACM. https://doi.org/10.1145/3375462.3375473

Jézégou, A. (2010). Community of inquiry in e-learning: A critical analysis of the Garrison and Anderson model. *The Journal of Distance Education / Revue de l'Éducation à Distance*, *24*(3), 1–18. http://www.ijede.ca/index.php/jde/article/view/707

Kanuka, H., Rourke, L., & Laflamme, E. (2007). The influence of instructional methods on the quality of online discussion. *British Journal of Educational Technology*, *38*(2), 260–271. https://doi.org/10.1111/j.1467-8535.2006.00620.x

Kaul, M., Aksela, M., & Wu, X. (2018). Dynamics of the community of inquiry (CoI) within a massive open online course (MOOC) for in-service teachers in environmental education. *Education Sciences*, *8*(2), Article 40. https://doi.org/10.3390/educsci8020040

Kovanović, V., Joksimović, S., Gašević, D., & Hatala, M. (2014). Automated cognitive presence detection in online discussion transcripts. In K. Yacef & H. Drachster (Eds.), *Workshop proceedings of LAK 2014*. Sun SITE Central Europe (CEUR). http://ceur-ws.org/Vol-1137/LA_machinelearning_submission_1.pdf

Kovanović, V., Joksimović, S., Poquet, O., Hennis, T., Čukić, I., De Vries, P., Hatala, M., Dawson, S., Siemens, G., & Gašević, D. (2018). Exploring communities of inquiry in massive open online courses. *Computers & Education*, *119*, 44–58. https://doi.org/10.1016/j.compedu.2017.11.010

Kovanović, V., Joksimović, S., Waters, Z., Gašević, D., Kitto, K., Hatala, M., & Siemens, G. (2016). Towards automated content analysis of discussion transcripts. In D. Gašević & G. Lynch (Eds.), *LAK '16: Proceedings of the sixth international conference on learning analytics and knowledge* (pp. 15–24). ACM. https://doi.org/10.1145/2883851.2883950

Liu, C.-J., & Yang, S. C. (2014). Using the community of inquiry model to investigate students' knowledge construction in asynchronous online discussions. *Journal of Educational Computing Research*, *51*(3), 327–354. https://doi.org/10.2190/EC.51.3.d

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.

McKlin, T. (2004). *Analyzing cognitive presence in online courses using an artificial neural network* [Doctoral dissertation, Georgia State University]. ScholarWorks @ Georgia State University. https://scholarworks.gsu.edu/msit_diss/1/

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press. https://doi.org/10.1017/CBO9780511894664

Mladenić, D. (2010). Feature selection in text mining. In C. Sammut & G. I. Webb (Eds.), *Encyclopedia of Machine Learning* (pp. 406–410). Springer US. https://doi.org/10.1007/978-0-387-30164-8_307

Neto, V., Rolim, V., Ferreira, R., Kovanović, V., Gašević, D., Dueire Lins, R., & Lins, R. (2018). Automated analysis of cognitive presence in online discussions written in Portuguese. In V. Pammer-Schindler, M. Pérez-Sanagustín, H. Drachsler, R. Elferink, & M. Scheffel (Eds.), *EC-TEL 2018: Lifelong technology-enhanced learning*. *Lecture notes in computer science, vol. 11082* (pp. 245–261). Springer International. https://doi.org/10.1007/978-3-319-98572-5_19

Park, C. L. (2009). Replicating the use of a cognitive presence measurement tool. *Journal of Interactive Online Learning*, *8*(2), 140–155. http://www.ncolr.org/jiol/issues/pdf/8.2.3.pdf

Phan, T., McNeil, S. G., & Robin, B. R. (2016). Students' patterns of engagement and course performance in a massive open online course. *Computers and Education*, *95*, 36–44. https://doi.org/10.1016/j.compedu.2015.11.015

Rourke, L., & Kanuka, H. (2009). Learning in communities of inquiry: A review of the literature. *Journal of Distance Education*, *23*(1), 19–48. http://www.ijede.ca/index.php/jde/article/view/474/815

Sadaf, A., & Olesova, L. (2017). Enhancing cognitive presence in online case discussions with questions based on the practical inquiry model. *American Journal of Distance Education*, *31*(1), 56–69. https://doi.org/10.1080/08923647.2017.1267525

Siemens, G. (2013). Massive open online courses: Innovation in education? In R. McGreal, W. Kinuthia, & S. Marshall (Eds.), *Open educational resources: Innovation, research and practice* (pp. 5–16). Commonwealth of Learning and Athabasca University. https://www.oerknowledgecloud.org/archive/pub_PS_OER-IRP_web.pdf

Ullmann, T. D. (2019). Automated analysis of reflection in writing: Validating machine learning approaches. *International Journal of Artificial Intelligence in Education*, *29*(2), 217–257. https://doi.org/10.1007/s40593-019-00174-2

University of Auckland. (n.d.) *Logical and critical thinking*. FutureLearn. https://www.futurelearn.com/courses/logical-and-critical-thinking

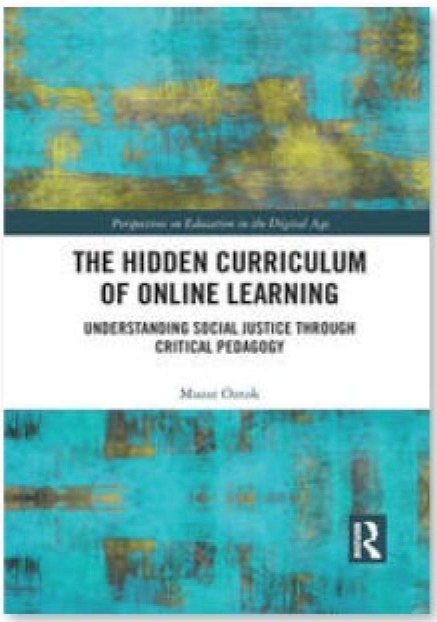
Xin, C. (2012). A critique of the community of inquiry framework. *The Journal of Distance Education*, *26*(1), 1–7. http://www.ijede.ca/index.php/jde/article/download/755/1333?inline=1

Athabasca University

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

Book Review: The Hidden Curriculum of Online Learning: Understanding Social Justice through Critical Pedagogy



Author: Murat Öztok (Routledge, 2019, 128 pages) ISBN: 9780429284052

Reviewed by: Alexandra Miller, *College of Education, Northern Arizona University, Flagstaff, United States*
Correspondence: aem432@nau.edu

Öztok's critical ethnography work, *The Hidden Curriculum of Online Learning: Understanding Social Justice Through Critical Pedagogy*, helps to fill a concerning gap in the online learning literature. Öztok argues that online learning literature is frequently positivist and apolitical with little emphasis on how learners make meaning of their experience or how learners are embedded in power hierarchies (p. 9). While this book contributes significantly to that epistemological gap in the online learning literature, it does not, nor should it, fill that hole entirely. Additional research is needed to explore learners' experiences in other contexts; perhaps more importantly, online learning scholars need to find ways to use research like this to improve learning experiences for students, especially those typically subjected to power differentials.

The purpose of *The Hidden Curriculum of Online Learning: Understanding Social Justice Through Critical Pedagogy* is to document how students experience online learning in terms of equity. Öztok carefully interrogates the term equity and how it operationalizes in online learning. Online learning is frequently thought of as an equalizer that provides access to education for those who might otherwise be excluded (Anderson, 2008; Harasim, 2000). However, as Öztok argues, this is equality more than equity (p. 6). This critical ethnography explores how online learning maintains cultural hegemony, as defined by Gramsci (Forgacs, 2000). In the book's conclusion, Öztok notes that his overall intention was to challenge how equity is defined in online learning (p. 112). Öztok meets that goal, which will hopefully help shape research designs going forward.

This book is divided into six chapters. The first chapter, "Genealogy of the Concepts and the Myths of Equity in Online Learning," provides the context for critical research in online learning. The second chapter, "How to Study Equity in Online Spaces: Situating the Theoretical Frameworks," lays out Öztok's argument and the theory that he integrates to support it. The next three chapters look at specific components of the theoretical framework. Chapter 3, "Writing Oneself into Online Being: The Art of Self-Representation and Impression Management," explores how learners craft their identity in an online environment, similar to Goffman's impression management theory (Goffman, 1956). Chapter 4, "Hierarchy of Privilege: Self as Curriculum of Diversity and Otherness," provides the richest ethnographic detail, looking at how diversity is performed in the classroom. Chapter 5, "Sociocultural Production of Self: Social Presence and Social Absence," dives into social presence. Öztok has published articles on social presence including a call for the term to be reconsidered in the online learning literature (Öztok & Kehrwald, 2017). The last chapter of this book, "Hidden Curriculum of Online Learning: Discourses of Whiteness, Social Absence, and Inequity," brings the entire argument together.

The most compelling argument in this book centers on the term social absence. The first substantive mention of social absence in the literature outside of Öztok's dissertation is in Öztok's (2014) presentation at the annual meeting of the Association of Internet Researchers. Previous uses of the term, such as Potter (2004), were glib references to complete non-participation in a class. In *The Hidden Curriculum of Online Learning: Understanding Social Justice Through Critical Pedagogy*, social absence is defined as "the extent to which particular identifications are not represented in one's online being" (p. 25). Öztok goes on to provide powerful interview excerpts where students describe leaving out key parts of their identity in efforts to fit in or make other students more comfortable. For some students, this means downplaying the non-Canadian aspects of their identity such as their Middle Eastern heritage. For others, downplaying their heritage is a consequence of being linked to the Canadian part of their hyphenated identity rather than standing alone as, for example, an Indian-Canadian. For the white Canadian students of British heritage in the study, social absence was not relevant; those students did not pick and choose aspects of their identity to present consciously as minority students did. The idea of social absence provides a useful heuristic to balance the frequently discussed concept of social presence in online learning.

The Hidden Curriculum of Online Learning: Understanding Social Justice Through Critical Pedagogy is rich in theory, but it offers few methodological insights. Öztok describes how he could not accomplish the Geertzian goal of ethnographic research as deeply hanging out (Geertz, 1998) and traditional ethnographic strategies had to be modified for the online learning context. He draws on analytics of learner use of the learning management system as a form of participant observation (p. 31). The reasoning behind this is sound, but little detail is provided. As ethnography of online learning is relatively new, the reader would have benefitted from more methodological detail than is provided by Öztok.

The Hidden Curriculum of Online Learning: Understanding Social Justice Through Critical Pedagogy is an exciting contribution to the online learning literature bringing a needed critical perspective. Hopefully this is the start of an expanded direction for online learning research. With a body of qualitative research exploring how students experience power hierarchies in online learning, the field will be able to focus on how to improve that experience for all learners.

References

Anderson, T. (Ed.). (2008). *The theory and practice of online learning* (2nd ed.). AU Press.

Forgacs, D. (Ed.). (2000). *An Antonio Gramsci reader: Selected writings 1916–1935*. Schocken Books.

Geertz, C. (1998). Deep Hanging Out. *The New York Review of Books, 45*(16), 69–72.

Goffman, E. (1956). *The presentation of self in everyday life*. Doubleday.

Harasim, L. (2000). Shift happens: Online education as a new paradigm in learning. *The Internet and Higher Education, 3*(1–2), 41–61. https://doi.org/10.1016/S1096-7516(00)00032-4

Öztok, M. (2014). Social presence and social absence: Socio-cultural production of self in online learning environments. *AoIR Selected Papers of Internet Research*, *4*. https://spir.aoir.org/ojs/index.php/spir/article/view/8874

Öztok, M., & Kehrwald, B. A. (2017). Social presence reconsidered: Moving beyond, going back, or killing social presence. *Distance Education, 38*(2), 259-266. https://doi.org/10.1080/01587919.2017.1322456

Potter, A. (2004). Interactive rhetoric for online learning environments. *The Internet and Higher Education, 7*(3), 183–198. https://doi.org/10.1016/j.iheduc.2004.06.002

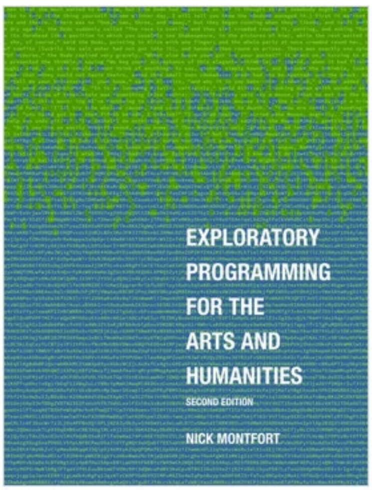
Athabasca University

CC BY

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 23, Number 2

May – 2022

Book Review: Exploratory Programming in the Arts and Humanities



Author: Nick Montfort (The MIT Press, 2021, 2nd ed., 363 pages)

ISBN: 9780262044608

Reviewed by: Kelly Hammond, *CUNY Graduate Center*

For humanists interested in incorporating programming into their practice or research, open-access (OA) resources abound. The Internet teems with how-to videos and textbooks—even full-blown OA courses—in open languages such as Python and R. These tutorials can be useful for faculty and students who can extrapolate with ease how a Python program, scraping jobs in the Houston area from monster.com, might be adapted to uncover new questions about women's literature in the late 1800s. For those seeking guidance rooted more directly in the humanities, there is a growing number of humanities-specific OA sites and journals, such as the Programming Historian, that provide lessons on skills useful for humanistic inquiry, such as how to geocode historical data or use stylometry to determine authorship. These resources can be great for faculty and students already comfortable with programming basics who have specific tasks they need to accomplish, but they do not lend themselves to much tinkering for the programming novice.

Nick Montfort's *Exploratory Programming in the Arts and Humanities*, whose second edition was released both in print and on MIT Press Open Access (available with a Creative Commons Attribution-NonCommercial-ShareAlike license), offers humanists something far more important than the programming he promises (and delivers) in the title: a mindset about how computational thinking and play can lead to discovery. He proves that programming, even the simple altering of someone else's code, can be fun, generative, socially transformative, and analytically powerful in humanities work. His book is useful as a complete textbook, as a course in and of itself, or as a dabbler's companion intended to open up possibilities. He writes both for those who have never programmed before as well as those who have. Further, he guides the reader to experiment with programming to explore text, images, and sound, and provides practical programming skills as well as a sense of structure that work across and between disciplines.

Structure and Content

Montfort begins with a grounding, philosophical introduction followed by a pragmatic chapter on installing the free software to be used throughout the book. He then presents fourteen instructional chapters—some devoted to general programming concepts as well as three specifically devoted to working with text, three to working with images, and one on working with sound. In each, he (a) introduces a concept, (b) asks readers to program along as he explains how the concept works, (c) presents free projects prompting readers to practice what they have learned through optional exercises, and (d) offers an essential concepts review for readers to solidify their understanding or spot areas that need a second look. Throughout it all, he provides tips on good practice, instilling habits that will save neophytes time and headaches as they increase their skill. He uses illustrations sparsely and wisely, generally to provide concrete visualization of complex ideas such as swapping pixels. These illustrations look hand drawn, like sketches on a napkin, lending a personal, tactile feel that many humanists crave.

Montfort also includes two not-to-be-missed appendices. The first, titled “Why Program?” shares provocative ideas about how programming can make us better thinkers and how it can shape a better world. Then, “Contexts for Learning” outlines ways the book can be used to teach ourselves and others, from what to do in a one-day workshop to how to sequence chapters for full semester- or quarter-long courses.

What’s new in the second edition? Montfort has updated both his instructions and the setup chapter to reflect newer versions of the free and open software packages used for the exercises. He has incorporated the use of free and open Jupyter Notebooks to demystify code interpretation and to allow readers to check their programs more easily as they build them. He has also reorganized the book, frontloading abstract concepts like statistics, probability, visualization, and classification before diving into higher-level chapters such as “Image III: Visual Design and Interactivity” and “Text III: Advanced Text Programming.” Importantly, has increased the frequency and visibility of his conviction that the best way to learn is to do.

Overall Impression and Relevance to the Field of Distance Education and E-learning

Exploratory Programming is a testament to what open-access can mean, especially in an e-learning environment. Used in full, it is a free course (that relies on free and open software) from a gifted MIT professor whose pedagogy is clear in structure and tone. He scaffolds, promotes predictive thinking, lauds collaborative learning, and urges readers to *do* not just to *read*. Used in part, it can be equally powerful. Thanks to Appendix B (“Contexts for Learning”), sections of the book can be easily adapted for a range of academic classes. Appendix A (“Why Program?”) could serve as a stand-alone philosophical introduction to any course with a required or optional programming component or unit. As online distance education assumes Internet connectivity, and not all students have Internet access at all times, Montfort provides guidance in Chapter 2 on how to program offline.

The book is also significant in that it broadens our scope as humanists. Montfort makes compelling arguments both within and across chapters that text, image, and sound speak to each other, especially through the language and structures of programming. A scholar of literature or history, hunting for open-access programming resources, may look only for tutorials on working with text, missing perhaps original opportunities to investigate visually or auditorily. The same is true for the scholars of art or archaeology who might limit themselves to working with images. Montfort helps the humanist see possibilities—analytical or creative—and then actually explore them.

The book's only real weakness is its limited cultural scope. References to Shakespeare, baseball, and *Spinal Tap* and the use of terms such as hose (to mean destroy) suggest an imagined reader who is perhaps white, American, and male. At times, Montfort is aware of these cultural assumptions, taking advantage of them to introduce some profound truths about cultural bias in computing. When introducing the conventional Hello, world! exercise, for example, he encourages users to alter the program to return phrases in other languages, not as a superficial nod to multilingual users, but rather as inquiry into the cultural assumptions built into the systems we use. Similarly, when exploring functions with if-else statements, he daringly asks the reader to imagine a world called Binaria where name endings indicate one of two genders. After leading the reader through this binary program, he makes the important point:

> Computers have been used to formalize and maintain traditional categories, including binary gender distinctions, ever since computers were invented. … Understanding how computer programs categorize, and how to build new and different systems, can have positive social potential when a classification issue like this one is involved. (Montfort, 2021, pp. 107–108)

Perhaps in the third edition, a wider audience can be assumed, and, even better, the book could be offered in fully open format, ripe for remixing by teachers and students who can insert their own cultural references, slang, and free project prompts.

Athabasca University

CC BY