Meina Zhu, Annisa Sari, and Curtis J. Bonk
Indiana University
meinzhu@iu.edu, annsari@uny.ac.id, cjbonk@indiana.edu

Abstract
This research examines research paradigms and massive open online courses (MOOCs) topics to increase our understanding of the MOOC phenomenon and methodological issues. We reviewed 197 studies published from October 2014 to July 2017 in two phases. Results revealed that: (a) quantitative methods were mainly used, (b) the primary data collection methods were surveys, platform database analysis, and interviews, (c) the majority of MOOC researchers used descriptive statistics to analyze data, (d) the research focus was mainly on learners, followed by design-focused, context and impact-focused, and instructor-focused studies, (e) most authors worked collaboratively within the same country, and (f) MOOCs from the United States were researched the most followed by the UK, Spain, and China. Implications for future research are offered.

Objectives/Purpose
Due to its popularity and potential impact to higher education, empirical research on MOOCs has quickly emerged during the past few years (Deng & Benckendorff, 2017; Veletsianos & Shepherdson, 2016). To better grasp the trends of MOOC studies and the impact of MOOCs, detailed and systematic analysis and synthesis of existing MOOC empirical studies in particular areas is required (Reich, 2015; Veletsianos & Shepherdson, 2016). Future systematic analysis of MOOC empirical studies might compare the research summaries and findings from different time periods (Veletsianos & Shepherdson, 2016). This study expands beyond previous MOOC reviews through an extensive examination of the topic areas, methods, geographic distribution, and country of MOOC delivery.

To this end, the researchers initially reviewed 146 empirical MOOC studies from October 2014 to November 2016 (Phase One). Next, they extended the study by reviewing 51 additional empirical MOOC studies from December 2016 to July 2017 (Phase Two). While Phase One of the study was published in The Internet and Higher Education journal (Zhu, Sari, & Lee, 2018), Phase Two is the extension of the previous study which is presented in this paper. This research aims to provide a review of research paradigms and topics of MOOCs to better understand the MOOC phenomenon and methodological issues related to it. The four research questions below guided our inquiry:
1. What are the research methods that researchers employed in empirical MOOC studies?
2. What are the research topics or focuses of previous MOOC studies?
3. How are researchers of empirical MOOC studies geographically distributed?
4. In terms of the delivery of MOOCs, what are the countries which are attracting the most research attention?

Theoretical Perspectives
Anyone with an Internet connection can enroll in a MOOC, obtain access to the resources, interact with peers, and reflect and share their knowledge with others (Kop, 2011; Koutropoulos et al., 2012). The emergence of MOOCs also enables higher education to be more accessible to learners worldwide. Not surprisingly, the MOOC phenomenon has received significant attention in higher education (Bali, 2014; Bullfin, Pangrazio, & Selwyn, 2014; Carver & Harrison, 2013; Jacobs, 2013; Liyanagunawardena et al., 2013) as well as the prevailing media. In 2016 alone, nearly 7,000 MOOCs were offered (Shah, 2016). Such upward trends continued the following year with over 9,400 MOOCs offered by more than 800 universities in 2017 (Lederman, 2018, Shah, 2018). Given these trends, it is not too surprising that there is enormous research attention and resources being placed in this new field.

During the past few years, researchers have conducted extensive reviews of the MOOC research literature as it has evolved, including Liyanagunawardena et al. (2013), Ebben and Murphy (2014), Hew and Cheung (2014), Jacoby (2014), Kennedy (2014), Raffaghello, Cucchiara, and Persico (2015), Veletsianos and Shepherdson (2015), Veletsianos and Shepherdson (2016), Deng and Benckendorff (2017), and (Zhu et al., 2018). These studies revealed that MOOC research topics and trends are continually and rapidly changing (Veletsianos & Shepherdson, 2016). For
example, early MOOC studies were mainly conceptual studies (Kennedy, 2014), whereas more recent MOOC research is increasingly empirically-based (Veletsianos & Shepherdson, 2016).

As a fast-emerging field, researchers and educators could benefit from a more comprehensive picture of the MOOC research methods and focus areas or topics. There is an increasing need to recap the empirical research literature on MOOCs because of the fact that: (a) tens of millions of people enroll in thousands of MOOCs each year (Shah, 2016); (b) MOOCs have inspired a discussion of a wide array of pressing topics and needs in education (e.g., sustainability, pedagogy, completion rates, self-directed learning, accreditation, quality, etc.) (Jona & Naidu, 2014); (c) researchers need to consider the appropriate methods for data collection and analysis (Raffagellie et al., 2015); (d) MOOC practices are too often leading MOOC theory, instead of the other way around (Jona & Naidu); (e) MOOCs are being utilized in a wide variety of ways and no longer should be seen as a homogenous practice (Jona & Naidu, 2014); and (f) there is wide and growing diversity of subject areas of MOOCs researchers (Liyanagunawardena et al., 2013; Veletsianos & Shepherdson, 2015).

Methods

Data Collection

The research team collected the data by searching electronic databases, Scopus, and peer-reviewed journal articles. This study includes Phase One (October 2014 - November 2016) and Phase Two (December 2016 - July 2017).

To ensure the relevance of the review, the researchers utilized six selection criteria. First, the study had to be an empirical study. Second, the study must mainly investigate educational aspects of MOOCs. Third, the Scopus database was the primary source of the article selection process due to its extensive large coverage (Hardcastle, 2011). In addition, several journals which were not indexed by Scopus were included since they were targeted by previous research studies (e.g., Veletsianos & Shepherdson, 2016) or were known to publish MOOC-related research. Fourth, the keywords “MOOC” and “Massive Online Open Course(s)” were used to screen titles, abstracts, and keywords of the literature. Fifth, the types of documents included in this review were journal articles; in effect, technical reports, white papers, conference papers, and news releases were not included. Finally, studies had to have been published in English.

For increasing validity, two researchers cross-checked the data, discussed any discrepancies, and reached consensus on analyses. The final number of papers included was 146 for Phase One, and 51 for Phase Two. The overall inter-rater agreement across all items was 91%. The recorded data of each study included a variety of dimensions such as authors and their affiliations, location of the authors, location of the MOOC delivery, year of publication, title, journal name, general analytic method (e.g., qualitative, quantitative, or mixed methods), data collection methods, data analysis methods, general study focus, special study focus, and article URL.

Data Analysis

To answer research question (RQ) 1, the researchers coded the papers using three general categories: quantitative research, qualitative research, and mixed-methods (Creswell, 2003). To analyze data collection methods, researchers utilized categories identified by Tashakkori and Teddlie (2003) such as interviews, surveys, focus groups, tests, and observations; in addition, discussion forum, platform database, and learning analytics which were added as possible forms of data collection or focuses in MOOC studies. We categorized data collected from the MOOC platform such as enrollment information, clickstream logs, percentage of finished assignments, and video/page views as platform database. For RQ2, to increase the validity, the researchers used five coding categories identified by Veletsianos and Shepherdson (2015); namely: (1) student-focused, (2) teacher-focused, (3) design-focused, (4) context and impact, and (5) other. These five categories effectively encapsulate the diverse areas of research on MOOCs. To answer RQ3, we calculated the locations of all the MOOC authors’ affiliations in this study as well as those for just the first author. For RQ4, the researchers calculated the countries of the MOOC being studied. For the studies which did not specify the location of MOOC delivery, the researchers coded them as “Global.”

Results

This research project has collected MOOC empirical studies which were published in 77 different journals. In total, 26 out of 197 articles (13%) in this review were published in 2014, 68 articles (35%) in 2015, 57 articles (29%) in 2016, and 46 articles (23%) in just the first seven months of 2017. Most articles were published in International Review of Research in Open and Distance Learning (IRRODL) (n=37), Computer & Education...
RQ1: What are the research methods that researchers employed in empirical MOOC studies?

Based on our analysis of the general research methods used in these 146 published MOOC studies in Phase One, 67 articles (46%) were quantitative, 52 studies (36%) were mixed methods, and the remaining 27 articles (18%) were qualitative (see Figure 1). Also detailed in Figure 1, in Phase Two, 22 out of 51 studies (43%) were quantitative, 20 studies (39%) were qualitative, and the remaining nine studies (18%) were mixed methods. Both phases show that quantitative methods were the most frequently used; however, the percentage of studies relying on qualitative methods has been increasing.

Data collection methods. We also analyzed the main data collection methods. In Phase One, most of the empirical MOOCs studies used one (47%) or two (29%) data collection methods. In addition, 17% of the studies (n=33) used three data collection methods, and 7% of them (n=14) used more than three data collection methods. The trends were similar in Phase Two; a majority of MOOCs studies used one (57%) or two (21%) data collection methods.

In Phase One, survey techniques were the used most frequently (n=87) (see Figure 2). Interestingly, 30 out of 146 studies (21%) relied on surveys as the sole data collection method, whereas 34 studies (23%) combined surveys with one other data collection method. Also noted in Figure 2, Phase Two revealed similar results as researchers primarily used surveys (n=25) and platform database (n=16).

Data analysis methods. Besides understanding data collection methods on MOOC research, data analysis is also vital. Figure 3 reveals that descriptive statistics were used the most in Phase One (i.e., 78%) (see also Zhu et al., 2018). Likewise, Figure 3 also shows that descriptive analysis was primarily employed in Phase Two (i.e., 78%). Results also revealed that thematic analysis and learning analytics started to increase. Yet, the qualitative analysis methods were not widely used in MOOC empirical studies.

RQ2: What are the research focuses in MOOC studies?

As described earlier, in terms of the general focus of empirical studies on MOOCs, we divided the research focus into five different categories. In Phase One, studies were primarily student-focused (n=74), followed by design-focused (n=48), context and impact (n=20), instructor-focused (n=5), and other (n=7) (see Figure 4). As noted in Figure 4, studies in Phase Two followed a similar trend, student-focused (n=28), design-focused (n=8), instructor-focused (n=5), context and impact (n=2), and other (n=8). In total, more than half of these studies (n=102) were more related to students. Among those student-focused studies, they tended to target students’ behaviors, motivation, satisfaction, performance, interaction, and retention (see Table 1). However, new topics such as communication patterns, the social structure of the discussion threads, and attitudinal change are emerging. Overall, one-fourth of MOOC studies (n=56) were design-focused and 11% focused on context and impact. Only 5% of MOOC studies were closely related to instructors.

Each research area used a variety of research methods in these 197 studies (see Table 2). Quantitative research was used by 54% of student-focused studies (n=55), 39% of design-focused studies (n=22), and 46% of studies of context and impact (n=10). However, in instructor-focused studies, 7 out of 10 used qualitative methods.

RQ 3: How are researchers of empirical MOOC studies geographically distributed?

Based on the location of the affiliations of the first authors of 197 MOOC studies, the five countries with the most empirical MOOC research were the United States (n=81), the UK (n=27), Spain (n=14), China (n=13), and Australia (n=11) (see Figure 5). In Phase Two, our findings also revealed that most MOOC researchers collaborated with others. Among these 51 studies, 18% (n=9) were independent author, whereas 35% of MOOC studies involved authors collaborating within their own institution and another 37% involved cross-institutional collaboration in the same country (see Figure 6). Only 10% of MOOC studies (n=5) entailed international collaboration.

RQ 4: In terms of the delivery of the MOOC, what are the countries which are attracting the most research?

In this sample of 197 MOOC research studies, most of them described the specific country of MOOC origin, but some studies that we analyzed did not specify the country. The research articles which did not specify the country in which the MOOC was offered typically were conducted by MOOC vendors such as Coursera and edX.
MOOCs from the U.S. were researched the most (n=66) followed by MOOCs from the UK (n=24), Spain (n=14), and China (n=11) (see Figure 7). Forty-four studies were conducted on MOOCs without a clear country of origin.

**Discussion and Significance of this Research**

According to our review of 197 empirical MOOC research articles published between 2014 and 2017, almost half of all MOOC research was conducted using quantitative methods. Such findings are aligned with previous research findings from Veletsianos and Shepherdson (2016) and others. However, the percentage of qualitative studies of MOOCs increased in Phase Two to 39%, compared to just 19% in Phase One research. Survey research was specified as the most frequently used data collection method. Platform database and interviews were the other two data collection methods extensively used in MOOCs. Given the heavy use of surveys and computer log data, it is not too surprising that descriptive statistics were widely used in MOOC research data analysis.

Across both phases of this research, clearly most of the existing research on MOOCs research is concentrated at the student level with quantitative methodology. However, the percentage of the design-focused studies decreased during Phase Two. Clearly, there is a need for MOOC researchers to begin to shift their focus towards more attention on the design of MOOCs from an instructor and instructional designer perspectives. In fact, they might explore the viewpoints of the entire course development, production, and evaluation team as well as those funding such efforts. Of course, many other studies are now vital. For instance, participants in specific MOOC courses intended to lead to skills needed in the twenty-first century, such as marketing analytics, Python programming, entrepreneurship, mental health, creativity, public speaking, business communication, or machine learning (Agarwal, 2017; Gaskell, 2016; Shah, 2017) might be monitored in short-term as well as longitudinal studies related to their employment histories, including promotions, awards, and career satisfaction. Research with employers and supervisors might also prove quite insightful.

In terms of the change of research methods, in Phase Two, the percentage of qualitative studies increased. As a result, the percentage of mixed-methods studies decreased. As the field evolves, however, more research using mixed-methods research may help to triangulate the data and offer deeper insights into decisions related to MOOC quality criteria (Margaryan, Bianco, & Littlejohn, 2015), cultural sensitivity, pedagogical issues (e.g., feedback mechanisms, course interaction components, collaboration strategies, multimedia engagement techniques, etc.), and various assessment practices. Such multi-pronged research can also help with the well-publicized issues related to learner retention rates in MOOCs (Belanger & Thornton, 2013; Chuang & Ho, 2016; Jordan, 2013; MOOC @ Edinburgh, 2013 – Report #1, 2013).

The location of MOOC researchers did not change much between the two phases with research in the United States remaining dominant. What is also clear across both phases of this research, however, is that research on MOOCs remains a growing area of interest across disciplines and regions of the world; those making in-roads have the potential to impact an enormous swath of the human population. Noticing trends in MOOC course design and delivery, course enrollments, and job and career benefits, MOOC researchers could play a pivotal role in not just better understanding the educational impact of MOOCs but in helping to align them with economic needs. As such, research on MOOCs during these two phases and beyond should prove both highly informative and valuable in establishing course and programmatic decisions within colleges and universities, policy directions for governments and other institutions offering MOOCs, and research and grant funding decisions.

We hope that these types of systematic reviews and other forms of research synthesis on MOOCs can lend valuable insights and perspectives. It is difficult to predict the future role of MOOCs, but past research can shed some light on it. Perhaps a decade or two from now, when sufficient experience in the design and delivery of MOOCs has taken place and the technology platforms in which they are offered on have evolved, enrollment in MOOCs will be as common as attending primary school or signing up for a professional development day.

As the existence of MOOCs has reached a decade, a follow-up and extension review is needed after the early MOOC reviews by Liyanagunawardena et al. (2013), Gašević et al. (2014), Veletsianos and Shepherdson (2015, 2016), Deng and Benckendorff (2017), Zhu et al. (2018), and several others. This study built upon our initial analyses by including more recently published MOOC studies and comparing the different phases to examine the trends of this fast growing field. For future research, we recommend expanding methodological approaches in MOOC research. In addition, examining MOOC research during the past decade by year might reveal significant and emerging trends. Some MOOC scholars might find their housing in crosscultural comparison research methods. Others might explore research methods that might be more localized to specific geographic regions, languages, and cultures. As one decade is now complete, we look forward to the next one and the advances and discoveries that will be made in MOOCs and associated MOOC research methods that will impact learners across the planet. For our part, we plan to continue our systematic analyses of MOOC studies and see where our efforts will lead.
References


Preview version of this paper. Content and pagination may change prior to final publication.


Acknowledgements:
The authors wish to explicitly thank Dr. Mimi Miyoung Lee from the University of Houston for her involvement in both phases of this research project, including many timely and valuable suggestions, comments, and insights.

Appendix

Table 1

Specific topical focus of MOOC studies (2014-2017) (n=197)

<table>
<thead>
<tr>
<th>Detailed Focus</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retention and completion/dropout</td>
<td>17</td>
</tr>
<tr>
<td>Motivation</td>
<td>15</td>
</tr>
</tbody>
</table>
Assessment/measurement/evaluation  13  
Engagement  12  
Instructional/MOOC design  12  
Learners’ satisfaction  11  
Communication/interaction  11  
Learners’ experience  10  
Performance/outcome  9  
Professional development  8  
Learners’ attitude  8  
Higher education  8  

Table 2

Research methods used in each research topic (2014-2017) (n=197)

<table>
<thead>
<tr>
<th>Focus</th>
<th>Quantitative</th>
<th>Qualitative</th>
<th>Mixed methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student-focused</td>
<td>55</td>
<td>16</td>
<td>31</td>
</tr>
<tr>
<td>Design-focused</td>
<td>22</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>Context and impact</td>
<td>10</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Instructor-focused</td>
<td>0</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

Note. Some studies have more than one focuses. And we did not include the “other” category in this table.
Figure 1. Research methods used in empirical MOOCs studies (Note: Phase One (2014 – 2016) (n=146); Phase Two (2016 – 2017) (n=51))

Figure 2. Data collection methods used in empirical MOOCs studies (Note: some studies contain more than one data collection method and this figure only includes the main data collection methods)
Figure 3. Specific data analysis methods for MOOC research (Note: some studies contain more than one data analysis method)

![Graph showing research focuses of empirical MOOCs studies]

- Student-focused: 48% (Phase Two), 16% (Phase One)
- Design-focused: 31% (Phase Two), 5% (Phase One)
- Others: 16% (Phase Two), 13% (Phase One)
- Instructor-focused: 10% (Phase Two), 3% (Phase One)
- Context and impact: 4% (Phase Two), 13% (Phase One)

Figure 4. Primary/general focus of MOOC delivery (Note: some studies contain more than one area of focus)

![Bar chart showing the primary areas of focus]

- Context and impact: 55% (Phase Two), 16% (Phase One)
- Instructor-focused: 31% (Phase Two), 10% (Phase One)
- Design-focused: 16% (Phase Two), 5% (Phase One)
- Others: 3% (Phase Two), 4% (Phase One)
- Student-focused: 0% (Phase Two), 13% (Phase One)

Figure 5. The location of the first author of MOOCs studies (2014 – 2017) (n=197) (Note: this figure only includes the main countries)

![Bar chart showing the geographical distribution of first authors]

- The U.S.: 81
- The UK: 27
- Spain: 14
- China: 13
- Australia: 12
- Canada: 7
**Figure 6.** Collaboration among the authors of MOOCs studies (2016 – 2017) (n=51)

**Figure 7.** Countries of MOOC delivery in which the research was conducted (2014 – 2017) (n=197) (Note: this figure only includes the main countries)