A Systematic Review of Questionnaire-Based Quantitative Research on MOOCs

Mingxiao Lu, Tianyi Cui, Zhenyu Huang, Hong Zhao, Tao Li et Kai Wang

Résumé de l'article

Massive open online courses (MOOCs) have attracted much interest from educational researchers and practitioners around the world. There has been an increase in empirical studies about MOOCs in recent years, most of which used questionnaire surveys and quantitative methods to collect and analyze data. This study explored the research topics and paradigms of questionnaire-based quantitative research on MOOCs by reviewing 126 articles available in the Science Citation Index (SCI) and Social Sciences Citation Index (SSCI) databases from January 2015 to August 2020. This comprehensive overview showed that: (a) the top three MOOC research topics were the factors influencing learners' performance, dropout rates and continuance intention to use MOOCs, and assessing MOOCs; (b) for these three topics, many studies designed questionnaires by adding new factors or adjustments to extant theoretical models or survey instruments; and (c) most researchers used descriptive statistics to analyze data, followed by the structural equation model, and reliability and validity analysis. This study elaborated on the relationship of research topics and key factors in the research models by building factors-goals (F-G) graphs. Finally, we proposed some directions and recommendations for future research on MOOCs.
A Systematic Review of Questionnaire-Based Quantitative Research on MOOCs

Mingxiao Lu¹, Tianyi Cui², Zhenyu Huang³, Hong Zhao¹, Tao Li¹, and Kai Wang¹
¹College of Computer Science, Nankai University, China; ²Business School, Nankai University, China; ³College of Business Administration, Central Michigan University, USA

Abstract

Massive open online courses (MOOCs) have attracted much interest from educational researchers and practitioners around the world. There has been an increase in empirical studies about MOOCs in recent years, most of which used questionnaire surveys and quantitative methods to collect and analyze data. This study explored the research topics and paradigms of questionnaire-based quantitative research on MOOCs by reviewing 126 articles available in the Science Citation Index (SCI) and Social Sciences Citation Index (SSCI) databases from January 2015 to August 2020. This comprehensive overview showed that: (a) the top three MOOC research topics were the factors influencing learners’ performance, dropout rates and continuance intention to use MOOCs, and assessing MOOCs; (b) for these three topics, many studies designed questionnaires by adding new factors or adjustments to extant theoretical models or survey instruments; and (c) most researchers used descriptive statistics to analyze data, followed by the structural equation model, and reliability and validity analysis. This study elaborated on the relationship of research topics and key factors in the research models by building factors-goals (F-G) graphs. Finally, we proposed some directions and recommendations for future research on MOOCs.

Keywords: MOOC, factors-goals graph (F-G graph), questionnaire-based survey, quantitative analysis, research topics
Introduction

Massive open online courses (MOOCs), an innovative technology-enhanced learning model, have offered educational opportunities to a vast number of learners, and have attracted much interest from educational researchers and practitioners around the world (Zhou, 2016). When COVID-19 suddenly broke out in early 2020, schools in many countries had to be closed to stop the spread of the pandemic according to media reports. MOOCs became a top choice for students studying online from home. Some stakeholders have suggested that MOOCs may have a groundbreaking impact on higher education, potentially making traditional physical universities obsolete (Shirky, 2013). While acknowledging the potential of MOOCs, some educators have expressed concerns about the pedagogical models based on information transmission that have been widely applied in MOOCs (Albert et al., 2015; Babori et al., 2019; Veletsianos & Shepherdson, 2016; Zhu et al., 2018). Despite the polarized debate, the number of MOOC courses offered and students enrolled has continued to grow, which has aroused the interest of researchers. There have been a substantial number of research studies and reports investigating various aspects and effective practices of MOOCs in recent times, some of which have focused on empirical research.

Questionnaire-based surveys can directly and quickly obtain information about the attitudes, behaviors, characteristics, and opinions of MOOC participants, all of which can be used as first-hand data for empirical research. Most questionnaire-based research has made use of measurement scales, with the collected answers quantitatively analyzed to extract value. Researchers considered various factors and used classical models and theories when they designed their questionnaires. Follow-up research is necessary to analyze and summarize this prior work. This paper explored the research topics and paradigms of questionnaire-based quantitative research on MOOCs. The main contribution is a graphical summary of the classical models and theories, as well as analysis of the key factors frequently considered in certain key topics.

Literature Review

Over the years, MOOCs have yielded many research publications and have attracted numerous types of review articles including systematic as well as critical reviews. Zhu et al. (2018) summarized the typical research topics and methods directed to MOOCs, as well as their geographical distribution, by reviewing 146 empirical studies of MOOCs published from 2014 to 2016. They summarized the typical research topics and methods through only a few statistical results in the form of numbers, bar charts, and pie charts. Rasheed et al. (2019) adopted a systematic mapping methodology to provide a fine-grain overview of the MOOC research domain by identifying the quantity and types of research, available results, and publication trends in educational aspects of MOOCs from 2009 to 2018. Their findings showed that most MOOC studies focused on addressing learners’ completion, dropout rates, and retention. Babori et al. (2019) examined the content of MOOC research in 65 peer-reviewed papers produced in five major educational technology research journals between 2012 and 2018. Their analysis revealed that these articles were mainly concerned with MOOCs’ objectives, prerequisites required for participation in MOOCs, and types of learning scenarios. In addition, empirical studies adopted a variety of conceptual frameworks that focused mainly on learning strategies. Montes-Rodriguez et al. (2019) examined the prevalence and characteristics of case studies on MOOCs, based on 92 articles selected from the Web of Science and Scopus. Their findings showed that even when searching solely for case studies, quantitative analysis was more prevalent for data collection and analysis in research on MOOCs.
The reviews cited above showed MOOC research trends and topics as rapidly evolving. Although the majority of early MOOC studies were mostly theoretical and conceptual, more empirical studies and topics have emerged in recent years. According to Fang et al. (2019) and Zhu et al. (2018) most empirical research on MOOCs has used quantitative methods for gathering and analyzing data. As a methodology, quantitative analysis is generally linked to interpretive paradigms that analyze the quantitative characteristics, relations, and changes of social phenomena. A key process in quantitative analysis is that of establishing a mathematical model to calculate various indicators and values of the research object based on statistical data. Therefore, how to effectively collect quantitative data is the basis of this methodology. For research on MOOCs, surveys, especially questionnaire-based surveys, have been the most frequently adopted method of data collection (Sanchez-Gordon & Lujan-Mora, 2017; Zhu et al., 2018).

**Research Questions**

Few studies have reviewed the questionnaire-based quantitative research about MOOCs and summarized the theories such research has been based on. A comprehensive picture of the methodologies adopted in these studies is needed in order to investigate the characteristics of research on MOOCs, including topic areas, theoretical models, and research methods. We reviewed questionnaire-based quantitative studies about MOOCs published from January 2015 to August 2020, in order to increase awareness of methodological issues and theoretical models in the MOOC research field. The following three research questions guided our review:

1. What research topics or focuses have been addressed in questionnaire-based quantitative MOOC studies?
2. What research models have been used for examining the critical topics in these MOOC studies?
3. What analysis methods were most often used in these MOOC studies?

**Research Methodology**

**Data Collection**

By using the keywords MOOC, MOOCs, massive open online course, and massive open online courses, we searched for articles from the Web of Science database as our source data. The attributes of each selected article included authors, title, year of publication, journal name, research focus, research model, analysis methodology, and article URL. We classified research methodologies as qualitative, quantitative, or mixed method (i.e., combining quantitative and qualitative approaches). In this study, we focused on articles with quantitative or mixed method research. We filtered the articles according to six ordered selection criteria, as shown in Table 1. Each criterion is a hard one, which means that an article was filtered out if it did not meet even one criterion. The filtering process comprised reading the title and abstract of each article and assigning a value of relevant or irrelevant. When the relevance was not evident from the title and abstract, we examined the article in detail, reading the methodology and results sections. A total of 126 articles about MOOCs were selected and verified, including 89 quantitative studies and 37 with mixed methods.
Table 1

Criteria for Selecting MOOC Articles

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Operational definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The article was retrieved from the SCI or SSCI database.</td>
</tr>
<tr>
<td>2</td>
<td>The article was published in English.</td>
</tr>
<tr>
<td>3</td>
<td>The article was published between January 2015 and August 2020.</td>
</tr>
<tr>
<td>4</td>
<td>The terms MOOC(s) or massive open online course(s) were used to screen titles, abstracts, and keywords.</td>
</tr>
<tr>
<td>5</td>
<td>The study mainly investigated the educational aspects of MOOCs.</td>
</tr>
<tr>
<td>6</td>
<td>The article reported on an empirical study using questionnaire-based survey data and quantitative analysis.</td>
</tr>
</tbody>
</table>

Data Analysis

To address our first research question, thematic content analysis was used to examine the key research topics in studies of MOOCs. First, researchers read the MOOC research articles and identified the specific research focuses of each paper; topics were then grouped into four categories, namely dropout rates and continuance intention to use MOOCs, learners’ performance, assessing MOOCs, and others. To answer research question two, related to the research models typically employed, we systematically presented the models by means of factors-goals (F-G) graphs. These graphs, which were first designed as a graphic device for this study, showed the correlation between research goals and influencing factors in order to provide a reference framework for building hypothesis models. F-G graphs provided a statistical baseline for accuracy, consistency, and representativeness to improve data quality. Finally, to answer the third research question, researchers counted the data analysis methods most often used in the quantitative studies.

Results and Analysis

Research Topics and Focuses in Questionnaire-Based Quantitative Studies on MOOCs

To examine the general topics and focuses of quantitative MOOC studies, we divided the key topics of 126 papers into four different categories: (a) dropout rates and continuance intention to use MOOCs ($n = 36; 28.57\%$); (b) factors influencing learners’ performance ($n = 45; 35.71\%$); (c) assessing MOOCs ($n = 29; 23.02\%$); and (d) others ($n = 16; 12.70\%$).

Dropout Rates or Continuance Intention to Use MOOCs

MOOCs might not be equally successful in keeping learners through to course completion, though they are successful in attracting and accommodating numerous learners. Some studies showed that only a small number of participants completed an entire course, and others quit partway through after experiencing a
few MOOC lessons (Shao, 2018; Yang et al., 2017). High dropout rates have been widely regarded as a serious issue for MOOCs (Bozkurt et al., 2017).

Most of the extant literature considered completion rate as a metric for evaluating the success or failure of a MOOC. It is vital to investigate the reasons why learners persist and complete their courses or drop out, so a large number of researchers have explored this issue through quantitative methods based on questionnaires. Both subjective and objective factors influenced MOOC participants’ retention and completion. The main subjective factors included learners’ preferences (Li et al., 2018), experience (Li et al., 2018; Zhao et al., 2020; Zhou, 2017), expectancy (Botero et al., 2018; Luik et al., 2019; Zhou, 2017), and psychological motivation (Botero et al., 2018; Yang & Su, 2017; Zhou, 2016). Objective factors included course quality (Hone & El Said, 2016; Yang et al., 2017), network externalities (Li et al., 2018), social motivation (Jung & Lee, 2018; Khan et al., 2018; Wu & Chen, 2017), and MOOC systems (Wu & Chen, 2017).

Factors Affecting MOOC Learners’ Performance

Dropout rate is not the only metric of the success of a MOOC. Learners have various motivations for taking online courses (Carlos et al., 2017), which can affect their attitude and intention to continue learning in MOOCs. The performance of learners attending a MOOC can be used as an essential reference for improving MOOC design and quality. Learners’ performance in MOOCs has been measured by course engagement, social interactions, sociability, and learning gains. Many studies have focused on the factors that influence learners’ performance (Carlos et al., 2017; Kahan et al., 2017; Soffer & Cohen, 2015; Zhang, 2016). From the articles reviewed in this study, we summarized the major factors affecting learners’ performance into four categories: motivation, self-regulated learning (SRL), attitudinal learning, and learning strategies.

Learners with different motivations for participating in a MOOC targeted different learning goals and strategies (de Barba et al., 2016; Watted & Barak, 2018). General participants were oriented toward acquiring knowledge and academic advancement, while university-affiliated students were also concerned with a need to obtain certificates. SRL is a learning strategy that influences MOOC learners’ academic performance. Independent learning in MOOCs calls for completing course content, making full use of platform resources, and allocating study time reasonably (Jansen et al., 2017; Kizilcec et al., 2017; Maldonado-Mahauad et al., 2018). The scale items of attitudinal learning conform to the following four-dimensional theoretical structure: cognitive learning, affective learning, behavioral learning, and social learning (Watson et al., 2016). Finally, learning strategy has been defined as a complex plan for a learning process that learners have purposefully and consciously formulated to improve their learning effectiveness and performance in MOOCs (Kizilcec et al., 2017; Maldonado-Mahauad et al., 2018).

Assessment of MOOCs

Some articles investigated the overall assessment of MOOCs, specifically evaluation of the teaching model, course structure and content design, the MOOC platform technology, and the benefits from participating in MOOCs. We divided the studies we examined into two categories: assessment from the perspective of learners, and assessment from teachers’ points of view. Some student-oriented research used learners’ perceived benefits to determine which course design better helped learners meet their goals (Jung et al., 2019; Lowenthal et al., 2018). Teacher-focused evaluation paid close attention to teaching skills and
challenges in MOOCs, as well as opportunities for future development (Donitsa-Schmidt & Topaz, 2018; Gan, 2018).

**Research Models for Examining the Key Topics in MOOC Studies**

Questionnaire-based quantitative research generally includes the following steps: (a) propose the research questions to be solved; (b) select an appropriate theoretical model or develop a new model, drawing on classical theories and the hypothetical relationship between factors; (c) design questionnaire items, usually in the form of a Likert scale, to measure the factors and variables in the research model; (d) collect the questionnaires from research subjects; and (e) analyze the collected data to verify the hypothesis model.

Building a hypothesis model is the foundation of quantitative research. In examining the sorts of models MOOC researchers have relied on, we described three F-G graphs to depict the correlation between the top three categories of research topics and the research models summarized from the 126 articles. Tables A1, A2, A3 in Appendix A provide background details from 32 typical articles for the top three topics, including article titles, research topics, theoretical models and factors involved in the questionnaire, and analysis methods. This data formed the foundations for drawing the F-G graphs for our study.

**F-G Graph for Learners’ Dropout Rates and Continuance Intention to Use MOOCs**

An F-G graph was built to demonstrate the correlation among the research models and the research goal of investigating factors that affect learners’ intentions to continue to use MOOCs. As shown in Figure 1, the F-G graph integrated the factors of research models frequently used in the articles we examined. The relationship hypotheses between factors are shown with straight arrows. The direction of an arrow points from an explanatory variable to a dependent variable. The factors in rounded rectangles are those that directly or indirectly affected learners’ intentions to continue to use MOOCs.
In the 126 articles we examined, most researchers designed questionnaire items by extending classical theoretical models, including: (a) the technology acceptance model (TAM; \( n = 12 \)); (b) the self-determination theory (SDT; \( n = 7 \)); (c) the task-technology fit (TTF; \( n = 4 \)); (d) the theory of planned behavior (TPB; \( n = 4 \)); (e) the unified theory of acceptance and use of technology (UTAUT; \( n = 3 \)); and (f) the information system (IS) success model (\( n = 2 \)). In Figure 1, the key factors from each model are enclosed within black dotted boxes. In addition, some studies enhanced these models by adding new elements or adjustments to further explain learners’ continuance intention to use MOOCs. In Figure 1, these new factors, often considered by the reviewed articles, are listed in black solid boxes. The specific explanations of these theoretical models are summarized in Table 2 and addressed in detail following the table.
Table 2

Classical Models That Address Dropout Rates or Continuance Intention to Use MOOCs

<table>
<thead>
<tr>
<th>Model</th>
<th>Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAM</td>
<td>Perceived usefulness and perceived ease of use determine the individual’s attitude toward a MOOC as well as the behavioral intention to use it (Davis, 1989).</td>
</tr>
<tr>
<td>SDT</td>
<td>A motivation theory to investigate how and why a particular human behavior occurs. Distinguishes between autonomous and controlled motivations in terms of the degrees of self-determination (Deci et al., 1999).</td>
</tr>
<tr>
<td>TPB</td>
<td>Explains three determinants of individual’s behavioral intentions: perceived behavioral control, subjective norms, and attitude toward the behavior (Ajzen, 1985).</td>
</tr>
<tr>
<td>TTF</td>
<td>Task characteristics and technology characteristics can affect the task-technology fit, which determines users’ performance and utilization (Goodhue et al., 2000).</td>
</tr>
<tr>
<td>UTAUT</td>
<td>Incorporates eight classical models or theories, including TAM, TPB, theory of reasoned action (TRA), the motivational model (MM), a model combining the technology acceptance model (C-TAM-TPB), the model of PC utilization (MPCU), innovation diffusion theory(IDT), and social cognitive theory (SCT) (Venkatesh et al., 2003).</td>
</tr>
<tr>
<td>IS</td>
<td>Users' satisfaction with an information system depends on six variables: system quality, information quality, perceived usefulness, net benefits to individuals, net benefits to organizations, and net benefits to society (Seddon, 1997).</td>
</tr>
</tbody>
</table>

In our analysis of factors in the TAM, attitude was considered a direct and positive factor that determined an individual’s intention and behavior. The TAM assumed that two main factors, perceived usefulness and perceived ease of use, determined an individual’s attitude toward a new MOOC technology as well as the behavioral intention to use it (Joo et al., 2018; Shao, 2018; Tao et al., 2019; Wu & Chen, 2017). To some extent, perceived usefulness also had a direct impact on the learner’s behavior.

The TPB aimed to explain that an individual could decide whether or not to continue learning in a MOOC according to his or her own free will, as affected by three factors—attitude, subjective norms, and perceived behavioral control (Khan et al., 2018; Shao, 2018; Sun et al., 2019; Zhou, 2016). The latter two were hypothesized to directly influence one’s attitude towards online learning. Subjective norms referred to the individual’s perception of social pressures. Perceived behavioral control, defined as the individual’s perceived ease or difficulty, had a direct impact on learning behavior.

Motivation significantly affected learners’ psychological and behavioral engagement, which is important to reduce the dropout rate of MOOCs. The SDT, a well-established motivation theory that has been widely adopted to investigate participants’ persistence in MOOCs, indicated that behavior may be encouraged not only by autonomous motivations but also by controlled motivations. It was found that meeting students’ needs for autonomy, competence, and relatedness can increase their intrinsic motivation and lead to their
active engagement in MOOCs (Castano-Munoz et al., 2017; Hone & El Said, 2016; Khan et al., 2018; Sun et al., 2019). In addition to the SDT, some new factors were put forward that affect learners’ motivation and persistence in MOOCs, such as an individual’s preference or interest, goal orientation, self-efficacy, and innovativeness (Jung & Lee, 2018; Tsai et al., 2018; Zhang et al., 2016). External motivational factors were also investigated, including social recognition, social influence, and environmental stimulus (Luik et al., 2019; Wu & Chen, 2017; Zhao et al., 2020; Zhou, 2017).

The UTAUT was applied as a basic framework for designing questionnaire items, integrating eight classical models (Botero et al., 2018; Zhao et al., 2020). The UTAUT proposed several hypotheses regarding the impact of four factors on behavioral intentions: (a) performance expectancy, (b) effort expectancy, (c) social influence, and (d) facilitating conditions. It also considered that learners’ gender, age, experience, and voluntariness of use affected these four factors.

The TTF was used to evaluate how information technology leads to learners’ performance and utilization in MOOCs, and to judge the match between the learning task and the characteristics of MOOC technology (Khan et al., 2018; Wu & Chen, 2017).

In the analysis of factors in the IS success model, system quality, course quality, and service quality were significant antecedents of learners’ continuance intention to use MOOCs (Yang et al., 2017). Some new factors that influenced MOOC quality were also considered. Network externalities affected users’ persistence through the mediation of system quality (Li et al., 2018). MOOC course quality was mainly determined by the course design including course content and course structure (Hone & El Said, 2016). Instructor and co-learners effects, such as interaction, support, and feedback, influenced both course quality and service quality (Hone & El Said, 2016).

**F-G Graph for Learners’ Performance in MOOCs**

A challenge for this study was to build an F-G graph to summarize various factors about learners’ performance in MOOCs in terms of aspects of the research models that were examined. After reviewing the articles, we divided these factors into four categories: motivation, SRL, attitudinal learning, and learning strategies, clearly shown in different colors in Figure 2. The factors in the rounded rectangles had direct or indirect impacts on learners’ performance in MOOCs. The direction of an arrow points from an explanatory variable to a dependent variable.
Many studies about learners’ performance have integrated existing survey instruments to design questionnaire items, such as the motivated strategies for learning questionnaire (MSLQ; \( n = 12 \)), the online self-regulated learning questionnaire (OSLQ; \( n = 7 \)), the meta-cognitive awareness inventory (MAI; \( n = 5 \)), and the learning strategies questionnaire (LS; \( n = 5 \)). The key factors in these instruments that have been considered in MOOC environments are enclosed with dashed boxes in Figure 2. In addition, Table 3 summarizes how factors about MOOC learners’ motivation and learning strategies have been addressed in these four instruments. A checkmark indicates that the questionnaire considered the corresponding factor.
Table 3

Factors Related to MOOC Learners’ Motivation and Learning Strategies in MSLQ, OSLQ, MAI, and LS

<table>
<thead>
<tr>
<th>Scale</th>
<th>Factor</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MSLQ</td>
</tr>
<tr>
<td>Motivation</td>
<td>Intrinsic goal orientation</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Extrinsic goal orientation</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Task value</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Control beliefs</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Self-efficacy</td>
<td>✓</td>
</tr>
<tr>
<td>Meta-cognitive strategies</td>
<td>Goal setting</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Strategic planning</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Task strategies</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Elaboration</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Critical thinking</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Meta-cognitive self-regulation</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Self-evaluation</td>
<td>✓</td>
</tr>
<tr>
<td>Resource management strategies</td>
<td>Time management</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Environment structuring</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Help seeking</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Strategy regulation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Effort regulation</td>
<td>✓</td>
</tr>
</tbody>
</table>
strategies section included cognitive and meta-cognitive strategies and resource management strategies.

The OSLQ was adopted to measure learners’ SRL ability and strategies, including goal setting, environment structure, task strategies, time management, help-seeking, and self-evaluation (Kizilcec et al., 2017; Lee et al., 2020; Martinez-Lopez et al., 2017).

The MAI was constructed to measure meta-cognitive awareness as classified into two categories—cognition knowledge and cognition regulation (Schraw & Dennison, 1994).

The LS questionnaire has been used to measure three learning strategies—cognitive learning, behavioral learning, and self-regulatory learning—as associated with learning gain in MOOCs (Warr & Downing, 2000). The factors within these strategies included elaboration, help-seeking, motivation control, and comprehension monitoring (self-evaluation), among others.

In the analysis of new factors not included in the classical questionnaires, attitudinal learning was investigated in order to study the relationship between learners’ inherent positive attitudes and their belief in being able to complete learning tasks well (Watson et al., 2018; Watson et al., 2016). Learners’ emotional state and self-perceived achievement when attending a MOOC has been shown to affect their attitudinal learning. Behavioral learning was mainly predicted by learners’ engagement with activities (Ding & Zhao, 2020). Some new factors affecting learners’ motivation were also explored, such as individual benefits, including career, personal, and educational benefits. Social influence, similar to situational interest, was also studied and included certain conditions or stimuli in the social environment, such as peers’ recommendation and teacher’s support (de Barba et al., 2016; Durksen et al., 2016; Gallagher & Savage, 2016). MOOC instructors can refer to the influencing factors listed in Figure 2 to design for learner-centered experiences in the MOOC space (Blum-Smith et al., 2021).

**F-G Graph for Assessment of MOOCs**

In the reviewed articles, some researchers investigated students’ and teachers’ overall evaluation of MOOCs before or after participating in their courses. Figure 3 is an F-G graph that illustrates our summary of research models for assessment of MOOCs. This analysis spanned four dimensions, namely, (a) learners’ evaluation, (b) learners’ perceived benefits from learning, (c) teachers’ evaluation, and (d) teachers’ perceived benefits from teaching. The factors in the rounded rectangles directly or indirectly affect the assessment of MOOCs by learners and teachers. The direction of an arrow points from an explanatory variable to a dependent variable.
Regarding evaluation by learners and teachers, in order to obtain feedback that contributed to improving MOOCs, most researchers collected opinions and suggestions from students and teachers about course design, including course content, course structure, and available resources (Gan, 2018), as well as teaching skills and methods (Gan, 2018; Kormos & Nijakowska, 2017; Lowenthal et al., 2018). Regarding teaching methods, students’ main concerns were feedback from and interaction with instructors and co-learners (Marta-Lazo et al., 2019). In addition, students’ views on criteria for evaluating academic performance were crucial to assessment of MOOCs (Robinson, 2016; Ruiz-Palmero et al., 2019; Sari et al., 2020; Teresa Garcia-Alvarez et al., 2018). Teachers were concerned about their course management skills, teaching challenges, and personal development (Donitsa-Schmidt & Topaz, 2018; Robinson, 2016).

Students evaluated MOOCs based on what they perceived as benefits, including academic achievement, expected certificates or rewards, progress of learning efficiency, and effort invested in acquiring new knowledge or practical skills (Jung et al., 2019; Ruiz-Palmero et al., 2019; Teresa Garcia-Alvarez et al., 2018). Through participating in MOOCs, learners gained tangible and intangible benefits that generally justified their expectations, usually coinciding with individuals’ plans to change their career, education, or life trajectory (Sablina et al., 2018).

Teachers’ perceived benefits from providing courses as MOOCs were the key factors when they evaluated MOOCs. Benefits consisted mainly of enriching their instructional practice and experience, professional development, and potential for lifelong learning (Donitsa-Schmidt & Topaz, 2018). A teaching-quality control system was proposed as a way to provide teachers with motivation for continuous teaching with
MOOCs, and to promote teachers’ self-confidence and self-efficacy (Gan, 2018).

**Analysis Methods Used Most Often in Research on MOOCs**

After collecting questionnaire data, researchers chose analysis methods according to their different research needs. Based on our summary of the analysis methods used in the 126 reviewed studies, 61 articles (48.41%) used descriptive analysis, 53 studies (42.06%) used a structural equation model (SEM), 48 articles (38.10%) performed reliability analysis, and 41 studies (32.54%) adopted validity analysis. Most articles used several quantitative analysis methods at the same time. Researchers used various statistical analysis software to assist the processes of data analysis, most often IBM SPSS ($n = 34$) and AMOS ($n = 14$).

In the research we investigated, descriptive statistics often dealt with demographic data including participants’ gender, age, educational background, and experience with MOOCs (Botero et al., 2018; de Barba et al., 2016; Farhan et al., 2019; Hone & El Said, 2016). Descriptive statistics of data characteristics included data frequency analysis, centralized trend analysis, dispersion analysis, distribution, and some basic statistical graphics.

Reliability analysis refers to the degree of consistency of the results obtained when a questionnaire repeatedly measures the same object. It is best to verify the reliability of the items before using a questionnaire instrument to collect data. In the articles we reviewed, Cronbach’s $\alpha$ was the most commonly used reliability coefficient (Kovanovic et al., 2017; Sun et al., 2019; Tsai et al., 2018; Zhao et al., 2020). The data collected through questionnaires was generally considered credible when Cronbach’s $\alpha$ was greater than 0.7.

Validity analysis determines the degree to which the measurement results of a questionnaire can accurately reflect what needs to be measured. Validity analysis comprises content validity and structural validity. In the studies we examined, researchers usually invited people with extensive development experience to check the content validity of their questionnaires (Jo, 2018; Zhou, 2017). Structural validity consisted of two main methods, namely exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). EFA was commonly used in item analysis for scale preparation to explore the model structure, while CFA was used in reliability and validity analysis of mature questionnaires to verify the structure of a model (Jansen et al., 2017; Luik et al., 2019; Shahzad et al., 2020).

SEM is a more often recommended analysis method when attitude-related variables are included in the hypothesis model. SEM is a statistical method to analyze the relationship between variables based on a covariance matrix of variables for multivariate data analysis. The SEM methods used most frequently in the studies we examined were the partial least squares method (PLS-SEM) (Hone & El Said, 2016; Shao, 2018; Yang et al., 2017; Zhao et al., 2020) and the maximum likelihood estimation method (MLE-SEM) (de Barba et al., 2016; Teo & Dai, 2019; Zhou, 2016). When the collected data had no significant distribution characteristics, researchers most often used PLS-SEM analysis.
Conclusion

Through systematic review and analysis of 126 questionnaire-based quantitative research articles on MOOCs published between January 2015 and August 2020, this study explored the research paradigms associated with this field including research topics, models, and data analysis methods.

Our findings show that MOOC research remains an important and growing field of interest for educational researchers. Empirical studies of MOOCs explored multiple issues, most of which were based on quantitative investigation and research. This paper divided the key topics of these reviewed articles into three different categories: (a) the determinants of learners’ dropout rate or continuance intention, (b) the relevant factors of learners’ performance, and (c) participants’ assessment of MOOCs. Most research focused on MOOC participants or learners, with a few researchers actively concentrating on MOOC instructors, curriculum design, and platform development. It may be promising for researchers to conduct more in-depth exploration of the characteristics and profiles of MOOC participants and instructors, the potential for personalized customization in MOOCs, and MOOC quality improvements.

As shown in this study, most questionnaire-based quantitative studies of MOOCs had a solid theoretical foundation, a standardized research process, and effective research methods. By understanding the research paradigms summarized and expanded in this study, researchers will be better able to carry out more empirical research while experimenting with research methods that have not yet been commonly used. This paper provides three F-G graphs to separately analyze the correspondence between research topics and factors involved in the models or hypotheses studies were based on. By referring to the F-G graphs, MOOC researchers can design more reasonable questionnaire items and collect high-quality data to better support data science research.

This study revealed several limitations of MOOC research as apparent in the studies we reviewed, including small sample size during data collection, lack of diversity among the survey participants, and the limitations inherent in traditional statistical analysis. Based on these limitations, we suggest three new directions for the future development of research on MOOCs.

First, we recommend expanding the scope of data collection and establishing big data sets. In some studies of MOOCs selected for this paper, the sample size for surveys was relatively small. Some research results failed to be persuasive, or the factors investigated had no significant impact on the research subjects. A preferable approach may be to expand the scope and target of data collection, and establish a large-scale database in the MOOC field, perhaps even worldwide. This would serve to make the data sources more objective, more universal, and more convincing (Ang et al., 2020).

Second, we suggest standardizing multi-sourced heterogeneous data about MOOCs. This is an essential feature of big data, since the survey data from different studies are based on different collection scales and standards. Standardized multi-sourced heterogeneity data can provide a solid data foundation and further insights for subsequent data analysis.

Finally, we recommend applying data mining and deep learning methods. In the articles we reviewed, data analysis methods were mostly limited to traditional statistical approaches. Data mining and deep learning emphasize correlation judgments between samples and infer the population from the standard data set
What is more, researchers can apply data mining and deep learning to analyze objective behaviors and subjective perceptions of MOOC learners and instructors, make feature profiles of users, and propose personalized optimization schemes (Geng et al., 2020; Cagiltay et al., 2020).
References


Robinson, R. (2016). Delivering a medical school elective with massive open online course (MOOC) technology. *Peerj, 4*(1), e2343. [https://doi.org/10.7717/peerj.2343](https://doi.org/10.7717/peerj.2343)


Tao, D., Fu, P., Wang, Y., Zhang, T., & Qu, X. (2019). Key characteristics in designing massive open online courses (MOOCs) for user acceptance: An application of the extended technology acceptance model. *Interactive Learning Environments, 44*, 1–14. https://doi.org/10.1080/10494820.2019.1695214


https://doi.org/10.1016/j.iheduc.2018.01.002
Appendix A

Details of 32 Typical Articles for the Top Three Topics

Table A1

12 Typical Articles About Dropout Rate or Continuance Intention to Use MOOCs

<table>
<thead>
<tr>
<th>Article</th>
<th>Questionnaire model and items design</th>
<th>Quantitative analysis</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yang &amp; Su, 2017</td>
<td>(1) TPB: Perceived behavior control, attitudes, subjective norms, behavior intention, actual behavior (2) TAM: Perceived ease of use (PEU), perceived usefulness (PU)</td>
<td>PLS-SEM Reliability analysis Validity analysis</td>
<td>272</td>
</tr>
<tr>
<td>Wu &amp; Chen, 2017</td>
<td>(1) TAM: PU, PEU, attitude toward using MOOCs, continuance intention to use (CIU) (2) TTF: Individual technology fit, task-technology fit (3) Social motivations: Social recognition, social influence (4) Features of MOOCs: Openness, reputation</td>
<td>PLS-SEM</td>
<td>252</td>
</tr>
<tr>
<td>Khan et al., 2018</td>
<td>(1) TTF: Task characteristics, technology characteristics (2) Social motivation: Social recognition, social influence (3) SDT: Perceived relatedness, autonomy, competence</td>
<td>Multivariate assumptions Kolmogorov-Smirnov test</td>
<td>414</td>
</tr>
<tr>
<td>Zhu et al., 2018</td>
<td>(1) TPB: Attitudes, perceived behavioral control, subjective norms, behavioral intention (2) SDT: Controlled motivation, autonomous motivation</td>
<td>CFA MLE-SEM</td>
<td>475</td>
</tr>
<tr>
<td>Yang et al., 2017</td>
<td>(1) IS: System quality, course quality, service quality (2) TAM: PU, PEU, CIU</td>
<td>PLS-SEM</td>
<td>294</td>
</tr>
<tr>
<td>Botero et al., 2018</td>
<td>UTAUT: Performance expectancy, effort expectancy, social influence, facilitating conditions, behavioral intention, attitudes towards behavior</td>
<td>SEM Descriptive statistics</td>
<td>587</td>
</tr>
<tr>
<td>Zhang et al., 2016</td>
<td>(1) TAM: PU, PEU, CIU (2) Perceived learner control, personal innovativeness, information technology, E-learning self-efficacy</td>
<td>PLS-SEM Validity analysis: AVE Cronbach’s alpha</td>
<td>214</td>
</tr>
</tbody>
</table>
## Table A2

### 10 Typical Articles About Learners’ Performance in MOOCs

<table>
<thead>
<tr>
<th>Factor</th>
<th>Article</th>
<th>Questionnaire model and items design</th>
<th>Methodology</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRL</td>
<td>Lee et al., 2020</td>
<td>OSLQ: Self-regulated learning strategies MSLQ: Self-efficacy, task value</td>
<td>Multiple regression Pearson’s correlation analysis</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td>Martinez-Lopez et al., 2017</td>
<td>OSLQ, MAI, and LS Goal setting, environment structuring, task strategies, management help, help-seeking, self-evaluation</td>
<td>Modified kappa Coefficient content validity Indexing (CVI) SEM</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>Jansen et al., 2017</td>
<td>MSLQ, OSLQ, MAI, and LS (1) Preparatory phase: Task definition, goal setting, strategic planning (2) Performance phase: Environmental structuring, time management, help-seeking, comprehension monitoring, task strategies, motivation control, effort regulation (3) Appraisal phase: Strategy regulation</td>
<td>EFA CFA Descriptive statistical analyses</td>
<td>162</td>
</tr>
<tr>
<td>Authors</td>
<td>Measures</td>
<td>Methods</td>
<td>Pages</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------------------------------------------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>Kizilcec et al., 2017</td>
<td>MSLQ, OSLQ, MAI, and LS Goal setting strategies, strategic planning, elaboration, help-seeking</td>
<td>Descriptive statistics, Spearman correlation coefficients, Fitted logistic regression</td>
<td>4831</td>
<td></td>
</tr>
<tr>
<td>de Barba et al., 2016</td>
<td>Motivation: (1) Individual interest, mastery-approach goals, value beliefs (2) Situational interest: Entering situational interest, maintaining situational interest</td>
<td>Cronbach’s alpha, Descriptive statistics, MLE-SEM, Bootstrap, Chi-square</td>
<td>862</td>
<td></td>
</tr>
<tr>
<td>Fryer &amp; Bovee, 2018</td>
<td>(1) Prior competence, prior computer use, teacher support, smartphone use (2) Ability beliefs, effort beliefs, task value</td>
<td>CFI (confirmatory fit index), RMSEA (root mean square error of approximation), MANOVA, Latent profile analysis (LPA)</td>
<td>642</td>
<td></td>
</tr>
<tr>
<td>Watted &amp; Barak, 2018</td>
<td>(1) Career benefits: Certificate (2) Personal benefits: Improving knowledge (3) Educational benefits: Research and professional advancement</td>
<td>Kruskal-Wallis analysis</td>
<td>377</td>
<td></td>
</tr>
<tr>
<td>Attitudinal learning</td>
<td>(1) Emotion, self-perceived achievement (2) Video engagement, assignment engagement</td>
<td>Reliability analysis, Cronbach’s alpha</td>
<td>378</td>
<td></td>
</tr>
<tr>
<td>Watson et al., 2018</td>
<td>Cognitive learning, affective learning, behavioral learning, social learning.</td>
<td>Descriptive statistics, CFA</td>
<td>1009</td>
<td></td>
</tr>
</tbody>
</table>
### Table A3

**10 Typical Articles About Assessment of MOOCs**

<table>
<thead>
<tr>
<th>Respondent</th>
<th>Article</th>
<th>Questionnaire</th>
<th>Methodology</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learners</td>
<td>Martin Nunez et al., 2017</td>
<td>Available resources, course forums, evaluations adequacy</td>
<td>Statistical analysis, MANOVA</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>Jung et al., 2019</td>
<td>(1) Course content, course structure, assessment method, learner-content interaction (2) Learner control, sense of progress, perceived effectiveness</td>
<td>Hierarchical linear regression</td>
<td>1364</td>
</tr>
<tr>
<td></td>
<td>Robinson, 2016</td>
<td>Teaching effectiveness, course objectives, overall rating, personal learning objectives, recommendation to others</td>
<td>t-tests, Cronbach’s α, Post-hoc power analysis</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Donitsa-Schmidt &amp; Topaz, 2018</td>
<td>(1) Flexibility and convenience (2) Learning opportunity, professional development</td>
<td>Descriptive analysis</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Meinert et al., 2018</td>
<td>Motivation, learning methods, course content</td>
<td>Descriptive analysis, Logistic regression</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Lowenthal et al., 2018</td>
<td>(1) Teaching motivation: Interest and passion, publicity and marketing, benefits and incentives (2) Teaching experience (3) Perception of MOOC educational value</td>
<td>Descriptive analysis</td>
<td>186</td>
</tr>
<tr>
<td></td>
<td>Kormos &amp; Nijakowska, 2017</td>
<td>Participants’ attitudes, self-confidence, concerns about teaching practices, self-efficacy beliefs</td>
<td>Principal component analysis, Regression factor scores, MANOVA, GLM</td>
<td>752</td>
</tr>
<tr>
<td></td>
<td>Sneddon et al., 2018</td>
<td>Course development, course evaluation, course delivery</td>
<td>Descriptive analysis</td>
<td>219</td>
</tr>
</tbody>
</table>