



# Trends and Issues in MOOC Learning Analytics Empirical Research: A Systematic Literature Review (2011–2021)

Meina Zhu<sup>1</sup> · Annisa R. Sari<sup>2,3</sup> · Mimi Miyoung Lee<sup>4</sup>

Received: 23 December 2021 / Accepted: 28 March 2022 / Published online: 11 April 2022  
© The Author(s) 2022

## Abstract

Learning analytics (LA) is a growing research trend and has recently been used in research and practices in massive open online courses (MOOCs). This systematic review of 166 articles from 2011–2021 synthesizes the trends and critical issues of LA in MOOCs. The eight-step process proposed by Okoli and Schabram was used to guide this systematic review in analyzing publication outlets, research purposes and methods, stakeholders, and researchers' geographic locations and subjects. The review showed that MOOC LA research is primarily published in top-tier journals, such as the *Journal of Learning Analytics*, *Journal of Computer Assisted Learning*, and *Computers & Education*, as well as conference proceedings. The review also revealed that LA in MOOCs was used more frequently for the purpose of research than for practice (i.e., learning and teaching). Approximately 60% of the reviewed studies adopted learners' log data and achievement data as the primary data sources. Statistics, machine learning, content analysis, social network analysis, text analysis, and data visualization were the top six specific data analysis techniques used in the MOOC LA studies. Regarding collaboration, more than half of the reviewed studies involved interdisciplinary collaborations, and approximately one-third involved international collaborations. We suggested future studies on MOOC LA interventions to improve learning and teaching practices, and the active interdisciplinary collaboration to increase the rigor of the studies and the dissemination of the knowledge. More detailed discussion and implications for research and practice are presented. This research provides insights on future research and practices regarding LA use in MOOCs.

---

✉ Meina Zhu  
meinazhuu@gmail.com

<sup>1</sup> Learning Design and Technology, Wayne State University, 365 Education Building, 5425 Gullen Mall, Detroit, MI 48202, USA

<sup>2</sup> Department of Accounting Education, Yogyakarta State University, Yogyakarta, Indonesia

<sup>3</sup> Instructional Systems Technology, Indiana University, Bloomington, IN, USA

<sup>4</sup> Department of Curriculum & Instruction, University of Houston, Houston, TX, USA

**Keywords** MOOCs · Learning analytics · Systematic review · Analysis techniques

## 1 Introduction

Massive open online courses (MOOCs) have become increasingly popular in education since 2008 as an open-source learning environment to help learners around the world enhance their knowledge and skills. MOOCs have already had a considerable impact on educational research. The large number of learners has provided opportunities for large-scale educational interventions and has introduced new research questions and methods, and techniques to educational research. For example, educators and researchers have recently started paying more attention to the challenges of MOOCs, such as the high dropout rate, low interaction, assessment, and personalization in MOOCs (Zhu et al., 2019, 2018). Using learning analytics (LA) in MOOC research and practice can help researchers and educators understand and address these issues.

Since 2011, LA studies have become increasingly popular in education, psychology, computer science, and data science (Hui & Kwok, 2019). LA is commonly used for studies aiming to understand and support learners' learning behavior based on large data. Therefore, research on MOOCs and LA is interconnected. Considerable data in MOOCs can be processed using diverse LA technology and techniques. For example, LA can use different levels of MOOC data (e.g., course level, program level, platform level) to improve teaching and learning (Drachler & Kalz, 2016; Lemay & Doleck, 2020; Mubarak et al., 2021) and instructional design (Shukor & Abdullah, 2019). However, despite the growing popularity of LA studies in MOOCs, the understanding of the publication outlets, research purposes and methods, primary stakeholders, researchers' institution locations, subject areas, and research context of MOOC LA research is limited. The limited understanding could hinder future research on MOOC LA studies. Therefore, the present systematic review fills this gap to guide future research. We first present the background of MOOCs and LA in a comprehensive literature review, followed by the method section describing the various systematic review approaches, the databases searched, keywords used in the literature search, and selection criteria.

## 2 Literature review

### 2.1 MOOC systematic review

A wide range of free and open education courses are available around the world and have received considerable attention as the number of MOOCs continues to grow. In 2020, 16,300 courses were offered by 950 universities, with enrollments of more than 180 million learners globally (excluding data from China) (Shah, 2020). These MOOCs have been offered by numerous providers, including Udacity, Coursera, FutureLearn, and edX (Taneja & Goel, 2014; Sari et al., 2020). MOOCs were introduced in 2007 and have continued to expand globally, evolving into various forms,

including cMOOCs, xMOOCs, and blended MOOCs (Creed-Dikeogu & Clark, 2013; Kop, 2011; Mota & Scott, 2014; Zhu, 2021).

With the increasing popularity of MOOCs, a substantial number of literature reviews of MOOC research have been conducted. The authors of the present study (2018; 2020) conducted a systematic review of MOOC empirical research focusing on a general analysis, including the topics, methods, publication outlets, and geographical distributions of authors and courses. Similar systematic review studies were also conducted by Joksimović et al. (2018), Calonge and Shah (2016), and Lambert (2020). Moreover, several systematic review studies have also targeted specific subtopics or foci, such as the languages used in MOOCs (Hidalgo & Abril, 2020; Sallam et al., 2020; Zainuddin et al., 2019), learner engagement (Guajardo Leal et al., 2019; Paton et al., 2018; Rincón-Flores et al., 2019), assessment methods (Alturkistani et al., 2020; Tenório et al., 2016; Wei et al., 2020), self-regulated learning (Alonso-Mencía et al., 2020; Lee et al., 2019; Min & Nasir, 2020; Wong et al., 2019), dropout and completion rates (Bezerra & Silva, 2017; Dalipi et al., 2018), MOOCs in higher education (Albelbisi et al., 2018; Al-Rahmi et al., 2019; Calonge & Shah, 2016), social aspects of MOOCs (Lambert, 2020; Rolfe, 2015; Van de Oudeweetering & Agirdag, 2018), and MOOCs in the Asia Pacific region (Albelbisi & Yusop, 2020; Li et al., 2017). Based on these MOOC research reviews, it is clear that the number of MOOC studies has persistently expanded (Martin et al., 2020; Zhu et al., 2018).

While some scholars have doubted the viability of the future of MOOCs, others have predicted continued growth (Yahoo, 2021). Furthermore, these experts have predicted that MOOCs will “grow by \$ 16.01 billion from 2021–2025, progressing at a CAGR of 32% during the forecast period” (Yahoo, 2021). Such optimism is supported by the continuous growth of MOOC enrollment numbers over the past decade. With continued growth, more diverse MOOC research and discussion will be needed to further address concerns such as long-term and short-term progress, the dropout rate, self-directed learning, learner engagement, and course quality (Jona & Naidu, 2014). The current, up-to-date systematic review of MOOC empirical studies aims to serve this purpose.

## 2.2 Learning analytics overview

The current era of big data has allowed a vast amount of data to be captured and stored digitally (Siemens & Long, 2011), increasing “the volume, variety, velocity and veracity of student data” (Prinsloo & Slade, 2017, p. 8). Using LA techniques to analyze the large data can produce results that help complement and improve the performance of institutions and organizations and provide personalized and learner-centered education (Asamoah et al., 2017; Jantti & Heath, 2016).

Although the definition of LA varies, a commonly used definition by the Society for Learning Analytics Research (SoLAR) defines it as “the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” (Siemens & Long, 2011, p.33). LA involves information retrieval, machine learning

techniques, statistics, and data visualization (Siemens, 2012). It is also an interdisciplinary area related to the fields of computer science, data science, statistics, behavior science, educational psychology, and instructional technology to enhance education (Tzimas & Demetriadis, 2021).

Initially, LA was primarily used to analyze trace data to capture and describe learner behaviors in online learning (Veenman, 2013). Gradually, it became more broadly applied, including predicting what learners are more likely to drop out of the courses (Sclater et al., 2016). Researchers have classified LA approaches into different categories, such as descriptive analytics, predictive analytics, and prescriptive analytics (Berland et al., 2014). Descriptive analytics reports the process or status of student learning using data sources, including surveys, assessment results, learning management activities, and learner demographic information. Descriptive LA helps educators understand learning behaviors (Wong & Chong, 2018). Predictive analytics aims to anticipate future learner behaviors and learning success. Prescriptive analytics primarily uses algorithms to not only predict learner success but also recommend instructional interventions based on the data (Baker & Siemens, 2015). For example, it can provide helpful suggestions to instructors, learners, and policymakers (Hwang et al., 2014). Currently, LA in higher education is mainly used to improve student learning, understand the motivation to decrease the dropout rate (Colvin et al., 2015; Glick et al., 2019), provide adaptive learning environments to enhance students' learning trajectory (Ifenthaler & Schumacher, 2016; Ifenthaler et al., 2019), and provide information for institutions to allocate resources and funds for learners' success (Leitner et al., 2017).

### 2.3 Systematic review on learning analytics

There has been a recent rapid increase in the number of systematic reviews of LA in education. One of the earliest systematic reviews was by Papamitsiou and Economides (2014), who analyzed 40 LA empirical studies published between 2008 and 2013 to capture the history chronologically. During the past several years, the systematic reviews on LA substantially increased. These systematic reviews included diverse foci, such as serious games (Alonso-Fernandez et al., 2019), machine learning (Korkmaz & Correia, 2019), LA interventions (Knobbout & Van Der Stapen, 2020; Larrabee S nderlund et al., 2019), visual LA (including LA dashboard) (Schwendimann et al., 2016; Vieira et al., 2018), visual LA and self-regulated learning (Matcha et al., 2019), LA and learning outcomes or performance (Blumenstein, 2020; Costa et al., 2020; Foster & Francis, 2020), data sources used in LA (Samuelson et al., 2019), LA and instructional design (Mangaroska & Giannakos, 2018), ethic issues in LA (Tzimas & Demetriadis, 2021), flipped classrooms (Algayres & Triantafyllou, 2020), mobile learning settings (Pishtari et al., 2020), and LA at different educational levels, such as for young children (Crescenzi-Lanna, 2020) and higher education (Avella et al., 2016; Ifenthaler & Yau, 2020), and LA in vocational education (Gedrimiene et al., 2020).

For example, Foster and Francis (2020) reviewed 34 articles published between 2007 and 2018 on the effectiveness of using LA to improve learning outcomes in higher education. They reported that 75% of the studies reported that LA was effective in increasing learners' learning outcomes. However, the authors suggested that the relationship between LA and learning outcomes needed further investigation. Recently, the ethics of using LA has also gained attention from researchers. Tzimas and Demetriadis (2021) reviewed 53 articles published (2011–2018) on LA ethics. They found that there are inadequate empirical guidelines on LA guidelines and suggested establishing policies or codes to monitor and evaluate LA practice ethics.

Although numerous systematic reviews on LA have been conducted, studies of a systematic review related to the use of LA in MOOCs are lacking. The aim of the present study was to systematically review the empirical research literature related to LA and MOOCs to identify the current status and trends in this field. The following research questions guided this study:

- (1) What are the publication outlets of MOOC LA research published in the past eleven years?
- (2) What are the research purposes and methods utilized in the MOOC LA research published in the past eleven years?
- (3) Who are the primary stakeholders of MOOC LA studies published in the past eleven years?
- (4) Where are the employers/institutions of MOOC researchers located, and what are the subject areas of the researchers of these MOOC LA studies published in the past eleven years?
- (5) Which countries and subjects were studied most in MOOC research in the past eleven years?

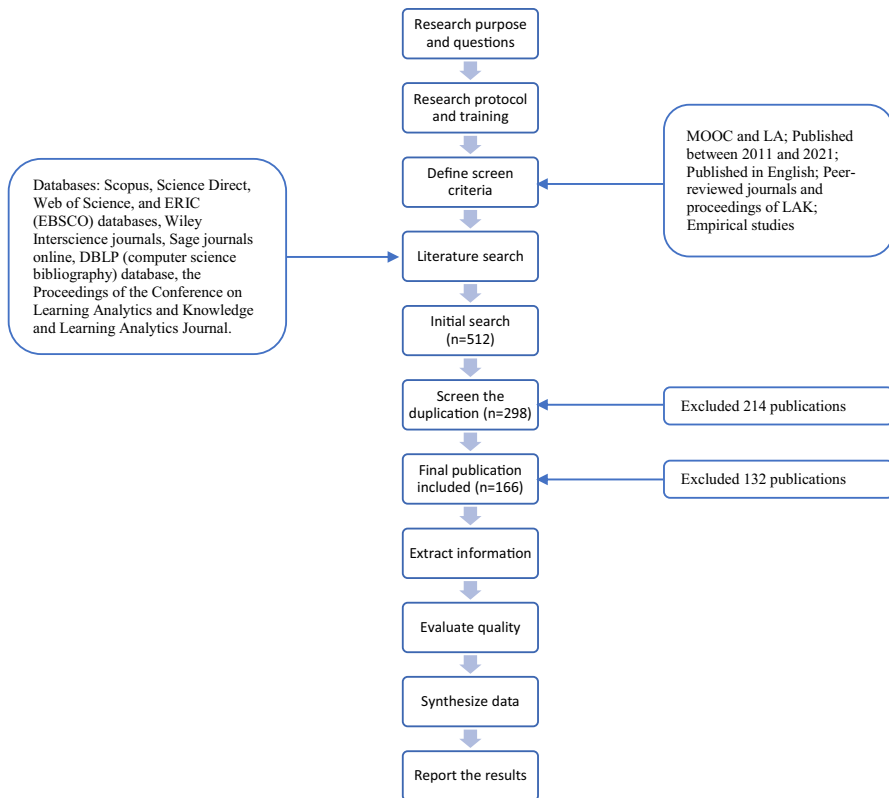
### 3 Methods

#### 3.1 Data collection

The present study followed the eight-step process proposed by Okoli and Schabram (2010). To conduct a rigorous systematic review, the following eight steps were followed (Okoli, 2015): (1) identify the review purpose; (2) create protocol and conduct training; (3) define screen criteria; (4) literature search; (5) extract information; (6) evaluate quality; (7) synthesize data; and (8) report the review results (see Fig. 1).

The flowchart in Fig. 1 presents the process of conducting the systematic review. Following the purpose and research questions stated earlier, the research team created a research protocol and detailed the procedures for conducting the review. To increase the validity of the study, a training session regarding the database and reviewing techniques was held with the reviewers. The selection criteria used in the review process are as follows:

- Studies were related to LA and MOOCs;



**Fig. 1** The systematic review process

- Studies were published from 2011 to the end of April 2021 because the first learning analytic conference was held in 2011 when this field garnered researchers' attention (Khalil & Ebner, 2016), and MOOCs started to gain popularity in 2011 (Zhu, 2021);
- Studies were published in English;
- Studies were empirical studies from peer-reviewed journal articles or the proceedings of the International Learning Analytics and Knowledge (LAK) conferences. These conference proceedings were chosen because the conferences serve as major forums for LA research and employ a rigorous selection review. Peer-reviewed journals were chosen for their higher standards of research objectivity and credibility (Utah State University Library, 2020) compared to non-peer-reviewed book chapters, blogs, and magazines, which were excluded.

The following steps were used for the literature search.

- (1) The first two authors did the initial search in various journal databases. One researcher searched articles in Scopus, Science Direct, Web of Science, and

ERIC (EBSCO) databases. The second researcher searched Wiley Interscience journals, Sage journals online, DBLP (computer science bibliography) database, and the Proceedings of the Conference on Learning Analytics and Knowledge and Learning Analytics Journal.

- (2) A combination of the search keywords “learning analytics and MOOCs” and “learning analytics and massive open online courses” were used to screen abstracts and titles of research articles on MOOC LA.
- (3) The initial results yielded 512 articles. To increase the validity, two researchers cross-checked the data and reached a consensus. After reviewing duplicates, 214 articles were deleted, resulting in 298 remaining articles.
- (4) Each researcher screened the full text of the 298 articles using the criteria, evaluated the articles, and crossed checked the other researchers’ results. A total of 166 articles met the review criteria with an overall inter-rater agreement of 95.2%.

Guided by the five research questions, the following information was extracted and recorded from each article: publication year, research purposes, general research approaches, data sources, data analysis techniques, stakeholders, author locations, author subject areas, MOOC locations, and MOOC subject areas. The names of the journals and the SCImago index were also recorded. The final stage of the systematic review included reporting and disseminating the findings of the papers, including the findings, discussions, limitations, implications, and conclusions.

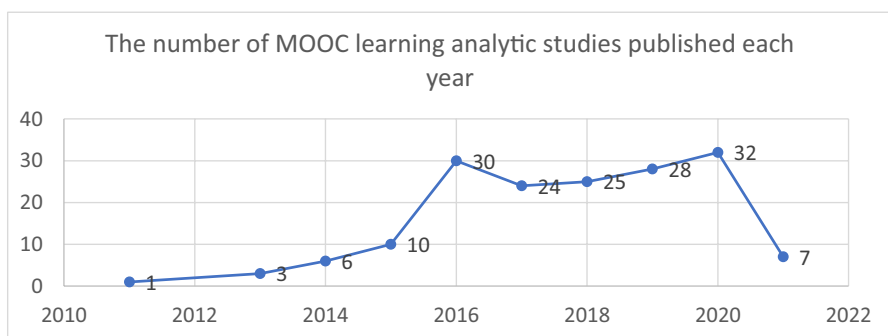
### 3.2 Data analysis

To answer Research Question (RQ) #1, we counted the number of MOOC LA publications from each publication outlet. In addition, the quartile rank for each journal was collected from the SCImago website. For RQ #2, we coded the research purposes using the following four categories: teaching, learning, research, and others. *Teaching* refers to studies that used LA techniques primarily to improve teaching practices; *learning* refers to the studies that primarily used LA techniques to improve learners’ learning in practice; *research* refers to the studies that only used LA techniques to conduct research; the studies that did not belong to these three categories were categorized as *others*. Regarding research methods, the authors analyzed the data sources and LA techniques. For data sources, the authors adapted the categories identified by Tashakkori and Teddlie (2003), such as interviews and surveys, and added new categories including log data, achievement data, demographic data, discussion forums, surveys, assignments, quizzes, videos, social media, and interviews. Log data in the present study refers to the data generated and recorded on learners’ behaviors and activities in the MOOC platform, such as clickstream logs and video/page views. Achievement data refers to data such as learners’ test scores and completion rates. For data analysis techniques, the present research adapted categories from Khalil and Ebner (2016) and Alonso-Fernandez et al. (2019) and classified them into two categories: qualitative and quantitative data. The qualitative data were categorized into the traditional qualitative analysis (e.g., thematic analysis,

content analysis, and discourse analysis) and LA text analysis techniques (e.g., text mining, semantic and linguistic analysis, and natural language processing). For quantitative data, we categorized the data into traditional statistics, social network analysis, and algorithm methods. Traditional statistics included regression, correlation analysis, ANOVA, MANOVA, chi-squared, and logistic regression. Algorithm methods included supervised learning (e.g., linear and logistic regression, regression and decision trees, support vector machines, Bayesian networks, neural networks, naive Bayes, and Bayesian knowledge tracing), unsupervised learning (e.g., correlation, clustering, factor analysis), and data visualization (e.g., performance metrics, heatmaps of interactions). For RQ #3 on the primary stakeholders, the authors categorized stakeholders into instructors, learners, providers (i.e., MOOC institutions, MOOC platforms), and instructional designers. Stakeholders represent to whom the research study findings most benefit. To answer RQ #4, we counted the countries and subject fields of all MOOC authors' affiliations in this study. For RQ #5, we counted the countries of users (students) and subject areas of the MOOCs being studied. For the studies that did not provide specific information on the MOOCs, the authors coded them as NA. Excel 16.49 and Python were used to calculate the numbers and create visual presentations.

## 4 Results

The authors collected 166 MOOC LA research studies (see Fig. 2). In terms of publication dates, most of the studies were published since 2016: one article (0.6%) was published in 2011, two articles (1.8%) in 2013, six articles (3.6%) in 2014, 10 (6.0%) in 2015, 30 (18.1%) in 2016, 24 (14.5%) in 2017, 25 (15.1%) in 2018, 27 (16.9%) in 2019, 32 (19.3%) in 2020, and 7 (4.2%) by the end of April 2021 (see Fig. 2). This indicates that the number of MOOC LA studies has progressively increased (by 2020). Given that only 2021 articles published by the end of April 2021 were included, we are unsure whether the number of publications in 2021 will show an increase or decrease.



**Fig. 2** The number of MOOC learning analytic studies published each year (2011–2021) (n = 166)



#### 4.1 Research Question #1 (RQ #1): What are the publication outlets of MOOC learning analytics research published in the past eleven years?

The 166 MOOC LA studies in this review were published in 72 different journals or proceedings. We also analyzed the journal rank in which the LA studies were published using quartile rankings (Q1–Q4). Quartile ranking refers to the impact factor of a journal in a specific subject/field from Q1 (the most prestigious) to Q4 (the least prestigious) (Rømer et al., 2020; Tóth & Demeter, 2021). More than 40% of MOOC LA articles were published in top-tier journals (Q1). Specifically, among the journals and proceedings we analyzed, 71 articles were published in Q1 journals, 23 articles in Q2 journals, 13 in Q3 journals, three in Q4 journals, and 54 articles in journals/proceedings that were not assigned in any quartile rank (see Fig. 3). The *Proceeding of the International Learning Analytics & Knowledge Conference* and the *Journal of Learning Analytics*, which are hosted by the *Society of Learning Analytics Research Association*, were the primary publication outlets for the MOOC LA studies.

The top 15 publication outlets included: *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (n = 13), *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (n = 10), *Journal of Learning Analytics* (n = 9; Q1), *Journal of Computer Assisted Learning* (n = 7; Q1), *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (n = 6), *Computers & Education* (n = 6; Q1), *Computers in Human Behavior* (n = 6; Q1), *Proceedings of the Eighth International Conference on Learning Analytics and Knowledge* (n = 5), *International Review of Research*

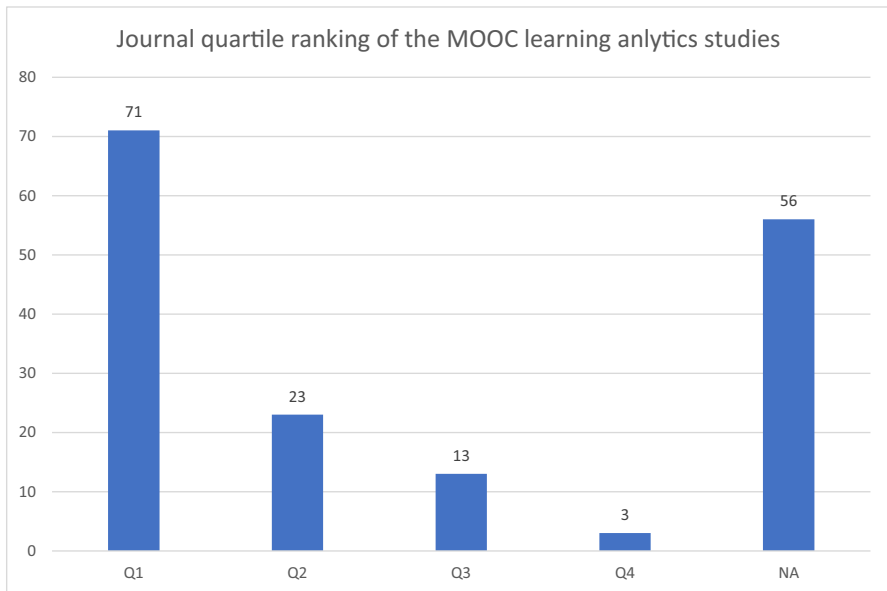
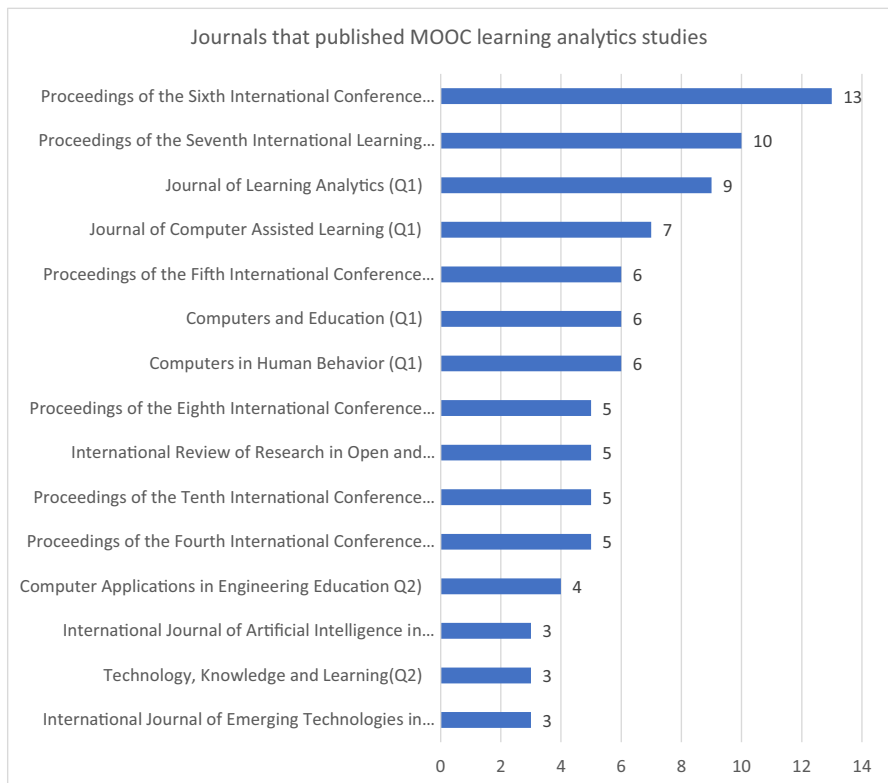
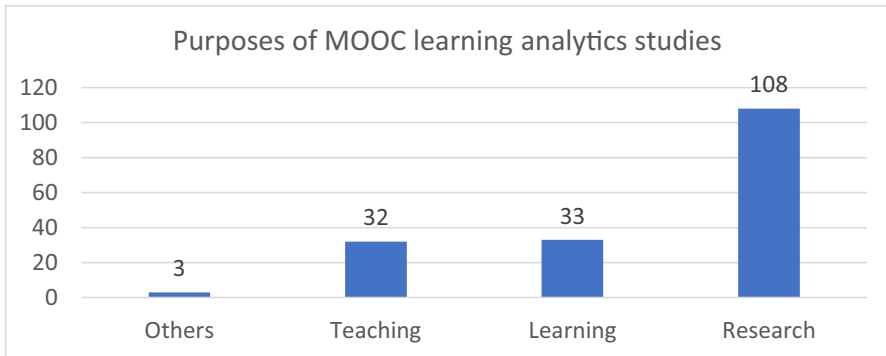


Fig. 3 Ranks of journals where MOOC learning analytic studies were published (2011–2021) (n = 166)

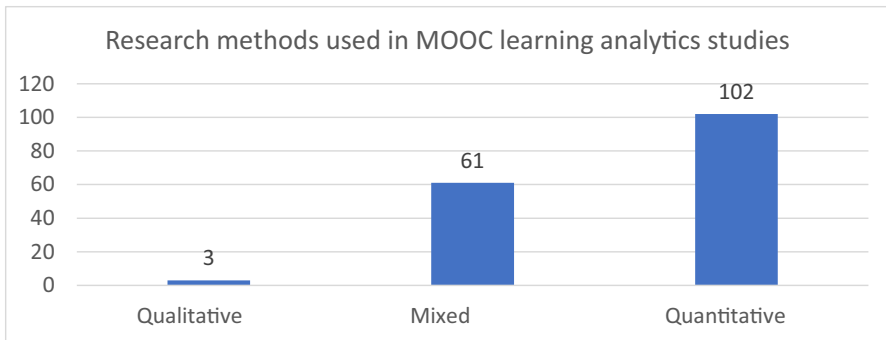
in *Open and Distance Learning* (n=5; Q1), *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (n=5), *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge* (n=5), *Computer Applications in Engineering Education* (n=4; Q2), *International Journal of Artificial Intelligence in Education* (n=3; Q1), *Technology, Knowledge and Learning* (n=3; Q2), and *International Journal of Emerging Technologies in Learning* (n=5; Q3) (see Fig. 4). Besides the conference proceedings, the *Journal of Learning Analytics* (n=9), *Journal of Computer Assisted Learning* (n=7), and *Computers & Education* (n=6) are the top three journals that published MOOC LA related articles. This showed that the *Proceeding of the International Learning Analytics & Knowledge Conference* and the *Journal of Learning Analytics*, which are hosted by the *Society of Learning Analytics Research Association*, were the primary publication outlets for MOOC LA studies.



**Fig. 4** Journals that published MOOC LA studies (the top 15 journals/proceedings) (2011–2021) (n = 166)



**Fig. 5** Purposes of MOOC learning analytic studies published (2011–2021) (n = 166)



**Fig. 6** Research methods used in MOOCs LA studies (2011–2021) (n = 166)

#### **4.2 RQ #2 What are the research purposes and methods utilized in the MOOC learning analytics research published in the past eleven years?**

Among the 166 reviewed MOOC LA studies, the majority used LA techniques for the primary purpose of research (n = 108) instead of teaching practice (see Fig. 5). The remaining studies used LA for education practices such as learning (n = 33) and teaching (n = 32). Given that fewer studies used analytics to directly improve teaching and learning practices compared to the research purposes, studies on how to use LA to improve educational practices need further research.

Regarding the research methods used, among the 166 articles, 102 articles (61.4%) were quantitative, 61 studies (36.7%) used mixed methods, and the remaining three articles (1.9%) were qualitative (see Fig. 6). It is not surprising that a majority of the MOOC LA studies were quantitative studies. Over one-third of the studies also included qualitative data in the research. The three qualitative studies primarily used a qualitative method to explore the MOOC LA phenomenon.

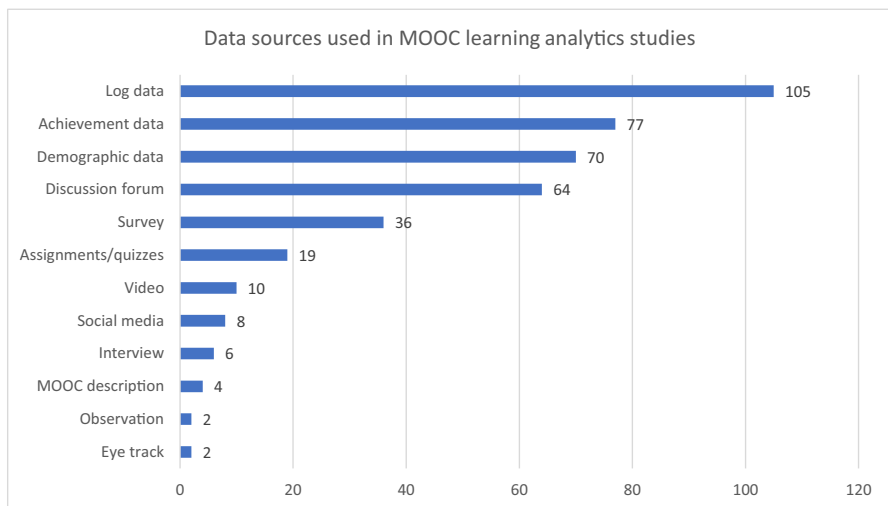
More specifically, among the various research methods, the studies in this review used diverse data sources. The most used data source was log data

( $n = 105$ ) followed by achievement data ( $n = 77$ ), demographic data ( $n = 70$ ), discussion forum ( $n = 64$ ), survey ( $n = 36$ ), assignments or quizzes ( $n = 19$ ), videos ( $n = 10$ ), social media ( $n = 8$ ), interviews ( $n = 6$ ), MOOC descriptions ( $n = 4$ ), observations ( $n = 2$ ), and eye-tracking data ( $n = 2$ ) (see Fig. 7). The results revealed that auto-generated data from MOOCs were the primary data sources in MOOC LA studies.

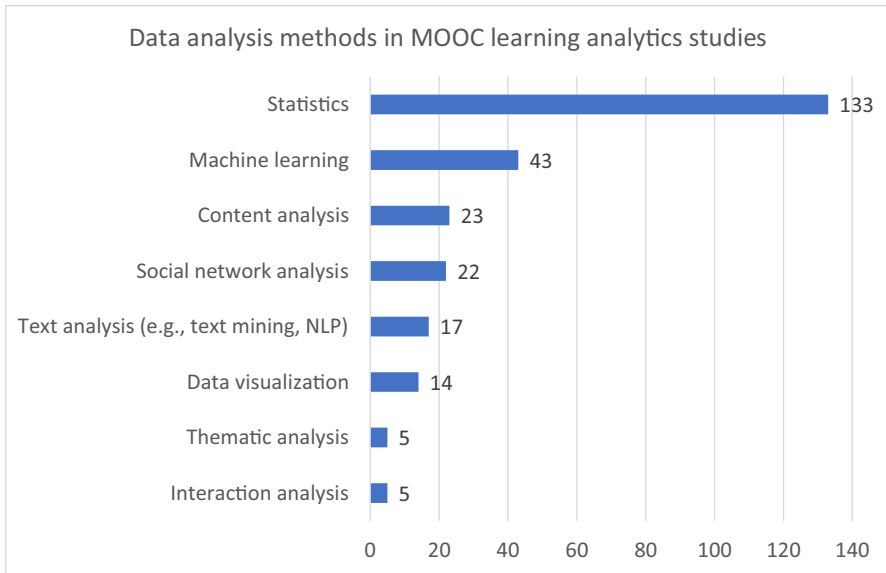
The top ten specific data analysis techniques used in the MOOC LA studies were statistics ( $n = 133$ ), machine learning ( $n = 43$ ), content analysis ( $n = 23$ ), social network analysis ( $n = 22$ ), text analysis ( $n = 17$ ), data visualization ( $n = 14$ ), thematic analysis ( $n = 5$ ), and interaction analysis ( $n = 5$ ) (see Fig. 8). The majority of the studies used statistical techniques, which included descriptive statistics, correlation analysis, regression analysis, and structural equation modeling. The second most used analytic technique was machine learning, which included supervised learning and unsupervised learning approaches.

### 4.3 RQ #3. Who are the primary stakeholders of MOOC learning analytics studies published in the past eleven years?

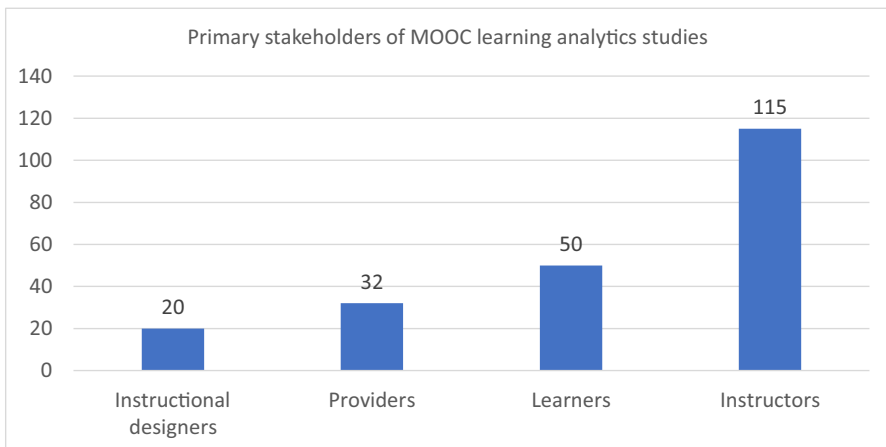
The primary stakeholders of MOOC LA studies were instructors ( $n = 115$ ), followed by learners ( $n = 50$ ), providers ( $n = 32$ ), and instructional designers ( $n = 20$ ) (see Fig. 9). Thus, it seems that MOOC LA studies aimed to benefit instructors the most.



**Fig. 7** Data sources used in MOOCs LA studies (2011–2021) ( $n = 166$ )



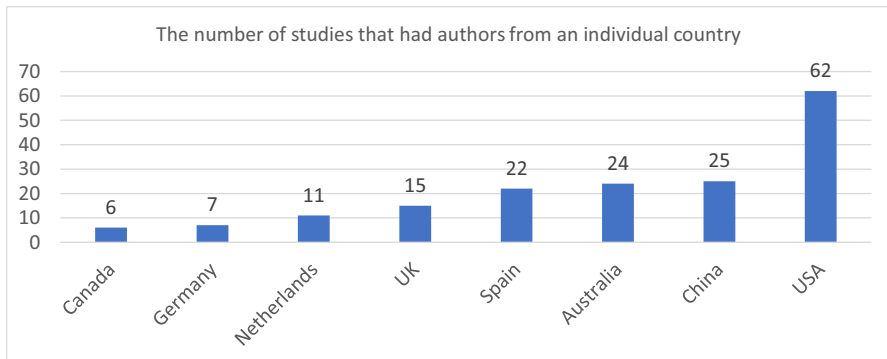
**Fig. 8** Top eight data analysis techniques used in MOOC LA studies (2011–2021) (n = 166) (Note: One article may have used more than one data analysis technique)



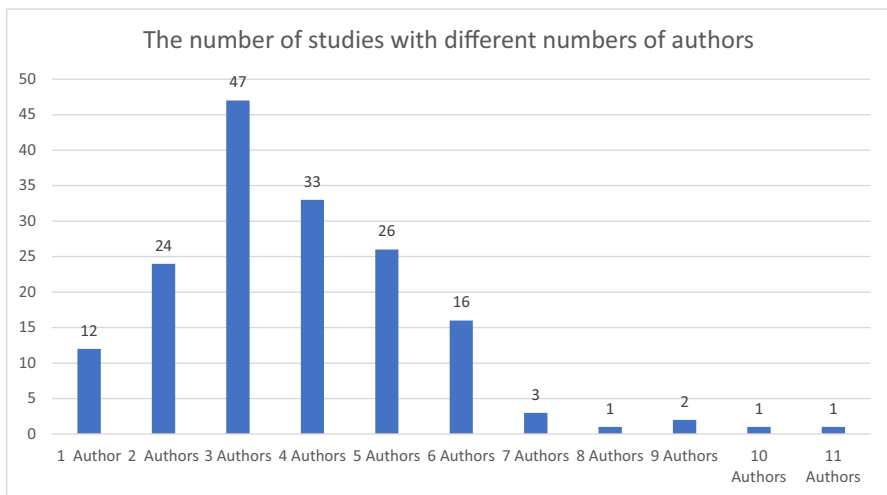
**Fig. 9** The primary stakeholders of MOOC LA studies (2011–2021) (n = 166)

#### **4.4 RQ #4. Where are the employers/institutions of MOOC researchers located, and what are the subject areas of the researchers of these MOOC learning analytics studies published in the past eleven years?**

Among the 166 MOOC LA studies, authors of 62 studies were from the USA, followed by authors from China (n = 25), Australia (n = 24), Spain (n = 22), UK



**Fig. 10** The number of studies with authors from an individual country (2011–2021) (n = 166)

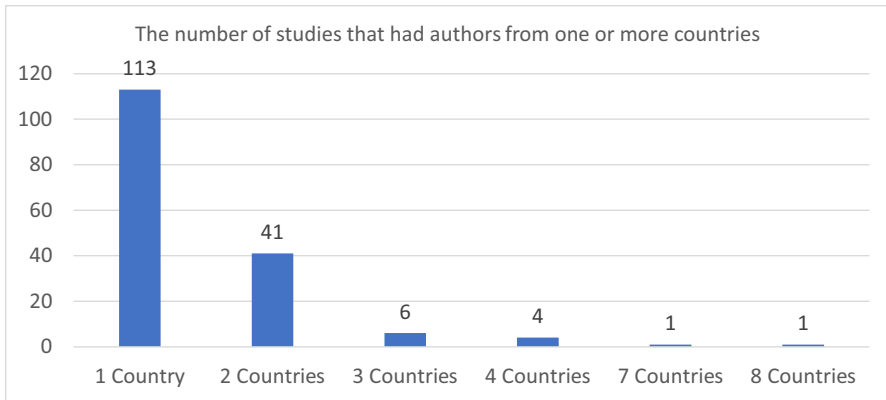


**Fig. 11** The number of studies with different numbers of authors (2011–2021) (n = 166)

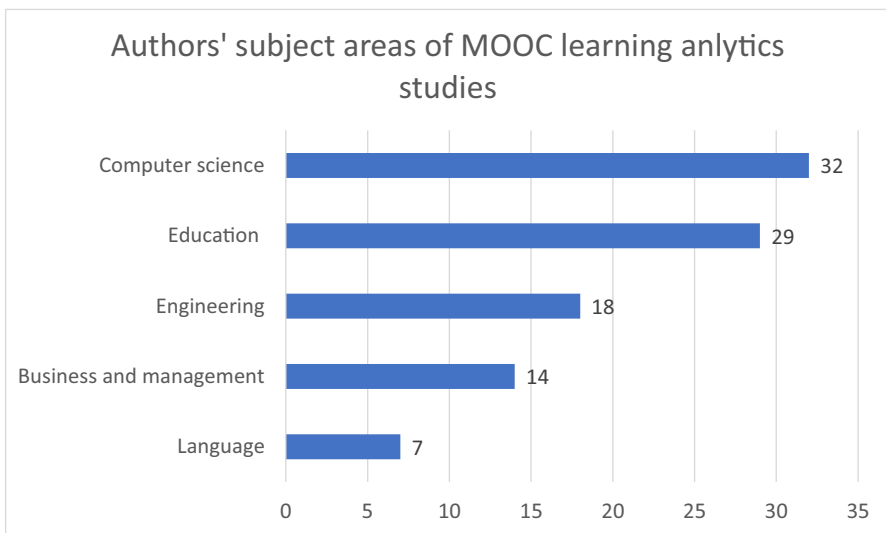
(n = 15), the Netherlands (n = 11), Germany (n = 7), and Canada (n = 6) (see Fig. 10). Therefore, most studies were conducted by authors from developed countries, with no authors from developing countries except China.

The number of authors who collaborated in each study also varied, but more than 92% of the MOOC LA studies were collaborations. Most of the 166 studies had three authors (n = 47), followed by four authors (n = 33), five authors (n = 26), two authors (n = 24), six authors (n = 16), a solo author (n = 12), seven authors (n = 3), nine authors (n = 2), eight authors (n = 1), ten authors (n = 1), and eleven authors (n = 1) (see Fig. 11). The most common collaborations included three or four authors in MOOC LA studies.

Regarding international collaboration, most collaborations were with authors from the same country. Most authors were from one (i.e., the same) country



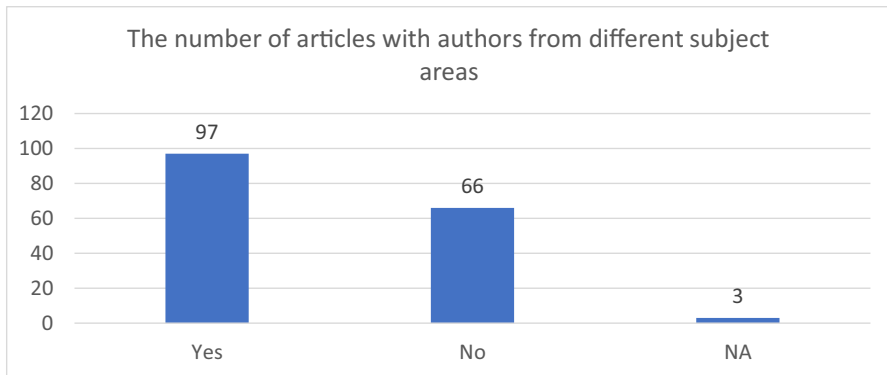
**Fig. 12** The number of studies that had authors from one or more countries (2011–2021) (n = 166)



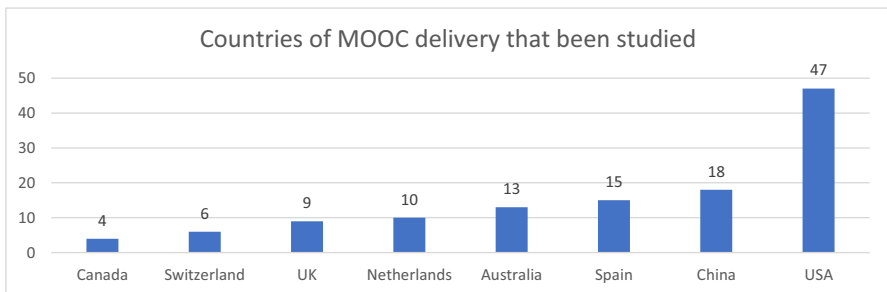
**Fig. 13** The top five subject areas of the authors of MOOC LA studies (2011–2021) (n = 166)

(n = 113), followed by authors collaborating from two countries (n = 41), three countries (n = 6), and four countries (4) (see Fig. 12). Notably, there were two studies with authors from seven or eight countries. In addition, approximately one-third of the MOOC LA studies were international collaborations.

Researchers from computer science and education fields were the primary disciplines conducting MOOC LA studies. Among the 166 studies, the top five subjects of the authors in the reviewed studies were computer science (n = 32), education (n = 29), engineering (n = 18), business and management (n = 14), and language (n = 7) (see Fig. 13). Moreover, over 58% of studies (n = 97) involved



**Fig. 14** The number of articles with authors from different subject areas (2011–2021) (n = 166)



**Fig. 15** The top eight countries of MOOC delivery that have been studied in the MOOCs LA studies (2011–2021) (n = 166)

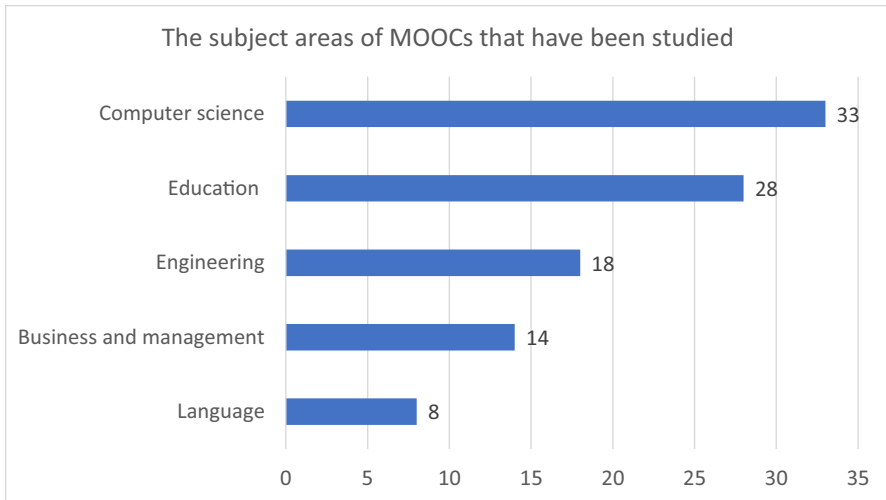
interdisciplinary collaborations (see Fig. 14), indicating that the majority of MOOC LA studies involve interdisciplinary collaborations.

#### **4.5 RQ #5 Which countries and subjects were studied most in MOOC research in the past eleven years?**

In terms of the countries that were the targets of the most MOOC research, the top eight countries included the USA (n = 47), followed by China (n = 18), Spain (n = 15), Australia (n = 13), the Netherlands (n = 10), the UK (n = 9), Switzerland (n = 6), and Canada (n = 4) (see Fig. 15). This corresponded with the locations of the authors' institutions. Developed countries and China were the targets of all of the MOOC LA research in the 166 reviewed studies.

In terms of the subject areas with the most MOOC research, the top five MOOC subjects were computer science (n = 33), followed by education (n = 28), engineering (n = 18), business and management (n = 14), and language (n = 8) (see Fig. 16). This aligns with the authors' subject backgrounds of computer science and education being the most popular. It makes sense that the authors from different subject





**Fig. 16** The top five subject areas that have been studied in MOOC LA studies (2011–2021) (n = 166)

areas would teach MOOCs in their subjects and subsequently conduct research on these topics.

## 5 Discussion

The nature of MOOCs with massive open access provides a good justification for utilizing LA to capture learners' learning behavior to evaluate or improve their learning process (Coffrin et al., 2014; Wibawa et al., 2021). Since 2011, the number of MOOC learning analytics studies has steadily increased with a sharp peak in 2016 (see Fig. 2). This finding is supported by a study conducted by Khalil and Ebner (2016) which found that the combined terms of Learning Analytics and MOOCs were the most cited in Google Scholar between 2013 and 2015. Both MOOCs and LA gained considerable attention in these years (Zhu et al., 2020; Shi & Cristea, 2018). Furthermore, based on our search, the number of articles published in 2017 decreased but gradually increased again until 2020. Although DeMatthews et al. (2020) stated that the research in education slowed down due to the lack of access to data, MOOC LA studies increased as more learners enrolled in MOOCs during the pandemic, and the log data became easier for researchers and educators to access. The ending point of our data collection for this paper was April 2021, which explains why only seven articles were found in 2021. We expect a much larger number for the entire year of 2021.

Our findings show that 72 different journals and proceedings published the 166 MOOC LA articles in this study. There are several possible explanations for this wide range of journals and conference proceedings. The fields of MOOC and LA are still in the initial phase, and few journals or conferences specifically focus on this field. In addition, LA emerged from different fields, including “business

intelligence, web analytics, educational data mining, and recommender systems” (Ferguson, 2012, p.304). Additionally, MOOC research is usually categorized under the umbrella of online learning.

It is also important to highlight that almost one-third (54 articles) of the 166 reviewed articles were published in journals or proceedings that are not assigned to any quartile rank (Q1-Q4). Fifty-four of the articles were published in the Learning Analytics and Knowledge (LAK) conference proceedings, which has not been assigned any quartile rank by the Scimago Journal Ranking. However, this conference is quite established and managed by the Society for Learning Analytics Research (SOLAR), which has members and is supported by many well-known institutions around the world.

Among the 166 MOOC LA research studies that were published in the past 11 years, a majority of the studied used LA for conducting research purposes rather than directly improving teaching and learning practices (see Fig. 5). Given the emerging needs for MOOCs and large-scale online courses, it is critical to put more effort into leveraging LA to improve MOOC teaching practices. In addition, among the 166 reviewed studies, almost two-thirds (61%) used quantitative methods, and 37% used mixed methods. This finding is consistent with the nature of MOOC learning analytics data, which mostly deal with numbers and mathematics, and statistics plays a large role in the data analysis techniques (see Fig. 8). The data used in the reviewed studies include log data (e.g., click numbers, view numbers), achievement data (completed tasks or exams), demographic data, survey, and assignment/quizzes scores (see Fig. 7). Although MOOCs provide considerable data with easy access that facilitates LA, data privacy and ethics issues demand more attention in future MOOC LA studies.

The primary stakeholders of MOOC LA research studies are the parties that benefit the most from the findings of these studies. Based on our findings, instructors and learners were the primary stakeholders for these studies, which explain why these stakeholders have been the focus of many MOOCs and LA articles (Foster & Francis, 2020; Ifenthaler & Yau, 2020; Kew & Tasir, 2020; Sønderslund et al., 2019; Zhu et al., 2020). Given that researchers have been interested in understanding the high dropout rates of tertiary learners, LA can help us understand these learners’ behavior (Kew & Tasir, 2020; Sønderslund et al., 2019). For instructors, LA has been used as part of their teaching reflection and to increase their awareness of students’ learning (Dazo et al., 2017). Instructors can benefit from the LA data as support to improve instruction and decide whether or not the students need early intervention (Erdemci & Karal, 2020). LA data may as well be used to assist MOOC designers and providers regarding how to refine their courses (Shukor & Abdullah, 2019). However, this study found that the studies primarily used LA for research rather than to directly inform teachers about interventions to improve teaching and learning. While on the contrary, the pair between researchers and instructional designers can steer to the LA approach and interventions that closely fit with the specific learning environment (Erkan et al., 2019). Ifenthaler and Yau (2020) explained that a benefit of LA is to maximize the learning experience. Thus, more research focusing on using LA interventions for educational practices in the future will be beneficial.

Prior research has found that LA researchers come from multidisciplinary subjects (Suthers & Verbert, 2013; Waheed et al., 2018), and some have conducted international collaborative research on LA (Society for Learning Analytics Research, n.d.). Our findings support these previous findings. Almost all of the researchers were from Western countries, including the US, Australia, Spain, and the UK (see Fig. 10). China was the only Asian country, among the top eight countries, in our studies to publish MOOC LA studies in English. Our finding suggests that there still remains a wide gap of researchers and publications numbers between countries in North America, Europe, and China with the rest of the world. Murugesan et al. (2017) summarize several factors that may cause the research and publication gaps between these countries, such as a short supply of mentors and funding, lack of writing abilities, and insufficient knowledge of publishing practices. Furthermore, we suggest increasing more opportunities for research collaboration and mentoring in the field of MOOC LA to bridge these gaps.

Although most of the studies were collaborations with researchers from the same country, one-third of the studies were collaborations with researchers from two or more countries (see Fig. 12), and more than half of the studies were interdisciplinary (see Fig. 14). Despite the fact that LA is used far more to analyze teaching and learning in the educational field, the research process often requires people with backgrounds from computer science or engineering fields to collect and analyze the LA data. Moreover, the research found that computer science, education, and engineering are the primary subject areas that are being studied in MOOC LA studies. Thus, collaboration among researchers from different subject areas is recommended. Researchers and institutions benefit from research collaboration, such as knowledge exchange and expertise transfer among researchers, cross-fertilization of research ideas, and a deeper understanding of how to approach a research area. Thus, expanding researchers' networks is needed to enhance our understanding of the outcomes of MOOC teaching and learning (Carroll et al., 2010; Gorska et al., 2020; John-Steiner et al., 1998).

Regarding the targets of most MOOC research, this study found that developed countries and China were the targets of all of the MOOC LA research in the 166 reviewed studies. This finding resonates with our prior systematic review's findings that the MOOCs from developed countries were dominantly studied (Zhu et al., 2020). MOOCs that are offered using non-English and created through regional initiatives get less recognition from mass media and scientific journals (King et al., 2018; Launois et al., 2019; Murugesan et al., 2017; Ruipérez-Valiente et al., 2022). As a consequence, less frame of reference of MOOC LA are able to get from these countries. Therefore, more MOOC LA research could shift to the Global South (Zhang et al., 2020). This finding showed here could help the developing countries to provide funding on MOOC LA related research and teaching practices.

## 6 Limitations and future research

Despite covering multiple MOOCs and researchers across the globe in the analysis, there are some limitations to this study. First, only peer-reviewed journal articles and conference proceedings from *Conference on Learning Analytics and Knowledge*

were included. Thus, some important discussions of MOOC LA in book chapters, institutional reports, newspapers, and dissertations could be missing from our systematic review. Future research could expand the outlets for more in-depth results. Second, future systematic reviews in this topic may also be expanded to add more research papers beyond the 2011–2021 time range to enrich the finding of MOOC LA practice. In addition, only publications in English were included in the review. Thus, meaningful research articles published in other languages (e.g., Chinese, Spanish) may have been overlooked in this research. Future systematic reviews could include articles published in diverse languages to obtain a more comprehensive picture of the MOOC LA status and trends worldwide, which could offer further insights to global MOOC LA research communities.

## 7 Conclusions

LA has been used extensively in education, in general, but has been used only recently in MOOCs over the past decade. A systematic review of MOOC LA has been needed to help us understand the emerging trends of this growing area of education. This study presents a systematic review of MOOC LA studies published from 2011 until the end of April 2021. This study shows that within an approximately ten years range, there have been only around 166 LA empirical studies that have been conducted in a MOOC setting. Given the large expansion of MOOCs, the number of studies in the MOOC LA field is relatively small and new. The number of MOOCs and MOOC learners is still growing rapidly, emphasizing the need for more research on MOOC LA so we can better understand the needs of learners. High-quality journals are still the first option for MOOC LA researchers to disseminate their work, and more than half of the articles were published in prestigious journals (Q1 and Q2 index). However, as MOOC LA research increases, we hope that more research journals will recognize the field and provide publication outlets for this important area. Current MOOC LA studies have primarily been used for research purposes. Thus, future studies on MOOC LA interventions can help improve learning and teaching practices, such as MOOC instructional design, self-directed learning activities, and effective and efficient assessments, etc. Fourth, in terms of the countries and subject areas of the MOOC locations and authors, the US and European countries and researchers in science fields still dominate this research. China is the only Asian country, among the top eight countries, publishing top research in MOOC LA in English. More research from various parts of Asian and African regions and especially from non-science fields are strongly encouraged. These new perspectives can offer richer discussion and valuable decisions for learning. In addition, MOOC LA research requires knowledge and skills from education, computer science, and data science fields. Thus, the active interdisciplinary collaboration will increase the rigor of the studies and the dissemination of the knowledge.

**Acknowledgements** Not applicable

**Funding** No funding.

**Data Availability** The datasets used and/or analyzed during the current study are not publicly available due to their personal and private nature but are available from the corresponding author on reasonable request.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Albelbisi, N. A., & Yusop, F. D. (2020). Systematic review of a Nationwide Mooc initiative in Malaysian higher education system. *Electronic Journal of e-Learning*, 18(4), 287–298. <https://doi.org/10.34190/EJEL.20.18.4.002>
- Albelbisi, N., Yusop, F. D., & Salleh, U. K. M. (2018). Mapping the factors influencing success of massive open online courses (MOOC) in higher education. *Eurasia Journal of Mathematics, Science and Technology Education*, 14(7), 2995–3012. <https://doi.org/10.29333/ejmste/91486>
- Algayres, M. G., & Triantafyllou, E. (2020). Learning Analytics in Flipped Classrooms: A Scoping Review. *Electronic Journal of e-Learning*, 18(5), 397–409. <https://doi.org/10.34190/EJEL.18.5.003>
- Alonso-Fernandez, C., Calvo-Morata, A., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2019). Applications of data science to game learning analytics data: A systematic literature review. *Computers & Education*, 141, 103–612. <https://doi.org/10.1016/j.compedu.2019.103612>
- Alonso-Mencia, M. E., Alario-Hoyos, C., Maldonado-Mahauad, J., Estévez-Ayres, I., Pérez-Sanagustín, M., & Delgado Kloos, C. (2020). Self-regulated learning in MOOCs: Lessons learned from a literature review. *Educational Review*, 72(3), 319–345. <https://doi.org/10.1080/00131911.2019.1566208>
- Al-Rahmi, W., Aldraiweesh, A., Yahaya, N., Kamin, Y. B., & Zeki, A. M. (2019). Massive open online courses (MOOCs): Data on higher education. *Data in Brief*, 22, 118–125. <https://doi.org/10.1016/j.dib.2018.11.139>
- Alturkistani, A., Lam, C., Foley, K., Stenfors, T., Blum, E. R., Van Velthoven, M. H., & Meinert, E. (2020). Massive open online course evaluation methods: Systematic review. *Journal of Medical Internet Research*, 22(4), e13851. <https://doi.org/10.2196/13851>
- Asamoah, D. A., Sharda, R., Hassan Zadeh, A., & Kalgotra, P. (2017). Preparing a data scientist: A pedagogic experience in designing a big data analytics course. *Decision Sciences Journal of Innovative Education*, 15(2), 161–190. <https://doi.org/10.1111/dsji.12125>
- Avella, J. T., Kebritchi, M., Nunn, S. G., & Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Online Learning*, 20(2), 13–29. <https://files.eric.ed.gov/fulltext/EJ1105911.pdf>
- Baker, R. S., & Siemens, G. (2015). Educational data mining and learning analytics. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences (2nd ed)* (pp. 253–272). Cambridge University Press.
- Berland, M., Baker, R. S., & Bilkstein, P. (2014). Educational data mining and learning analytics: Applications to constructionist research. *Technology, Knowledge and Learning*, 19(1–2), 205–220. <https://doi.org/10.1007/s10758-014-9223-7>

- Bezerra, L. N., & Silva, M. T. (2017). A review of literature on the reasons that cause the high dropout rates in the MOOCS. *Revista Espacios*, 38(05). <https://www.revistaespacios.com/a17v38n05/17380511.html>
- Blumenstein, M. (2020). Synergies of Learning Analytics and Learning Design: A Systematic Review of Student Outcomes. *Journal of Learning Analytics*, 7(3), 13–32. <https://doi.org/10.18608/jla.2020.73.3>
- Calonge, D. S., & Shah, M. A. (2016). MOOCs, graduate skills gaps, and employability: A qualitative systematic review of the literature. *The International Review of Research in Open and Distributed Learning*, 17(5), 67–90. <https://doi.org/10.19173/irrodl.v17i5.2675>
- Carroll, J. K., Albada, A., Farahani, M., Lithner, M., Neumann, M., Sandhu, H., & Shepherd, H. L. (2010). Enhancing international collaboration among early career researchers. *Patient Education and Counseling*, 80(3), 417–420. <https://doi.org/10.1016/j.pec.2010.06.020>
- Coffrin, C., Corrin, L., de Barba, P., & Kennedy, G. (2014, March). Visualizing patterns of student engagement and performance in MOOCs. In *Proceedings of the fourth international conference on learning analytics and knowledge* (83–92). <https://doi.org/10.1145/2567574.2567586>
- Costa, L. A., Pereira Sanches, L. M., Rocha Amorim, R. J., Nascimento Salvador, L. D., & Santos Souza, M. V. D. (2020). Monitoring academic performance based on learning analytics and ontology: A systematic review. *Informatics in Education*, 19(3), 361–397. <https://doi.org/10.15388/infedu.2020.17>
- Colvin, C., Rodgers, T., Wade, A., Dawson, S., Gasevic, D., Buckingham Shum, S., et al. (2015). *Student retention and learning analytics: A snapshot of Australian practices and a framework for advancement*. Canberra, ACT: Australian Government Office for Learning and Teaching.
- Creed-Dikeogu, G., & Clark, C. (2013). Are you MOOC-ing yet? A review for academic libraries. *Kansas Library Association College and University Libraries Section Proceedings*, 3(1), 9–13.
- Crescenzi-Lanna, L. (2020). Multimodal Learning Analytics research with young children: A systematic review. *British Journal of Educational Technology*, 51(5), 1485–1504. <https://doi.org/10.1111/bjet.12959>
- Dalipi, F., Imran, A. S., & Kastrati, Z. (2018, April). MOOC dropout prediction using machine learning techniques: Review and research challenges. In *2018 IEEE Global Engineering Education Conference (EDUCON)* (1007–1014). IEEE. <https://doi.org/10.1109/EDUCON.2018.8363340>
- Dazo, S. L., Stepanek, N. R., Chauhan, A., & Dorn, B. (2017, May). Examining instructor use of learning analytics. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (2504–2510). <https://doi.org/10.1145/3027063.3053256>
- DeMatthews, D., Knight, D., Reyes, P., Benedict, A., & Callahan, R. (2020). From the field: Education research during a pandemic. *Educational Researcher*, 49(6), 398–402. <https://doi.org/10.3102/0013189X20938761>
- Drachsler, H., & Kalz, M. (2016). The MOOC and learning analytics innovation cycle (MOLAC): A reflective summary of ongoing research and its challenges. *Journal of Computer Assisted Learning*, 32(3), 281–290. <https://doi.org/10.1111/jcal.12135>
- Erdemci, H., & Karal, H. (2020). Examination of instructors' experiences for the use of learning analytics. *The International Journal of Information and Learning Technology*, 38(1), 21–31. <https://doi.org/10.1108/IJILT-05-2020-0076>
- Erkan, E., Gómez-Sánchez, E., Dimitriadis, Y., Bote-Lorenzo, M. L., Asensio-Pérez, J. I., & Álvarez-Álvarez, S. (2019). Aligning learning design and learning analytics through instructor involvement: A MOOC case study. *Interactive Learning Environments*, 27(5–6), 685–698. <https://doi.org/10.1080/10494820.2019.1610455>
- Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5–6), 304–317. <https://doi.org/10.1504/IJTEL.2012.051816>
- Foster, C., & Francis, P. (2020). A systematic review on the deployment and effectiveness of data analytics in higher education to improve student outcomes. *Assessment & Evaluation in Higher Education*, 45(6), 822–841. <https://doi.org/10.1080/02602938.2019.1696945>
- Gedrimiene, E., Silvola, A., Pursiainen, J., Rusanen, J., & Muukkonen, H. (2020). Learning analytics in education: Literature review and case examples from vocational education. *Scandinavian Journal of Educational Research*, 64(7), 1105–1119. <https://doi.org/10.1080/00313831.2019.1649718>
- Gorska, A., Korzynski, P., Mazurek, G., & Pucciarelli, F. (2020). The role of social media in scholarly collaboration: An enabler of international research team's activation? *Journal of Global*

- Information Technology Management*, 23(4), 273–291. <https://doi.org/10.1080/1097198X.2020.1817684>
- Glick, D., Cohen, A., Festinger, E., Xu, D., Li, Q., & Warschauer, M. (2019). Predicting success, preventing failure. In D. Ifenthaler, D.-K. Mah, & J.Y.-K. Yau (Eds.), *Utilizing learning analytics to support study success* (pp. 249–273). Springer.
- Hui, Y. K., & Kwok, L. F. (2019). A review on learning analytics. *International Journal of Innovation and Learning*, 25(2), 197–222. <https://doi.org/10.1504/IJIL.2019.097673>
- Hidalgo, F. J. P., & Abril, C. A. H. (2020). MOOCs: Origins, concept and didactic applications: A systematic review of the literature (2012–2019). *Technology, Knowledge and Learning*, 25(4), 853–879. <https://doi.org/10.1007/s10758-019-09433-6>
- Hwang, G. J., Hung, P. H., Chen, N. S., & Liu, G. Z. (2014). Mindtool-assisted in-field learning (MAIL): An advanced ubiquitous learning project in Taiwan. *Educational Technology & Society*, 17(2), 4–16. <https://www.jstor.org/stable/pdf/jeductechsoci.17.2.4.pdf>
- Ifenthaler, D., Mah, D.-K., & Yau, J.Y.-K. (2019). Utilising learning analytics for study success. Reflections on current empirical findings. In D. Ifenthaler, J.Y.-K. Yau, & D.-K. Mah (Eds.), *Utilizing learning analytics to support study success* (pp. 27–36). Springer.
- Ifenthaler, D., & Tracey, M. W. (2016). Exploring the relationship of ethics and privacy in learning analytics and design: Implications for the field of educational technology. *Educational Technology Research and Development*, 64(5), 877–880. <https://doi.org/10.1007/s11423-016-9480-3>
- Ifenthaler, D., & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. *Educational Technology Research and Development*, 64(5), 923–938. <https://doi.org/10.1007/s11423-016-9477-y>
- Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*, 68(4), 1961–1990. <https://doi.org/10.1007/s11423-020-09788-z>
- John-Steiner, V., Weber, R. J., & Minnis, M. (1998). The challenge of studying collaboration. *American Educational Research Journal*, 35(4), 773–783. <https://doi.org/10.3102/00028312035004773>
- Joksimović, S., Poquet, O., Kovanović, V., Dowell, N., Mills, C., Gašević, D., ... & Brooks, C. (2018). How do we model learning at scale? A systematic review of research on MOOCs. *Review of Educational Research*, 88(1), 43–86. <https://doi.org/10.3102/0034654317740335>
- Jona, K., & Naidu, S. (2014). MOOCs: Emerging research. *Distance Education*, 35(2), 141–144. <https://doi.org/10.1080/01587919.2014.928970>
- Kew, S. N., & Tasir, Z. (2021). Learning Analytics in Online Learning Environment: A Systematic Review on the Focuses and the Types of Student-Related Analytics Data. *Technology, Knowledge and Learning*, 1–23. <https://doi.org/10.1007/s10758-021-09541-2>
- Khalil, M., & Ebner, M. (2016). What is learning analytics about? A survey of different methods used in 2013–2015. *arXiv preprint arXiv:1606.02878*.
- King, M., Luan, B., & Lopes, E. (2018). Experiences of Timorese language teachers in a blended massive open online course (MOOC) for continuing professional development (CPD). *Open Praxis*, 10(3), 279–287. <https://doi.org/10.5944/openpraxis.10.3.840>
- Knobbout, J., & Van Der Stappen, E. (2020). Where is the learning in learning analytics? A systematic literature review on the operationalization of learning-related constructs in the evaluation of learning analytics interventions. *IEEE Transactions on Learning Technologies*, 13(3), 631–645. <https://doi.org/10.1109/TLT.2020.2999970>
- Kop, R. (2011). The challenges to connectivist learning on open online networks: Learning experiences during a massive open online course. *The International Review of Research in Open and Distributed Learning*, 12(3), 19–38. <http://www.irrod.org/index.php/irrod/article/view/882/1689>
- Korkmaz, C., & Correia, A. P. (2019). A review of research on machine learning in educational technology. *Educational Media International*, 56(3), 250–267. <https://doi.org/10.1080/09523987.2019.1669875>
- Lambert, S. R. (2020). Do MOOCs contribute to student equity and social inclusion? A systematic review 2014–18. *Computers & Education*, 145, 103693. <https://doi.org/10.1016/j.compedu.2019.103693>
- Larrabee Sønderlund, A., Hughes, E., & Smith, J. (2019). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology*, 50(5), 2594–2618. <https://doi.org/10.1111/bjet.12720>
- Launois, P., Allotey, P., Reidpath, D., Maher, D., Certain, E., & Ross, B. (2019). Lessons learnt from a professional development MOOC: Engaging culturally and linguistically diverse learners from low-and middle-income countries. *European Journal of Open, Distance and E-learning*, 22(2).

- Lee, D., Watson, S. L., & Watson, W. R. (2019). Systematic literature review on self-regulated learning in massive open online courses. *Australasian Journal of Educational Technology*, 35(1). <https://doi.org/10.14742/ajet.3749>
- Leitner, P., Khalil, M., & Ebner, M. (2017). Learning analytics in higher education—A literature review. In A. Pe a-Ayala (Ed.), *Learning analytics: Fundamentals, applications, and trends* (1–23). Springer. [https://doi.org/10.1007/978-3-319-52977-6\\_1](https://doi.org/10.1007/978-3-319-52977-6_1)
- Guajardo Leal, B. E., Navarro-Corona, C., & Valenzuela González, J. R. (2019). Systematic mapping study of academic engagement in MOOC. *International Review of Research in Open and Distributed Learning*, 20(2). <https://doi.org/10.19173/irrodl.v20i2.4018>
- Jantti, M., & Heath, J. (2016). What role for libraries in learning analytics? *Performance Measurement and Metrics*, 17(2), 203–210. <https://doi.org/10.1108/PMM-04-2016-0020>
- Lemay, D. J., & Doleck, T. (2020). Grade prediction of weekly assignments in MOOCs: Mining video-viewing behavior. *Education and Information Technologies*, 25(2), 1333–1342. <https://doi.org/10.1007/s10639-019-10022-4>
- Li, X., Chen, Y., & Gong, X. (2017). MOOCs in China: A review of literature, 2012–2016. *New ecology for education—communication X learning*, 21–32. [https://doi.org/10.1007/978-981-10-4346-8\\_3](https://doi.org/10.1007/978-981-10-4346-8_3)
- 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>
- Martin, F., Dennen, V. P., & Bonk, C. J. (2020). A synthesis of systematic review research on emerging learning environments and technologies. *Educational Technology Research and Development*, 68(4), 1613–1633. <https://doi.org/10.1007/s11423-020-09812-2>
- Matcha, W., Gašević, D., & Pardo, A. (2019). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, 13(2), 226–245. <https://doi.org/10.1109/TLT.2019.2916802>
- Min, H., & Nasir, M. K. M. (2020). Self-Regulated Learning in A Massive Open Online Course: A Review of Literature. *European Journal of Interactive Multimedia and Education*, 1(2), 1–6. <https://doi.org/10.30935/ejimed/8403>
- Mubarak, A. A., Cao, H., & Ahmed, S. A. (2021). Predictive learning analytics using deep learning model in MOOCs' courses videos. *Education and Information Technologies*, 26(1), 371–392. <https://doi.org/10.1007/s10639-020-10273-6>
- Murugesan, R., Nobes, A., & Wild, J. (2017). A MOOC approach for training researchers in developing countries. *Open Praxis*, 9(1), 45–57. <https://www.learntechlib.org/p/181406/>
- Mota, R., & Scott, D. (2014). *Education for innovation and independent learning*. Elsevier.
- Okoli, C. (2015). A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems Research*, 37(43), 879–910. <https://doi.org/10.17705/ICAIS.03743>
- Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. *Sprouts: Working Paper on Information Systems*, 10(26), 1–49.
- Papamitsiou, Z. K., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology & Society*, 17(4), 49–64. <https://www.jstor.org/stable/pdf/jeductechsoci.17.4.49.pdf>
- Paton, R. M., Fluck, A. E., & Scanlan, J. D. (2018). Engagement and retention in VET MOOCs and online courses: A systematic review of literature from 2013 to 2017. *Computers & Education*, 125, 191–201. <https://doi.org/10.1016/j.compedu.2018.06.013>
- Pishtari, G., Rodríguez-Triana, M. J., Sarmiento-Márquez, E. M., Pérez-Sanagustín, M., Ruiz-Calleja, A., Santos, P., ... & Váljataga, T. (2020). Learning design and learning analytics in mobile and ubiquitous learning: A systematic review. *British Journal of Educational Technology*, 51(4), 1078–1100. <https://doi.org/10.1111/bjet.12944>
- Prinsloo, P., & Slade, S. (2017). An elephant in the learning analytics room. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference on—LAK'17*, 46–55. <https://doi.org/10.1145/3027385.3027406>
- Rincón-Flores, E. G., Montoya, M. S. R., & Mena, J. (2019, October). Engaging MOOC through gamification: Systematic mapping review. In *Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality* (600–606). <https://doi.org/10.1145/3362789.3362831>



- Rolfé, V. (2015). A systematic review of the socio-ethical aspects of Massive Online Open Courses. *European Journal of Open, Distance and E-Learning (EURODL)*, 18(1), 52–71. <https://www.ceeol.com/search/article-detail?id=848923>
- Rømer, T., Hansen, M. T., & Helge, J. W. (2020). An analysis of the productivity and impact of clinical PhD theses from the University of Copenhagen. *Danish Medical Journal*, 67(5), A12190731.
- Ruipérez-Valiente, J. A., Staubitz, T., Jenner, M., Halawa, S., Zhang, J., Despujol, I., ... & Reich, J. (2022). Large scale analytics of global and regional MOOC providers: Differences in learners' demographics, preferences, and perceptions. *Computers & Education*, 180, 104426. <https://doi.org/10.1016/j.compedu.2021.104426>
- Sallam, M. H., Martín-Monje, E., & Li, Y. (2020). Research trends in language MOOC studies: A systematic review of the published literature (2012–2018). *Computer Assisted Language Learning*, 1–28. <https://doi.org/10.1080/09588221.2020.1744668>
- Samuelsen, J., Chen, W., & Wasson, B. (2019). Integrating multiple data sources for learning analytics—review of literature. *Research and Practice in Technology Enhanced Learning*, 14(1), 1–20. <https://doi.org/10.1186/s41039-019-0105-4>
- Sari, A. R., Bonk, C. J., & Zhu, M. (2020). MOOC instructor designs and challenges: What can be learned from existing MOOCs in Indonesia and Malaysia?. *Asia Pacific Education Review*, 21(1), 143–166. <https://doi.org/10.1007/s12564-019-09618-9>
- Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., ... & Dillenbourg, P. (2016). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*, 10(1), 30–41. <https://doi.org/10.1109/TLT.2016.2599522>
- Slater, N., Peasgood, A., & Mullan, J. (2016). *Learning analytics in higher education: A review of UK and international practice*. Bristol: JISC.
- Shah, D. (2020, December 14). *The Second Year of the MOOC: A Review of MOOC Stats and Trends in 2020*. Class Central. <https://www.classcentral.com/report/the-second-year-of-the-mooc/>
- Shi, L., & Cristea, A. I. (2018, January). Demographic indicators influencing learning activities in MOOCs: Learning analytics of futurelearn courses. In *Proceedings of the 27th International Conference on Information Systems Development: Designing Digitalization, ISD 2018*. <https://dro.dur.ac.uk/25774/>
- Shukor, N. A., & Abdullah, Z. (2019). Using learning analytics to improve MOOC instructional design. *International Journal of Emerging Technologies in Learning (iJET)*, 14(24), 6–17. <https://www.learntechlib.org/p/217038/>
- Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46, 30–32. <https://doi.org/10.17471/2499-4324/195>
- Siemens, G. (2012). Learning analytics: envisioning a research discipline and a domain of practice. *LAK'12: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. <https://doi.org/10.1145/2330601.2330605>
- Society for Learning Analytics Research. (n.d.). <https://www.solaresearch.org/about/>.
- Sønderlund, A. L., Hughes, E., & Smith, J. (2019). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology*, 50(5), 2594–2618. <https://doi.org/10.1111/bjet.12720>
- Suthers, D., & Verbert, K. (2013, April). Learning analytics as a "middle space". In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (1–4). <https://doi.org/10.1145/2460296.2460298>
- Taneja, S., & Goel, A. (2014). MOOC providers and their strategies. *International Journal of Computer Science and Mobile Computing*, 3(5), 222–228.
- Tenório, T., Bittencourt, I. L., Isotani, S., & Silva, A. P. (2016). Does peer assessment in on-line learning environments work? A systematic review of the literature. *Computers in Human Behavior*, 64, 94–107. <https://doi.org/10.1016/j.chb.2016.06.020>
- Tóth, J., & Demeter, M. (2021). Prestige and Independence-Controlled Publication Performance of Researchers at 14 Hungarian Research Institutions between 2014 and 2018: A Data Paper. *KOME: an International Journal of Pure Communication Inquiry*, 9(1), 41–63. <https://doi.org/10.17646/KOME.75672.61>
- Tashakkori, A., & Teddlie, C. (2003). *Handbook on mixed methods in the behavioral and social sciences*. Sage.

- Tzimas, D., & Demetriadis, S. (2021). Ethical issues in learning analytics: a review of the field. *Educational Technology Research and Development*, 69(2), 1101–1133. <https://doi.org/10.1007/s11423-021-09977-4>
- Van de Oudewetering, K., & Agirdag, O. (2018). MOOCs as accelerators of social mobility? A systematic review. *Journal of Educational Technology & Society*, 21(1), 1–11. <https://www.jstor.org/stable/26273863>
- Veenman, M. V. J. (2013). Assessing metacognitive skills in computerized learning environments. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 157–168). Springer.
- Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education*, 122, 119–135. <https://doi.org/10.1016/j.compedu.2018.03.018>
- Waheed, H., Hassan, S. U., Aljohani, N. R., & Wasif, M. (2018). A bibliometric perspective of learning analytics research landscape. *Behaviour & Information Technology*, 37(10–11), 941–957. <https://doi.org/10.1080/0144929X.2018.1467967>
- Wei, X., Saab, N., & Admiraal, W. (2020). Assessment of cognitive, behavioral, and affective learning outcomes in massive open online courses: A systematic literature review. *Computers & Education*, 163, 104097. <https://doi.org/10.1016/j.compedu.2020.104097>
- Wibawa, B., Siregar, J. S., Asrorie, D. A., & Syakdiyah, H. (2021, April). Learning analytic and educational data mining for learning science and technology. In *AIP Conference Proceedings* (Vol. 2331, No. 1, p. 060001). AIP Publishing LLC.
- Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G. J., & Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A systematic review. *International Journal of Human-Computer Interaction*, 35(4–5), 356–373. <https://doi.org/10.1080/10447318.2018.1543084>
- Wong, A., & Chong, S. (2018). Modelling adult learners' online engagement behaviour: Proxy measures and its application. *Journal of Computers in Education*, 5(4), 463–479. <https://doi.org/10.1007/s40692-018-0123-z>
- Yahoo. (2021, February 25). *Global MOOCs (Massive Open Online Course) Market 2021–2025: Market is Poised to Grow by \$16.01 Billion, Progressing at a CAGR of 32%*. Yahoo. Retrieved from <https://www.yahoo.com/now/global-moocs-massive-open-online-150900130.html>
- Zainuddin, G., Danuri, M. S. N., Ali, A. M., Ahmad, M. I., Sokri, N. E. A., Jaffar, M. N., ... & Hashim, H. (2019). A systematic literature review on massive open online course for language learning. *Creative Education*, 10(12), 3195. <https://doi.org/10.4236/ce.2019.1012243>
- Zhang, K., Bonk, C. J., Reeves, T. C., & Reynolds, T. H. (Eds.). (2020). *MOOCs and open education in the Global South: Challenges, successes, and opportunities*. NY: Routledge. <https://doi.org/10.4324/9780429398919>.
- Zhu, M. (2021). Enhancing MOOC learners' skills for self-directed learning. *Distance Education*, 42(3), 441–460. <https://doi.org/10.1080/01587919.2021.1956302>
- Zhu, M., Bonk, C., & Sari, A. (2019). Massive open online course instructor motivations, innovations, and designs: Surveys, interviews, and course reviews. *Canadian Journal of Learning and Technology*, 45(1). <https://www.learntechlib.org/p/208592/>
- Zhu, M., Sari, A., & Lee, M. M. (2018). A systematic review of research methods and topics of the empirical MOOC literature (2014–2016). *The Internet and Higher Education*, 37, 31–39. <https://doi.org/10.1016/j.iheduc.2018.01.002>
- Zhu, M., Sari, A. R., & Lee, M. M. (2020). A comprehensive systematic review of MOOC research: Research techniques, topics, and trends from 2009 to 2019. *Educational Technology Research and Development*, 68(4), 1685–1710. <https://doi.org/10.1007/s11423-020-09798-x>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.