A Thematic Analysis of Research Trends in MOOCs
Dissertations and Theses (2008-2021)

Yibo Fan
Wayne State University
42 W Warren Ave, Detroit, MI 48202
yibofan@wayne.edu

Liangyue Lu
Grambling State University
GSU Box 4305, Grambling, LA 71245
lul@gram.edu

DOI: 10.6084/m9.figshare.21385578

Abstract
This study examined the topical and methodological trends in 78 dissertations and master’s theses about MOOCs published between 2008 and 2021 by ProQuest Dissertation and Theses Database to address the scarcity in existing literature reviews on young scholars’ research efforts. Six major topical trends were identified from the perspectives of different stakeholders of MOOCs. Methodological trends described the research methods employed and how they related to research topics, data collection and data analysis methods.

Keywords: massive open online courses (MOOCs), dissertations and theses, thematic analysis, literature review, online education

Few of the previous reviews on MOOCs (e.g., Deng et al., 2019), if not any, targeted the MOOCs dissertation and thesis studies, an important body of literature reflecting the research interests of young scholars and their advisors (Davies et al., 2010; Drysdale et al., 2013). In this review, thesis studies were those written for the master’s degree and dissertations for the doctoral degree. Specifically, it addressed two research gaps existing in the current reviews: (1) it offered an updated and a more comprehensive understanding of the research trends as reflected in MOOCs graduate research by covering the wide year range from 2008 to 2021; (2) it provided an important perspective which, if taken together with previous reviews, could help researchers develop a holistic understanding of the current state of the field.

To address the scarcity in existing literature reviews on young scholars’ research efforts, this study aims to highlight the topical and methodological issues and make corresponding recommendations for future research endeavor. Four research questions that guided this inquiry were as follows:

1. What research topics have been explored in the MOOCs graduate research (2008-2021)?
2. What research methods have been employed by the MOOCs graduate research (2008-2021)?
3. What data collection methods have been used in the MOOCs graduate research (2008-2021)?
4. What data analysis methods have been used in the MOOCs graduate research (2008-2021)?

Methodology

Data Collection
The search for relevant literature was performed in the “ProQuest Dissertation and Theses Database (PQDT)”. A list of search terms and procedures were used during the search process:

- The keywords used for the search were “MOOC*” and “massive open online course*”.
- The search was limited to “abstract” only.
- The year range covered was from 2008 to 2021.
- The language was limited to English only.

A two-stage data screening process was followed to screen the 222 initially returned records. At the first stage, the title and abstract of each study were carefully read to determine that MOOCs were studied rather than simply mentioned, yielding 98 studies that were potentially eligible for the second-stage screening, excluding two studies not available in full-text.

At the second stage, the researcher read the full text of each manuscript, guided by two major purposes: (1) to determine whether the studies studied issues on MOOCs; (2) to determine whether the description of method section provides sufficient detail so that adequate information could be extracted to address the research questions.

After completing the two-stage screening process, 78 studies were identified as eligible for review in this paper.

Data Analysis
This study modified the five research strands of MOOC research identified in Veletsianos and Shepherdson (2016), and used the following six coding categories to address RQ1: (1) learner-focused, (2) instructor-focused, (3) institution-focused, (4) design-focused, (5) employer-focused, (6) provider-focused.

For RQ2, three widely accepted categories: quantitative, qualitative, and mixed-methods research (Creswell & Plano Clark, 2018) were adopted, and they have been used in several previous reviews (Almasi & Zhu, 2020; Zhu et al., 2018). Apart from the three general categories, the researcher also noticed the use of design-based research (DBR) (Brown, 1992) during the second-stage screening process. Given the increasing recognition and adoption of DBR by the research community (Anderson & Shattuck, 2012; Raffaghello et al., 2015), and the need to reflect the methodological diversity in the reviewed studies, this study used DBR as the fourth category to describe the research methods adopted in the reviewed studies.

Regarding RQ3, this study employed some of the categories identified in Veletsianos and Shepherdson (2016) such as surveys, and added new categories like platform data, document review, artifacts, and physiological signals.

To address RQ4, the researcher categorized the methods into more general items, for instance, methods like regression and analysis of variance were coded under “inferential statistics”. The final identified data analysis methods were cross checked so that they were mutually exclusive.
Findings and Discussion

RQ1: What research topics have been explored in the MOOCs graduate research (2008-2021)?

Six major research topics were identified based on the examination of research purpose and research questions in each study. To reflect the diversity of topics, each major research topic was further categorized into certain number of sub-topics, which will be presented in more detail in the following sections.

More than half of the reviewed studies explored issues related to learners in MOOCs (65.4%), which was in line with previous review studies (e.g., Deng & Benckendorff, 2017; Raffaghelli et al., 2015). Other major research topics that followed the learner-related issues were instructor-focused (11.5%), institution-focused (11.5%), design-focused (6.4%), employer-focused (3.8%), and provider-focused topics (1.3%).

**Learner-focused**

Completion. There has been at least one MOOCs dissertation or thesis each year examining completion-related issues from 2014 to 2020. Much research effort investigated the factors correlated with MOOCs completion, including self-directed learning (Schulze, 2014), cultural indicators (Alabdullaziz, 2015), types of assessment (Papathoma, 2015), and the incorporation of a variety of multimedia materials and guided discussions (Montgomery, 2016), etc.

Learning experience. Studies in this category targeted learning experience issues from various perspectives, which included: the investigation of learners’ general experience in MOOCs (Morris, 2014), the examination of learners’ participation patterns (Stager, 2016), and the study of the barriers and challenges learners encountered in MOOCs (Cox, 2018), etc.

Motivation. Six out of ten studies in this category explored learners’ motivation to enroll in MOOCs (e.g., Alabdullaziz, 2015). Other research topics included the investigation of the relationship between learners’ motivation and learning outcomes (Wang, 2017) and participation in MOOCs (Haniya, 2019), and the drivers that motivated learners to complete MOOCs (Cox, 2018), etc.

Perceptions. Five of the seven studies in this category examined learners’ perceptions of their experience in MOOCs (e.g., Kilgore, 2018), including the perceptions of peer interactions (Loizzo, 2015), and engagement and achievement in MOOCs (Morris, 2014).

Engagement. Studies in this category investigated the factors impacting learner engagement in MOOCs, including learners’ trait complexes like personality and achievement goal orientations (Torres, 2016), the MOOC design features (Gore, 2018), and the integration of MOOCs into campus courses in Saudi women’s higher education (Almutairi, 2018), etc.

Other trends. Studies coded under “learning outcomes” investigated the factors contributing to learner performance in MOOCs, including learners’ motivation (Wang, 2017), pre-test (Janelli, 2019), self-regulatory strategies (Maldonado, 2019), and feedback (e.g., Kulkarni, 2015). “Peer interaction” studies mainly focused on learners’ interaction pattern in MOOCs (Kellogg, 2014), the promotion of peer interaction (Hill, 2015), and how peer interaction related to learner performance in MOOCs (Huesman, 2019).

**Instructor-focused**

Compared with “learner-focused” topics, the number of studies that examined instructor-related issues was extremely low ($N = 9$). The results suggested that the majority of “instructor-
focused” studies explored issues related to instructors’ perceptions of MOOCs (33.3%), motivation to teach MOOCs (22.2%) and experience in teaching MOOCs (22.2%).

**Institution-focused**
Nine studies specifically focusing on issues related to higher education institutions covering three sub-topics: the impact of MOOCs on the landscape of higher education (66.7%), the integration of MOOCs into the existing instruction paradigm (44.4%), and the institution disposition (22.2%, e.g., decision-making process and expectations).

**Design-focused**
Five studies specifically addressed matters on the design of MOOCs. The identified sub-topics under this category included: the use of instructional design models and theories in MOOCs design, the design of MOOCs to improve accessibility for disabled learners, and the collaborative design process and experience of MOOCs.

**Employer-focused**
Three studies were coded for this major research topic. Only one sub-topic attending to the employers’ perceptions of, and acceptance of MOOCs was identified. For instance, Outland (2014) addressed hiring managers’ perceptions of courses offered in MOOCs, and the potential positive or negative impact of MOOC course-taking on candidates’ employability.

**Provider-focused**
One last identified major research topic examined MOOC provider perceptions of assessment, accountability, and accreditation of MOOCs (May, 2015).

The findings suggested that (1) the assessment in MOOCs should be more learner-centered, (2) learners in MOOCs are the major stakeholders and (3) MOOCs could be institutionally accredited when offered as a part of degree program. Notably, “provider” is to be distinguished from “institution” reviewed above. Provider in this study refers specifically to the MOOC platforms hosting the MOOC courses produced by different higher education “institutions”.

**RQ2: What research methods have been employed by the MOOCs graduate research (2008-2021)?**

Among the 78 eligible studies, nearly half of them were quantitative in nature, 29 studies were coded as qualitative, and 15 studies employed the mixed-methods approach. It was worth noting that one study specifically declared the use of DBR as its guiding approach (e.g., Li, 2015), in which both quantitative and qualitative data were collected. Qualitative methods were not the least used as indicated by previous reviews (e.g., Zhu et al., 2018, 2020), instead, they were only second to the quantitative methods, which were suggested in previous reviews as the most frequently used methods (e.g., Raffaghelli et al., 2015; Veletsianos & Shepherdson, 2016; Zhu et al., 2018, 2020).

Mixed-methods were the third most employed methods, which should be concerning as the small number of such studies may reflect the superficial magnitude of complexity of the research design in MOOC studies (Gašević et al., 2014; Greene et al., 1989). Future studies should balance the use of different research methods in order to achieve the methodological
diversity advocated by Veletsianos and Shepherdson (2016) and to increase the validity of research findings (Greene et al., 1989).

DBR was the least adopted method in the reviewed studies. As an emerging research methodology that has the potential to guide better educational research, DBR has received considerable attention from researchers (Amiel & Reeves, 2008; Design-Based Research Collective, 2003). MOOCs offer a purely naturalistic setting in which learners are diverse in demographics and locations (Breslow et al., 2013). The successful implementation of DBR requires a naturalistic setting (Design-Based Research Collective, 2003; McKenney & Reeves, 2012), starts with the discussion of practical problems between practitioners and researchers (Amiel & Reeves, 2008), and usually involves multiple rounds of iterations (Amiel & Reeves, 2008; Anderson & Shattuck, 2012). The scarcity of DBR studies in MOOCs could possibly be attributed to the high demand of time and effort for the multiple iterations (Gašević et al., 2014). Moreover, given the large investment in developing and offering MOOCs, some MOOCs are not offered multiple times, which makes it more challenging for researchers to conducted DBR that requires several iterations in MOOCs (Gašević et al., 2014). However, it would be possible to implement several interventions in different subpopulations (Kizilcec et al., 2013) among the registered learners in MOOCs to make up for inadequate opportunities of multiple iterations (Gašević et al., 2014).

RQ3: What data collection methods have been used in the MOOCs graduate research (2008-2021)?

Survey was the most used data collection method which was employed in more than half of the reviewed studies. Survey was used in studies for collecting learners’ demographic information (Martin, 2015); measuring learners’ motivation (Li, 2015), self-directed learning readiness (Schulze, 2014), and self-efficacy (Branson, 2017); collecting stakeholders’ perceptions of MOOCs (e.g., Stein, 2016), etc. Other identified data collection methods included semi-structured interviews (33 studies, 42%); platform data (24 studies, 30.8%), document review (18 studies, 23.1%), and observation (eight studies, 10.3%).

The researcher also examined the number of data collection methods employed by each study. About half of the reviewed studies (48.7%) used only one data collection method, including survey (42.1%), platform data (31.6%), and semi-structured interview (18.4%). About one third of the studies (29.5%) employed two data collection methods, in which survey was utilized in 16 of the 23 studies (69.6%), platform data in 11 studies (47.8%), semi-structured interview in 10 studies (43.5%), document review in 4 studies (17.4%), observation in 2 studies (8.7%), and focus groups in 2 studies (8.7%). Survey was the most frequently employed method in either one-method or two-method studies. The pattern was a bit different in studies employing three methods, in which semi-structured interview was used in 14 studies, followed by document review in 12 studies, survey in 10 studies, and observation in 4 studies. Only two of the 78 studies used four data collection methods.

Since all data collection methods are inherently biased and limited in scope, the use of only one method will inevitably bias the research findings, thus diminishing the validity of the research studies (Greene et al., 1989). It is, therefore, suggested in this study that two or more methods that have “offsetting biases” (Greene et al., 1989, p. 256) be used in the future research endeavors to enrich the methodological diversity.
RQ4: What data analysis methods have been used in the MOOCs graduate research (2008-2021)?

A total of 12 major data analysis methods were identified. Inferential statistics were the most frequently adopted data analysis methods (48.7%), followed by descriptive statistics (39.7%), coding (25.6%), thematic analysis (21.8%), and content analysis (9.0%). The findings suggested that the inferential instead of descriptive statistics were the most frequently adopted data analysis methods, which was in stark contrast with previous review effort (e.g., Zhu et al., 2018, 2020) and may reflect the complexity of research designs of the reviewed studies. Only seven studies adopted the more advanced and versatile methods like learning analytics and educational data mining (LA/EDM) (Gašević et al., 2014). LA/EDM methods could potentially enable the researchers to make full use of learners’ trace data by analyzing their learning activities in MOOCs (Gašević et al., 2014), thus providing a more intricate profile of how learning takes place in MOOCs. Future studies may consider an increasing use of LA/EDM in alignment with sound educational theories to provide a more comprehensive and meaningful interpretation of learning and teaching in MOOCs (Gašević et al., 2014).

Regarding the number of data analysis methods adopted in each study, the results suggested that half of the 78 studies used one data analysis method, and about one third of studies employed two data analysis methods. Many one-method studies employed qualitative type of data analysis methods like coding and thematic analysis. Statistical analysis methods including descriptive and inferential statistics were more often adopted than qualitative data analysis methods in studies employing more than two data analysis methods. In effect, 37 of 39 (94.9%) studies using two or more data analysis methods adopted the statistical analysis methods. Veletsianos and Shepherdson (2016) warned that the heavy dependence on particular research methods could largely constrain our understanding of MOOCs, therefore, they called for “an expansion of the methodological approaches used in MOOC research” (p. 214). More importantly, the triangulation of multiple data sources would enhance the validity of research findings (Greene et al., 1989).

Limitations

Since this study reviewed MOOCs graduate research available in ProQuest Dissertation and Theses Database (PQDT), in which a large percentage of studies are from universities in the U.S. and Canada, this study may be “geographically limited” to some extent. Future studies are encouraged to expand this review by including MOOCs graduate research written in other languages such as Chinese as recent review effort demonstrated that China came only after the U.S. in generating empirical MOOC research (Zhu et al., 2020). Also, future studies may consult other databases dedicated to the collection of dissertations and theses within specific institutions.

Conclusion

This supplementary review was consistent with previous systematic review effort in that it also suggested that most graduate research studies targeted learner-focused issues (e.g., Raffaghelli et al., 2015; Zhu et al., 2018, 2020) adopting quantitative research methods (e.g., Raffaghelli et al., 2015). As opposed to previous reviews (e.g., Zhu et al., 2018, 2020), qualitative methods were the second most widely used methods rather than least used. When it comes to data analysis methods, this review contrasted with existing reviews (e.g., Zhu et al., 2018, 2020) in that inferential instead of descriptive statistics were most frequently adopted. Moreover, the findings also revealed that half of the studies employed only one data analysis
method, and this would limit our comprehensive understanding of MOOCs (Veletsianos & Shepherdson, 2016).

What is unique about this review related to its identification of the scarcity of studies employing the design-based research method (Amiel & Reeves, 2008; Design-Based Research Collective, 2003), which has been increasingly accepted as an alternative research approach in addition to quantitative, qualitative, and mixed-methods approaches (Anderson & Shattuck, 2012; Raffaghelli et al., 2015). Given its deep focus in solving practical problems by working with practitioners as well fine-tuning theories (Amiel & Reeves, 2008; Anderson & Shattuck, 2012), DBR has great potential in solving some serious issues (i.e., how to engage learners in MOOCs) facing MOOCs that other research approaches fall short of. Therefore, future studies are encouraged to employ DBR approach to explore critical issues such as designing engaging courses, thus lowering the attrition rate in MOOCs. The paucity of graduate research studies adopting DBR approach should be alarming as it may signal that the young scholars are not prepared to conduct DBR studies. This study is timely and important as it points out this issue that has not been identified in other review efforts. Future studies may continue expanding the repertoire of methodological approaches in MOOCs research (i.e., more use of methods like educational data mining) and employing DBR approach, as well as focusing on topics that are critical for the sustainability and continuous growth of MOOCs.
References

Alabdullaziz, F. (2015). *Cultural diversity in massive open online courses: The correlation between cultural indicators and students’ attrition* [University of Northern Colorado]. https://www.proquest.com/pqdt/docview/1734044714/1AC88122456E4648PQ/1?accountid=14925


Almutairi, F. M. (2018). *The impact of integrating MOOCs into campus courses on student engagement* [University of Southampton]. https://eprints.soton.ac.uk/429737/1/final_version.pdf


https://doi.org/10.18178/ijiet.2017.7.8.939


Kellogg, S. (2014). *Patterns of peer interaction and mechanisms governing social network structure in three massively open online courses for educators* [North Carolina State University]. https://www.proquest.com/pqdt/docview/1642784941/B47BF96A0A54966PQ/1?accountid=14925

133
Kilgore, P. A. (2018). *Adult college students’ perceptions about learning mathematics via developmental mathematical xMOOCs* [University of South Florida].
https://www.proquest.com/pqdt/docview/2049683054/6A428BA8B6D94AE0PQ/1?accountid=14925

https://dl.acm.org/doi/pdf/10.1145/2460296.2460330

https://www.proquest.com/pqdt/docview/2459635988/185D29BEF9304EB2PQ/1?accountid=14925

https://www.proquest.com/pqdt/docview/1769046180/D3A63301EA1246A8PQ/1?accountid=14925

Loizzo, J. L. (2015). *Adult learners’ perceptions of MOOC motivation, success, and completion: A virtual ethnographic study* [Purdue University].
https://www.proquest.com/pqdt/docview/1712965103/F45FD9F794C54E0APQ/1?accountid=14925

https://www.proquest.com/pqdt/docview/2455950942/CBD5487C139C483FPQ/1?accountid=14925

https://www.proquest.com/pqdt/docview/1700208746/520771EB4C14F89PQ/1?accountid=14925

https://www.proquest.com/pqdt/docview/2056419734/C1A80420A8BB4D57PQ/1?accountid=14925


https://www.proquest.com/pqdt/docview/1851276630/EE434AF53FE04041PQ/1?accountid=14925

Morris, I. L. (2014). *An exploratory analysis of motivation and engagement in massive online open courses (MOOCs)* [University of California, Riverside].
https://www.proquest.com/pqdt/docview/1656449552/9F8357C4248F4456PQ/1?accountid=14925

Outland, J. C. (2014). *Examining the market positioning of massive open online courses to maximize employer acceptance* [Northcentral University].
https://www.proquest.com/pqdt/docview/1615853828/5B584637717429FPQ/1?accountid=14925

Papathoma, T. (2015). *Investigating different types of assessment in massive open online courses*

Schulze, A. S. (2014). Massive open online courses (MOOCs) and completion rates: Are self-directed adult learners the most successful at MOOCs? [Pepperdine University]. https://www.proquest.com/pqdt/docview/1549976283/7C73676149824FA6PQ/1?accountid=14925


Torres, W. J. (2016). *Course participation, performance, and completion by adult learners in a massive open online course (MOOC): Trait complexes, interest, and non-ability determinants* [Rice University]. https://www.proquest.com/pqdt/docview/2001550921/A23723769C47405FPQ/1?accountid=14925


Wang, Y. (2017). Demystifying learner success: Before, during, and after a massive open online course [Columbia University]. https://www.proquest.com/pqdt/docview/1887965559/33FA24554BAB4C46PQ/1?accountid=14925
