

Where is Research on Massive Open Online Courses Headed? A Data Analysis of the MOOC Research Initiative



Dragan Gašević¹, Vitomir Kovanović^{2,1}, Srećko Joksimović^{2,1}, and George Siemens³

¹Athabasca University, Canada, ²Simon Fraser University, Canada, ³University of Texas at Arlington, USA

Abstract

This paper reports on the results of an analysis of the research proposals submitted to the MOOC Research Initiative (MRI) funded by the Gates Foundation and administered by Athabasca University. The goal of MRI was to mobilize researchers to engage into critical interrogation of MOOCs. The submissions – 266 in Phase 1, out of which 78 was recommended for resubmission in the extended form in Phase 2, and finally, 28 funded – were analyzed by applying conventional and automated content analysis methods as well as citation network analysis methods. The results revealed the main research themes that could form a framework of the future MOOC research: i) student engagement and learning success, ii) MOOC design and curriculum, iii) self-regulated learning and social learning, iv) social network analysis and networked learning, and v) motivation, attitude and success criteria. The theme of social learning received the greatest interest and had the highest success in attracting funding. The submissions that planned on using learning analytics methods were more successful. The use of mixed methods was by far the most popular. Design-based research methods were also suggested commonly, but the questions about their applicability arose regarding the feasibility to perform multiple iterations in the MOOC context and rather a limited focus on technological support for interventions. The submissions were dominated by the researchers from the field of education (75% of the accepted proposals). Not only was this a possible cause of a complete lack of success of the educational technology innovation theme, but it could be a worrying sign of the fragmentation in the research community and the need to increased efforts towards enhancing interdisciplinarity.

Keywords: Massive online open courses; MOOC; content analysis; MOOC research analysis; MOOC Research Initiative; education research

Introduction

Massive open online courses (MOOCs) have captured the interest and attention of academics and the public since fall of 2011 (Pappano, 2012). The narrative driving interest in MOOCs, and more broadly calls for change in higher education, is focused on the promise of large systemic change. The narrative of change is some variant of:

Higher education today faces a range of challenges, including reduced public support in many regions, questions about its role in society, fragmentation of the functions of the university, and concerns about long term costs and system sustainability.

In countries like the UK and Australia, broad reforms have been enacted that will alter post-secondary education dramatically (Cribb & Gewirtz, 2013; Maslen, 2014). In the USA, interest from venture capital raises the prospect of greater privatization of universities (GSV Advisors, 2012). In addition to economic questions around the sustainability of higher education, broader socio-demographic factors also influence the future of higher education and the changing diversity of the student population (OECD Publishing, 2013).

Distance education and online learning have been clearly demonstrated to be an effective option to traditional classroom learning¹. To date, online learning has largely been the domain of open universities, separate state and provincial university departments, and for-profit universities. Since the first offering of MOOCs and by elite universities in the US and the subsequent development of providers edX and Coursera, online learning has now become a topical discussion across many campuses². For change advocates, online learning in the current form of MOOCs has been hailed as transformative, disruption, and a game changer (Leckart, 2012). This paper is an exploration of MOOCs; what they are, how they are reflected in literature, who is doing research, the types of research being undertaken, and finally, why the hype of MOOCs has not yet been reflected in a meaningful way on campuses around the world. With a clear foundation of what the type of research actually happening in MOOCs, based on submissions to the MOOC Research Initiative³, we are confident that the conversation about how MOOCs and online learning will impact existing higher education can be moved from a hype and hope argument to one that is more empirical and research focused.

¹ <http://nosignificantdifference.org>

² In this paper, we consider MOOCs to belong to the broader field of online education and learning and that their research should be built on and expand the existing body of research knowledge of online education and learning.

³ <http://www.moocresearch.com>

Massive Online Open Courses (MOOCs)

Massive open online courses (MOOCs) have gained media attention globally since the Stanford MOOC first launched in fall of 2011. The public conversation following this MOOC was unusual for the education field where innovations in teaching and learning are often presented in university press releases or academic journals. MOOCs were prominent in the *NY Times*, *NPR*, *Time*, *ABC News*, and numerous public media sources. Proclamations abounded as to the dramatic and significant impact that MOOCs would have on the future of higher education. In early 2014, the narrative has become more nuanced and researchers and university leaders have begun to explore how digital learning influences on campus learning (Kovanović, Joksimović, Gašević, Siemens, & Hatala, 2014; Selwyn & Bulfin, 2014). While interest in MOOCs appears to be waning from public discourse, interest in online learning continues to increase (Allen & Seaman, 2013). Research communities have also formed around learning at scale⁴ suggesting that while the public conversation around MOOCs may be fading, the research community continues to apply lessons learned from MOOCs to educational settings.

MOOCs, in contrast to existing online education which has remained the domain of open universities, for-profit providers, and separate departments of state universities, have been broadly adopted by established academics at top tier universities. As such, there are potential insights to be gained into the trajectory of online learning in general by assessing the citation networks, academic disciplines, and focal points of research into existing MOOCs. Our research addresses how universities are approaching MOOCs (departments, research methods, and goals of offering MOOCs). The results that we share in this article provide insight into how the gap between existing distance and online learning research, dating back several decades, and MOOCs and learning at scale research, can be addressed as large numbers of faculty start experimenting in online environments.

MOOC Research

Much of the early research into MOOCs has been in the form of institutional reports by early MOOC projects, which offered many useful insights, but did not have the rigor – methodological and/or theoretical expected for peer-reviewed publication in online learning and education (Belanger & Thornton, 2013; McAuley, Stewart, Siemens, & Cormier, 2010). Recently, some peer reviewed articles have explored the experience of learners (Breslow et al., 2013; Kizilcec, Piech, & Schneider, 2013; Liyanagunawardena, Adams, & Williams, 2013). In order to gain an indication of the direction of MOOC research and representativeness of higher education as a whole, we explored a range of articles and sources. We settled on using the MOOC Research Initiative as our dataset.

⁴ <http://learningatscale.acm.org>

MOOC Research Initiative (MRI)

The MOOC Research Initiative was an \$835,000 grant funded by the Bill & Melinda Gates Foundation and administered by Athabasca University. The primary goal of the initiative was to increase the availability and rigor of research around MOOCs. Specific topic areas that the MRI initiative targeted included: i) student experiences and outcomes; ii) cost, performance metrics and learner analytics; iii) MOOCs: policy and systemic impact; and iv) alternative MOOC formats. Grants in the range of \$10,000 to \$25,000 were offered. An open call was announced in June 2013. The call for submissions ran in two phases: 1. short overviews of 2 pages of proposed research including significant citations; 2. full research submissions, 8 pages with influential citations, invited from the first phase. All submissions were peer reviewed and managed in Easy Chair. The timeline for the grants, once awarded, was intentionally short in order to quickly share MOOC research. MRI was not structured to provide a full research cycle as this process runs multiple years. Instead, researchers were selected who had an existing dataset that required resources for proper analysis.

Phase one resulted in 266 submissions. Phase two resulted in 78 submissions. A total of 28 grants were funded. The content of the proposals and the citations included in each of the phases were the data source for the research activities detailed below.

Research Objectives

In this paper, we report the findings of an exploratory study in which we investigated (a) the themes in the MOOC research emerging in the MRI proposals; (b) research methods commonly proposed for use in the proposals submitted to the MRI initiative, (c) demographics (educational background and geographic location) characteristics of the authors who participated in the MRI initiative; (d) most influential authors and references cited in the proposals submitted in the MRI initiative; and (e) the factors that were associated with the success of proposals to be accepted for funding in the MRI initiative.

Methods

In order to address the research objectives defined in the previous section, we adopted the content analysis and citation network analysis research methods. In the remainder of this section we describe both of these methods.

Content Analysis

To address research objectives a and b, we performed content analysis methods. Specifically, we performed both automated a) and manual b) content analyses. The choice of content analysis was due to the fact that it provides a scientifically sound

method for conducting an objective and systematic literature review, thus enabling for the generalizability of the conclusions (Holsti, 1969). Both variations of the method have been used for analysis of large amounts of textual content (e.g., literature) in educational research.

Automated content analysis of research themes and trends.

Given that content analysis is a very costly and labor intensive endeavor, the automation of content analysis has been suggested by many authors and this is primarily achieved through the use of scientometric methods (Brent, 1984; Cheng et al., 2014; Hoonlor, Szymanski, & Zaki, 2013; Kinshuk, Huang, Sampson, & Chen, 2013; Li, 2010; Sari, Suharjito, & Widodo, 2012). Automated content analysis assumes the application of the computational methods – grounded in natural language processing and text mining – to identify key topics and themes in a specific textual corpus (e.g., set of documents, research papers, or proposals) of relevance for the study. The use of this method is especially valuable in cases where the trends of a large corpus need to be analyzed in “real-time”, that is, short period of time, which was the case of the study reported in this paper and specifically research objective c. Not only is the use of these automated content analysis methods cost-effective, but it also lessens the threats to validity and issues of subjectivity that are typically associated with the studies based on content analysis. Among different techniques, the one based on the word co-occurrence – that is, words that occur together within the same body of written text, such as research papers, abstracts, titles or parts of papers – has been gaining the widespread adoption in the recent literature reviews of educational research (Chang, Chang, & Tseng, 2010; Cheng et al., 2014). As such, the use of automated content analysis was selected for addressing research objective c.

In order to perform a content analysis of the MRI submissions, we used particular techniques adopted from the disciplines of machine learning and text mining. Specifically, we based our analysis approach on the work of Chang et al. (2010) and Cheng et al. (2014). Generally speaking, our content analysis consisted of the three main phases:

1. extraction of relevant *key concepts* from each submission,
2. *clustering* submissions to the important research themes, and
3. *in-depth analysis* of the produced clusters.

For extraction of key concepts from each submission, we selected Alchemy, a platform for semantic analyses of text that allows for extraction of the informative and relevant set of concepts of importance for addressing research objective c, as outlined in Table 1. In addition to the list of relevant concepts for each submission, Alchemy API produced the associated relevance coefficient indicating the importance of each concept for a given submission. This allowed us to rank the concepts and select the top 50 ranked

concepts for consideration in the study. In the rare cases when Alchemy API extracted less than 50 concepts, we used all of the provided concepts.

After the concept extraction, we used the agglomerative hierarchical clustering in order to define N groups of similar submissions that represent the N important research themes and trends in MOOC research, as aimed in research objective c. Before running the particular clustering algorithm we needed to: i) define a representation of each submission, ii) provide a similarity measure that is used to define submission clusters, and iii) choose appropriate number of clusters N . As we based the clustering on the extracted keywords using Alchemy API, our representation of each submission was a vector of concepts that appeared in a particular submission. More precisely, we created a large submission-concept matrix where each row represented one submission, and each column represented one concept, while the values in the matrix (M_{IJ}) represented the relevance of a particular concept J for a document I . Thus, each submission was represented as an N -dimensional row vector consisting of numbers between 0.00 and 1.00 describing how relevant each of the concepts was for a particular submission. The concepts that did not appear in the particular submission had a relevance zero, while the concepts that were actually present in the submission text had a relevance value greater than zero and smaller or equal to one.

With respect to the similarity measure, we used the popular cosine similarity which is essentially a cosine of the angle θ between the two submissions in the N -dimensional space defined by all unique concepts. It is calculated as dot product of two vectors divided by the products of their ℓ_2 norms. For two submissions A and B , and with the total of n different concepts (i.e., the length of vectors A and B was n – the number of concepts extracted from A and B), it is calculated as follows:

$$\text{similarity}(A, B) = \cos(\theta) = \frac{A \times B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

Agglomerative hierarchical clustering algorithms work by iteratively merging smaller clusters until all the documents are merged into a single big cluster. Initially, each document is in a separate cluster, and based on the provided similarity measure the most similar pairs of clusters are merged into one bigger cluster. However, given that the similarity measure is defined in terms of two documents, and that clusters typically consist of more than one document, there are several strategies of measuring the similarity of clusters based on the similarity of the individual documents within clusters. We used the GAAClusterer (i.e., Group Average Agglomerative) hierarchical clustering algorithm from the NLTK python library that calculates the similarity between each pair of clusters by averaging across the similarities of all pairs of documents from two clusters.

Table 1

Concept Categories for Describing Clusters

| Category | Description | Example |
|------------------------|--|--|
| Topics | The most frequent keywords that identify topics mentioned in the specific cluster. | Intelligent tutoring systems; Educational technology; Networked contexts |
| Theory/ Approach | Keywords that identify specific theory recognized within documents in each cluster. | Competence-based education; Social constructivist method |
| Environment | MOOC platform identified within the cluster. | Coursera; edX; MiriadaX |
| Domain | Keywords that represent a specific domain of a MOOC course. | STEM disciplines; Red Cross; Health Sciences |
| Data sources | Keywords representing data used for studies within the cluster. | Engagement data; Qualitative data; Study logs |
| Measures and variables | Keywords representing measures used for studies within the cluster. | Student outcome measures; Early motivation measures; |
| Analysis techniques | Keywords representing various analysis used for studies within the cluster. | Parallel multi-method analysis; Nonparametric statistical analysis; |
| Research instruments | Keywords representing various instruments used to collect data for studies within the cluster. | In-depth interviews; Focus group interview; Questionnaire |
| Use of control group | Identifies whether Control groups are used in at least one study within the cluster. | Control group |

The output of the clustering algorithm was a tree, which described the complete clustering process. We evaluated manually the produced clustering tree to select the clustering solution with the N most meaningful clusters for our concrete problem. In the phase one of the MRI granting process we discovered nine clusters, while in the second phase we discovered five clusters.

Finally, in order to assess the produced clusters and select the key concepts in each cluster, we created a concept-graph consisting of the important concepts from each cluster. The nodes in a graph were concepts discovered in a particular cluster, while the links between them were made based on the co-occurrence of the concepts within the

same document. More precisely, the undirected link between two concepts was created in case that both of them were extracted from the same document. To evaluate the relative importance of each concept we used the betweenness centrality measure, as the key concepts are likely the ones with the highest betweenness centrality. Besides the ranking of the concepts in each cluster based on their betweenness centrality, we manually classified all important concepts into one of the several categories that are shown in Table 1. Provided categories represent important dimensions of analysis and we describe each of the clusters based on the provided categories of key concepts. Thus, when we describe a particular cluster, we cover all of the important dimensions to provide the holistic view of the particular research trend that is captured in that cluster.

Content analysis of important characteristics of authors and submissions.

A manual content analysis of the research proposals was performed in order to address research objective b. The content analysis afforded for a systematic approach to collect data about the research methods and the background of the authors. These data are then used to cross-tabulate with the research themes found in the automated content analysis (i.e., research objective a) and citation analysis (i.e., research objective c). Specifically, each submission was categorized into one of the four categories in relation to research objective a:

1. *qualitative method*, which meant that the proposal used a qualitative research method such as grounded theory;
2. *quantitative method*, which meant that a proposal followed some of the quantitative research methods on data collected through (Likert-scale based) surveys or digital traces recorded by learning platforms in order to explore different phenomena or test hypotheses;
3. *mixed-methods*, which reflected a research proposals that applied some combination of qualitative and quantitative research methods;
4. *other*, which comprised of the research proposals that did not explicitly follow any of these methods, or it was not possible to determine from their content which of the three methods they planned to use.

For all the authors⁵ of submitted proposals to the MRI initiative, we collected the information related to their home discipline and the geographic location associated with their affiliation identified in their proposal submissions in order to address research objective c. Insight into researchers' home discipline was obtained from the information provided with a submission (e.g., if a researcher indicated to be affiliated with a school

⁵ Information about the geographic location as extracted from the application forms submitted by the authors to EasyChair, a software system used for the submission and review process.

of education, we assigned education as the home discipline for this research). In cases when such information was not available directly with the proposal submission, we performed a web search, explored institutional websites, and consulted social networking sites such as LinkedIn or Google Scholar.

Citation Analysis and Success Factors

The citation analysis was performed to address research objective d. It entailed the investigation of the research impact of the authors and papers cited in the proposals submitted to the MRI initiative (Waltman, van Eck, & Wouters, 2013). In doing so, the counts of citations of each reference and author, cited in the MRI proposals, are used as the measures of the impact in the citation analysis. This method was suitable, as it allowed for assessing the influential authors and publications in the space of MOOC research.

Citation network analysis – the analysis of so-called co-authorship and citation networks have gained much adoption lately (Tight, 2008) – was performed in order to assess the success factor of individual proposals to be accepted for funding in the MRI initiative, as set in research objective e. This way of gauging the success was a proxy measure of the quality and importance of the proposals, as aimed in research objective e. As such, it was appropriate to be used as an indicator of specific topics based on the assessment of the international board of experts who reviewed the submitted proposals.

Social network analysis was used to address research objective e. In this study, social networks were created through the links established based on the citation and co-authoring relationships, as explained below. The use of social network analysis has been shown as an effective way to analyze professional performance, innovation, and creativity. Actors occupying central networks nodes are typically associated with the higher degree of success, innovation, and creative potential (Burt, Kilduff, & Tasselli, 2013; Dawson, Tan, & McWilliam, 2011). Moreover, structure of social networks has been found as an important factor of innovation and behavior diffusion. For example, Centola (2010) showed that the spread of behavior was more effective in networks with higher clustering and larger diameters. Therefore, for research objective e, we expected to see the association between the larger network diameter and the success in receiving funding.

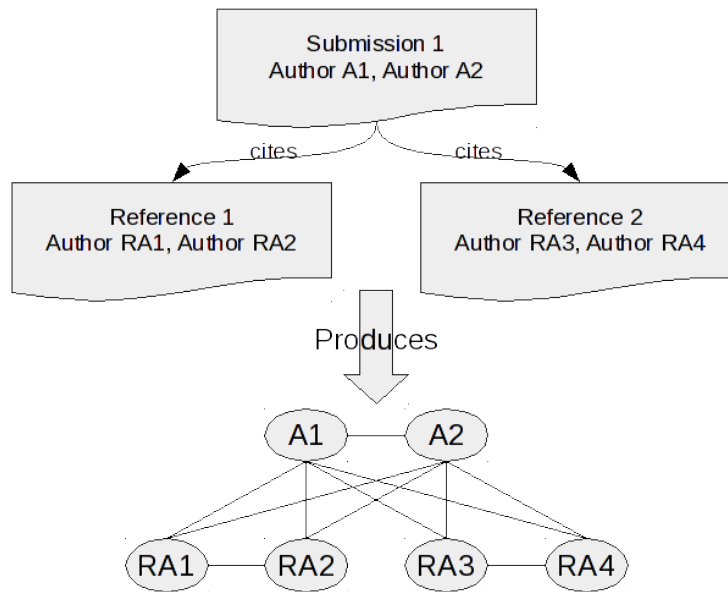


Figure 1. The citation networks – connecting the authors of a research proposal (A1 and A2) with the authors of two cited references (RA1, RA2, RA2 and RA4).

In this study, we followed a method for citation network analysis suggested by Dawson et al. (2014) in their citation network analysis of the field of learning analytics. Nodes in the network represent the authors of both submissions and cited references, while links are created based on the co-authorship and citing relations. Figure 1 illustrates the rules for creating the citation networks in the simple case when a submission written by the two authors references two sources, each of them with two authors as well.

We created a citation network for each cluster separately and analyzed them by the following three measures commonly used in social network analysis (Bastian, Heymann, & Jacomy, 2009; Freeman, 1978; Wasserman, 1994):

1. *degree*, the number of edges a node has in a network,
2. *diameter*, the maximum eccentricity of any node in a network, and
3. *path*, the average graph-distance between all pairs of nodes in a network.

All social networking measures were computed using the Gephi open source software for social network analysis (Bastian et al., 2009). The social networking measures of each cluster were then correlated (Spearman's ρ) with the acceptance ratio – computed as a ratio of the number of accepted proposals and the number of submitted proposals – for both phases of the MRI initiative.

Results

Phase 1 Results

Phase 1 research themes.

In order to evaluate the direction of the MOOC related research, we looked at the most important research themes in the submitted proposals. Table 2 shows the detailed descriptions of the discovered research themes and their acceptance rates, primary research fields of authors, as well as the average number of authors and citations on each submission. In total, there were nine research themes with a similar number of submissions, from 19 (i.e., “Mooc Platforms” research theme) to 40 (i.e., “Communities” and “Social Networks” research themes). Likewise, submissions from all themes had on average slightly more than 2 authors and from 7 to 9 citations. However, in terms of their acceptance rates, we can see much bigger differences. More than half of the papers from the “Social Networks” research theme moved to the second phase and finally 25% of them were accepted for funding, while none of the submissions from the “Education Technology Improvements” theme was accepted for funding.

Furthermore, Table 3 shows the main topics and research approaches used in each research theme, while Table 4 shows the most important methodological characteristics of each research theme.

Table 2

Phase 1 Research Themes

| Theme | Size | Accepted 2 nd round | Accepted funding | Authors avg. (SD) | Citations avg. (SD) | Major fields |
|--|------------|-----------------------------------|------------------------|-------------------------|------------------------|--|
| Cluster 1 Ed. Tech. Improvements | 23 | 4 (17.4 %) | 0 (0 %) | 2.7 (1.1) | 7.3 (3.6) | Education (36) Business (8) |
| Cluster 2 Processes | 26 | 10 (38.5 %) | 2 (7.7 %) | 2.6 (1.7) | 6.2 (2.8) | Education (38) Computer Science (8) |
| Cluster 3 High Ed. Institutions and MOOCs | 25 | 5 (20.0 %) | 1 (4.0 %) | 2.1 (1.1) | 9.0 (5.5) | Education(16) Social Sciences (9) |
| Cluster 4 Motivation and Behavioral Patterns | 29 | 13 (44.8 %) | 4 (13.8 %) | 2.1 (0.9) | 6.9 (4.6) | Education (29) Computer Science (8) |
| Cluster 5 Mobile and Adaptive Learning | 35 | 8 (22.9 %) | 4 (11.4 %) | 2.2 (1.2) | 8.3 (6.3) | Education (27) Computer Science (8) |
| Cluster 6 Learner Performance | 24 | 5 (20.8 %) | 2 (8.3 %) | 2.4 (1.5) | 8.3 (6.6) | Education (18) Industry (10) |
| Cluster 7 MOOC Platforms | 19 | 2 (10.5 %) | 1 (5.3 %) | 2.2 (1.1) | 9.1 (7.0) | Education (13) Technology (6) Industry (6) |
| Cluster 8 Communities | 40 | 9 (22.5 %) | 4 (10.0 %) | 2.3 (1.2) | 6.8 (4.8) | Education (42) Industry (15) |
| Cluster 9 Social Networks | 40 | 22 (55.0 %) | 10 (25.0 %) | 2.2 (1.2) | 8.3 (5.9) | Education (34) Computer Science (15) |
| Total | 261 | 78 (29.9 %) | 28 (10.7 %) | | | |

Table 3

Phase 1 Research Themes Topics and Theoretical Approaches

| Theme | Topics | Theoretical approaches |
|--|--|---|
| Cluster 1 Ed. Tech. Improvements | Intelligent tutoring systems Educational technology Networked contexts Deeper learning experience | Behavioral leadership theory Grounded theory Data-driven approach Design-based research Rapid prototyping approach |
| Cluster 2 Processes | Teaching-learning process Intellectual property issues Collaborative learning Forum discussion Social learning approach Self-regulated learning Learner engagement | Connectivist approach Descriptive research study Mixed method approach Thematic analysis Semiotic social theory Agile development models Longitudinal research |
| Cluster 3 High Ed. Institutions and MOOCs | Student perception Student achievement Highly-motivated students Higher education Online social worlds Collaborative activity | Competence-based education Social constructivist method Cognitive-behaviorist approach Innovation diffusion theory Ethnographic approach Flipped classroom style class |
| Cluster 4 Motivation and Behavioral Patterns | Student engagement Discussion forum entries Student motivation Student behavioral patterns Social media Blended learning courses Retention analysis | Exploratory study Cognitive science research Field research methods Flipped classroom model Problem based learning Theory of planned behavior |
| Cluster 5 Mobile and Adaptive Learning | Collaboration Mobile learning Content drop-out pattern Social networking Emergent learning Personal learning env. Learner engagement | Social learning theory Thematically based approach Social psychology Action research MSLQ cognitive strategy phenomenological study Flipped classroom concept |
| Cluster 6 Learner Performance | Personality data Educational technology Student demographics Course completion Student performance Gamification techniques | Flipped Classroom model Problem-based learning |
| Cluster 7 MOOC Platforms | Traditional education Instructional design Higher education practice xMOOC model | Problem-based learning Blended learning approach Psychometric theory Design-based research |
| Cluster 8 Communities | Online communities Discussion forums Completion rates Educational technology Self-directed learning | Ethnographic approach Mixed methods Design-based research approach Evidence-Based Learning Networked Learning Framework |

| | | |
|------------------------------|--|--|
| | Formal-learning environment Technology-enhanced learning Innovative business models Better retention | Behaviourism theory Connectivist theory |
| Cluster 9 Social Networks | Social network analysis Peer-to-peer interaction Peer assessment Student success Higher education Peer tutoring Discussion Forums Social learning Student motivation | Interdisciplinary approach Phenomenological study Design Based Research Flipped classroom Game theory simulation Actor network theory |

Table 4

Phase 1 Research Themes Data Analysis Characteristics

| Theme | Data sources and measures | Analysis techniques | Instruments |
|---|--|---|--|
| Cluster 1 Ed. Tech. Improvements | Engagement data Study logs Activity logs Feedback data Student success measures | Post-test implementation surveys Data classification Association rule mining Granular taxonomy Big data analytics | In-depth interviews Focus group interviews Online surveys |
| Cluster 2 Processes | Conversational data Narrative data Clickstream data Linguistic data Formative evaluation data Open research data | Cross-case analysis Critical literature survey Interactive language analysis Discourse analysis | Self-assessment instruments Focus groups Instructor survey Student surveys |
| Cluster 3 High Ed. Institutions and MOOCs | Social Media Rich qualitative data MOOC-related data Descriptive data Field data Post-instruction outcome measures | Meta-analysis method Focused content analysis Comparative analysis Meta-narrative analysis | Focus groups Interviews Survey instruments Questionnaires Participant observation Field notes |
| Cluster 4 Motivation and Behavioral Patterns | International mobility statistics Web traffic statistics Performance data Tracking log data Behavioral data Observational data Clickstream data Student outcome | Graph analysis Deep linguistic analyses Behavioral analysis Structural analysis Natural language processing Time series analysis | Interviews Student surveys Quizzes |

| | | | |
|---|---|--|--|
| | measures Early motivation measures | | |
| Cluster 5 Mobile and Adaptive Learning | Activity log data Discursive data Email tracking data Social graph data Client-side offline data Social psychological measures | Online ethnography Trace analysis | Surveys Questionnaires Participant observations Phenomenological inquiry |
| Cluster 6 Learner Performance | Student survey data Clickstream data Student performance data Learner data Activity logs | Latent Dirichlet analysis Comparative analysis Clickstream analytics Learner analytics Comparative analytics | Memorization tests Interviews Surveys Focus groups Feedback questionnaires |
| Cluster 7 MOOC Platforms | Log data | Performance data analysis Content analysis | Surveys Interviews Self-assessments Performance assessment Summative assessment |
| Cluster 8 Communities | Interview transcripts Online artifacts Assessment data | In-depth analysis Text Analysis Systematic discourse analysis Frame analysis Critical analysis | Focus groups Surveys Semi-structured interviews |
| Cluster 9 Social Networks | Learner interaction data Phenomenological data EEG-MOOC usage data Course completion data Engagement measures | Cross-case analysis Phenomenological analysis Evidence-based research Content analysis | Exit surveys Qualitative surveys Phenomenological interviews Phenomenological inquiry Interviews End-of-course surveys |

Table 5

Phase 1 Distribution of Research Methodologies

| Methodology | Submissions | Authors avg. (SD) | Citations avg. (SD) |
|--------------|---------------|-------------------|---------------------|
| Mixed | 96 (36.2%) | 2.4 (1.3) | 8.2 (5.0) |
| Qualitative | 74 (27.9%) | 2.1 (1.1) | 8.6 (6.4) |
| Quantitative | 80 (30.2%) | 2.4 (1.3) | 6.6 (4.8) |
| Unknown | 15 (5.7%) | 1.7 (0.9) | 7.1 (5.0) |
| Total | 265 (100.00%) | 2.3 (1.2) | 7.7 (5.4) |

Phase 1 research methods.

Table 5 shows the distribution of submissions per each methodology together with the average number of authors and citations per submission. Although the observed differences are not very large, we can see that the most common research methodology type is mixed research, while the purely qualitative research is the least frequent.

Phase 1 demographic characteristics of the authors.

Table 6 also shows the five most common primary research fields for submission authors. Given that some of the authors were not from academia, we included an additional field entitled “Industry” as a marker for all researchers from the industry field. We can see that researchers from the field of education represent by far the biggest group, followed by the researchers from the industry and computer science fields. Table 7⁶ shows a strong presence of the authors of the proposals from North America in Phase 1. They are followed by the authors from Europe and Asia, who combined had a much lower representation than the authors from North America. The authors from other continents had a much smaller presence, with very low participation of the authors from Africa and South America and with no author from Africa who made it to Phase 2.

⁶ The numbers of authored and accepted proposals are decimal, as some proposals had authors from different continents. For example, if a proposal had two authors from North America and one author from Africa, the number of authored proposals for North America would be 0.67 and for Africa 0.33.

Table 5
Phase 1 Top 5 Research Fields

| Field | Authors |
|------------------|---------|
| Education | 251 |
| Industry | 58 |
| Computer Science | 58 |
| Social Sciences | 32 |
| Engineering | 30 |

Table 6
Phase 1 Geographic Distribution of the Authors

| Continent | Authors | Authored proposals | Accepted proposals |
|---------------|---------|--------------------|--------------------|
| Africa | 4 | 3 | 0 |
| Asia | 87 | 34.38 | 3.67 |
| Australia/NZ | 23 | 10.33 | 6 |
| Europe | 137 | 60.51 | 15.83 |
| North America | 305 | 153.26 | 54.5 |
| South America | 9 | 4.5 | 1 |

Phase 1 citation analysis.

With respect to citation analysis, we extracted the list of most cited authors and papers. We counted an author's – authors of both MIR submissions and the papers cited in the submissions were included – citations as a sum of all of the authors' paper citations, regardless of whether the author was the first author or not. Figure 2 shows the list of most cited authors, while Table 8 shows the list of most cited papers in the first phase of the MRI initiative.

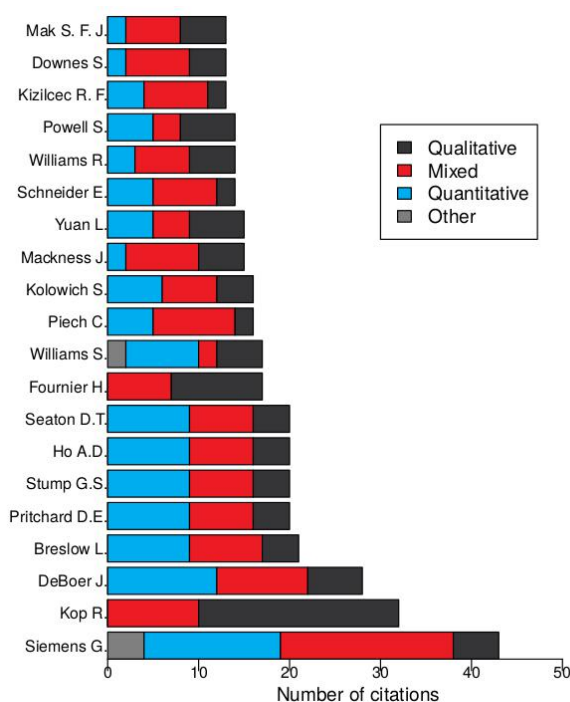


Figure 2. Phase 1 most cited authors.

Table 7

Phase 1 Most Cited papers

| Paper name | Citation count |
|---|----------------|
| Breslow, L., Pritchard, D., DeBoer, J., Stump, G., Ho, A. and Seaton, D. (2013). <i>Studying Learning in the Worldwide Classroom: Research into edX's First MOOC.</i> | 28 |
| Yuan, L. and Powell, S. (2013). <i>MOOCs and open education: Implications for higher education.</i> | 14 |
| Kizilcec, R. F., Piech, C. and Schneider, E. (2013). <i>Deconstructing Disengagement: Analyzing Learner Subpopulations in Massive Open Online Courses.</i> | 14 |
| Kop, R., Fournier, H. and Sui Fai Mak, J. (2011). <i>A pedagogy of abundance or a pedagogy to support human beings? Participant support on Massive Open Online Courses.</i> | 14 |
| Siemens, G. (2005). <i>Connectivism: A Learning Theory for the Digital Age.</i> | 13 |
| Daniel, J., (2012). <i>Making sense of MOOCs: Musings in a maze of myth, paradox and possibility.</i> | 13 |
| Mackness, J., Mak, S. and Williams, R. (2010). <i>The ideals and reality of participating in a MOOC.</i> | 11 |
| Pappano, L. (2012). <i>The year of the MOOC.</i> | 9 |

Finally, we extracted for each research theme a citation network from all Phase 1 submissions. Table 9 shows the graph centrality measures for the citation networks of each of the research themes.

Phase 1 success factors.

We looked at the correlations between the centrality measures of citation networks (Table 9) and the second phase acceptance rates. Spearman's rho revealed that there was a statistically significant correlation between the citation network diameter and number of submissions accepted into the second round ($\rho_s = .77$, $n=9$, $p < .05$), a statistically significant correlation between citation network diameter and second round acceptance rate ($\rho_s = .70$, $n=9$, $p < .05$), and a statistically significant correlation between citation network path and number of submissions accepted into the second round ($\rho_s = .76$, $n=9$, $p < .05$). In addition, a marginally significant correlation between citation network path length and second phase acceptance rate was also found ($\rho_s = .68$, $n=9$, $p = 0.05032$). These results confirmed the expectation stated in the citation analysis section that research proposals with the broader scope of the covered literature were more likely to be assessed by the international review board as being of higher quality and importance. Further implications of this result are discussed in the Discussion section.

Table 8

Phase 1 Citation Network Metrics

| Theme | Average degree (SD) | Diameter | Average shortest path (SD) | Density |
|---|---------------------|----------|----------------------------|---------|
| Cluster 1 Ed. Tech. Improvements | 4.8 (6.2) | 6 | 3.1 (1.6) | 0.018 |
| Cluster 2 Processes | 5.4 (5.7) | 12 | 4.6 (2.3) | 0.026 |
| Cluster 3 High Ed. Institutions and MOOCs | 4.2 (6.3) | 8 | 4.0 (1.5) | 0.021 |
| Cluster 4 Motivation and Behavioral Patterns | 3.6 (5.6) | 9 | 5.6 (2.3) | 0.013 |
| Cluster 5 Mobile and Adaptive Learning | 4.8 (7.7) | 8 | 3.8 (1.2) | 0.016 |
| Cluster 6 Learner Performance | 5.5 (7.1) | 7 | 3.9 (1.8) | 0.028 |
| Cluster 7 MOOC Platforms | 5.6 (8.9) | 8 | 4.1 (1.9) | 0.026 |
| Cluster 8 Communities | 5.7 (5.7) | 10 | 4.6 (1.8) | 0.023 |
| Cluster 9 Social Networks | 4.3 (7.1) | 10 | 5.1 (2.0) | 0.01 |
| Total | 5.8 (8.7) | 17 | 5.2 (1.5) | 0.003 |

Phase 2 Results

Following the analysis of the first phase of MRI, we analyzed the total of 78 submissions that were accepted into the second round of evaluation.

Phase 2 research themes.

Following the analysis of popular research themes, we applied the same automated content analysis method to the submissions that were accepted into the second phase. We found five research themes (Table 10) that were the focus of an approximately similar number of submissions. In order to give a better insight in the discovered research themes, we provide a list of extracted keywords which were related to the topic of investigation and their theoretical approaches (Table 11), and also a list of extracted keywords related to the data sources, analysis techniques, and used metrics (Table 12).

Research theme 1: engagement and learning success

The main topics in this cluster are related to learners' participation, engagement, and behavioral patterns in MOOCs. Submissions in this cluster aimed to reveal the most suitable methods and approaches to understanding and increasing retention, often relying on peer learning and peer assessment. Studies encompassed a wide variety of courses (e.g., biology, mathematics, writing, EEG-enabled courses, art, engineering, mechanical) on diverse platforms. However, most of the courses, used in the studies from this cluster, were offered on the Coursera platform.

Table 9

Phase 2 Research Themes

| Theme | Size | Accepted funding | Authors avg. (SD) | Citations avg. (SD) | Major Fields | Qualitative | Mixed | Quantitative |
|--|-----------|--------------------|-------------------|---------------------|--|-------------|-------|--------------|
| Cluster 1 Engagement and Learning Success | 14 | 6 (42.9 %) | 2.2 (1.3) | 15.0 (9.8) | Education (14) Computer Science (4) Engineering(3) | 1 | 3 | 10 |
| Cluster 2 MOOC Design and Curriculum | 14 | 2 (14.3 %) | 2.9 (2.1) | 20.2 (13.7) | Education (19) Computer Science (7) Engineering(4) | 3 | 5 | 6 |
| Cluster 3 Self-Regulated Learning and Social Learning | 15 | 6 (40.0 %) | 2.3 (0.9) | 21.7 (9.2) | Education(25) Computer Science (3) | 8 | 6 | 1 |
| Cluster 4 SNA and Networked Learning | 19 | 9 (47.4 %) | 2.1 (0.8) | 20.7 (15.6) | Education (23) Computer Science (5) | 2 | 12 | 5 |
| Cluster 5 Motivation, Attitude and Success Criteria | 16 | 5 (31.2 %) | 2.8 (1.1) | 23.1 (9.2) | Education (25) Engineering (5) Social Sciences(4) | 5 | 7 | 4 |
| Total | 78 | 28 (35.8 %) | | | | | | |

Table 10

Phase 2 Topics and Theoretical Approaches of Discovered Research Themes

| Theme | Topics | Theoretical approaches |
|---|---|---|
| Cluster 1 Engagement and Learning Success | Student engagement Academic progress User behavior Actual participation Peer assessment High school students | Theory of planned behavior Motivational messages Flipped Classroom |
| Cluster 2 MOOC Design and Curriculum | Collaborative practices Participant observation Higher education Course implementation models Program evaluation Student-level analytics MOOC design Treatment group Online discussions Learning behavior | Flipped Classroom Interest-oriented learning Community-based learning Quality education resources Self-regulated learning Constitutive complexity theory Self-directed online learning MOOCulus HMM approach CoI framework Social interdependence theories |
| Cluster 3 Self-regulated and Social Learning | Social sciences Higher education Self-regulated learning At-risk learners Social learning Educational resources | Complexity theory Social learning theory Self-regulated learning Instructional design research Self-determination theory Goal theory Flipped classrooms |
| Cluster 4 SNA and Networked Learning | Social network analysis Learners interaction Higher education Discussion forums Online interactions Specific learner profiles Network formulation Asynchronous interaction Network structure P2P interactions | CSCL Summative assessment strategy Design-based research approach Complex connectivist learning Social Cognitive Theory Simple topic modeling Mixed Membership Stochastic Blockmodels |
| Cluster 5 Motivation, Attitude and Success Criteria | Learner motivation Intrinsic motivation Learning design Completion rates Teaching strategies High satisfaction rates Faculty attitude Evaluation plans | Data elicitation methodology Agile research methodology Adaptive learning design Actor network theory |

Research theme 2: MOOC design and curriculum

Research proposals in this cluster were mostly concerned with improving learning process and learning quality and with studying students' personal needs and goals. Assessing educational quality, content delivery methods, MOOC design and learning conditions, these studies aimed to discover procedures that would lead to better MOOC design and curriculum, and thus improving learning processes. Moreover, many visualization techniques were suggested for investigation in order to improve learning quality. Courses suggested for the use in the proposed studies from this cluster were usually delivered by using the edX platform and the courses were in the fields of mathematics, physics, electronics and statistics. The cluster was also characterized by a diversity of data types planned for collection – from surveys, demographic data, and grades to engagement patterns and to data about brain activity.

Table 11

Phase 2 Research Characteristics of Discovered Research Themes

| Theme | Data sources and measures | Analysis techniques/instruments |
|---|--|--|
| Cluster 1 Engagement and Learning Success | Students demographic characteristics EEG dataset TBP measures SAT scores Final grading score Mental state EEG brain activity Engagement patterns Latent patterns | Qualitative peer assessment Unsupervised learning Probabilistic Soft Logic Design-based research approach MOOC-scale peers grading Surveys Wireless EEG headset Quizzes Pre/post-tests |
| Cluster 2 MOOC Design and Curriculum | Student achievement data edX user data Case study data Assessment data Trace data Complex SQL data Activity Summary Data Preliminary clickstream analysis Complete clickstream data Archival data Educational metrics Students time allocation Students active participation | Assessment-based outcome measures Hidden Markov model Survey Interviews Qualitative field work Post-course surveys Open-ended narrative questions Student background surveys |
| Cluster 3 Self-regulated and Social Learning | Online discourses Survey responses Course behavior data Discussion forum data Diversity-related learning outcomes | Frame analysis Critical discourse analysis Content analysis Empirical qualitative research Association rule mining Mindset survey question |

| | | |
|---|---|--|
| | Mindset score | Qualitative research interviews Entry survey In-depth interviews |
| Cluster 4 SNA and Networked Learning | Qualitative data collection Transactional data Social media data MOOC interaction data Click stream data Network analysis data Descriptive data Interactional data Course outcome data Coursera-based course data Longitudinal network data Longitudinal relational data Completion data Social graph data MOOCs learner metrics personality metrics social metrics standard statistic measure | Survival analysis Mixed research methods Collaborative Behaviors Analysis Interaction analysis Post-course data analysis Qualitative analysis Scale data analysis Probabilistic graphical models Text mining techniques with social network Learner analytics Quantitative research methods Real time analysis Focus groups Interviews Surveys D questionnaires Small group interviews |
| Cluster 5 Motivation, Attitude and Success Criteria | Publicly available data Course activity data Qualitative data Student performance data | Classification Confirmatory factor analysis Trace analysis Cluster analysis Structural equation modeling Qualitative data analysis Case study approach Interviews Surveys Open-ended assignments |

Research theme 3: Self-regulated learning and social learning

Self-regulated learning, social learning, and social identity were the main topics discussed in the third cluster. Analyzing cognitive (e.g., memory capacity and previous knowledge), learning strategies and motivational factors, the proposals from this cluster aimed to identify potential trajectories that could reveal students at risk. Moreover, this cluster addressed issues of intellectual property and digital literacy. There was no prevalent platform in this cluster, while courses were usually in fields such as English language, mathematics and physics.

Research theme 4: SNA and networked learning

A wide diversity in analysis methods and data sources is one of the defining characteristics of this cluster (Table 12). Applying networked learning and social network analysis tools and techniques, the proposals aimed to address various topics,

such as identifying central hubs in a course, or improving possibilities for students to gain employment skills. Moreover, learners' interaction profiles were analyzed in order to reveal different patterns of interactions between learners and instructors, among learners, and learners with content and/or underlying technology. Neither specific domain, nor platform was identified as dominant within the fourth cluster.

Research theme 5: Motivation, attitude and success criteria

The proposals within the fifth cluster aimed to analyze diverse motivational aspects and correlation between those motivational facets and course completion. Further, researchers analyzed various MOOC pedagogies (xMOOC, cMOOCs) and systems for supporting MOOCs (e.g., automated essay scoring), as well as attitudes of higher education institutions toward MOOCs. Another stream of research within this cluster was related to principles and best practices of transformation of traditional courses to MOOCs, as well as exploration of reasons for high dropout rates. The Coursera platform was most commonly referred to as a source for course delivery and data collection.

Phase 2 research methods.

Table 13 indicates that mixed methods was the most common methodological approach followed by purely quantitative research, which was used just slightly more than qualitative research. This suggests that there was no clear “winner” in terms of the adopted methodological approaches, and that all three types are used with a similar frequency. Also, the average number of authors and citations shows that the submissions mixed methods tended to have slightly more authors than quantitative or qualitative submissions, and that quantitative submissions had a significantly lower number of citations than submissions adopting both mixed and qualitative methods.

Table 10 shows that the submissions centered around engagement and peer assessment (i.e., cluster 1) used mainly quantitative research methods, while submissions dealing with self-regulated learning and social learning (i.e., cluster 3) exclusively used qualitative and mixed research methods. Finally, submissions centered around social network analysis (i.e., cluster 4) mostly used mixed methods, while submissions dealing with MOOC design and curriculum (i.e., cluster 2), and ones dealing with motivation, attitude and success criteria (i.e., cluster 5) had an equal adoption of all the three research methods.

Table 12

Phase 2 Distribution of Research Methodologies

| Methodology | Submissions | Authors avg. (SD) | Citations avg. (SD) |
|--------------|-------------|-------------------|---------------------|
| Mixed | 33 (42.3%) | 2.7 (1.5) | 21.8 (13.2) |
| Qualitative | 19 (24.4%) | 2.1 (0.9) | 22.8 (12.10) |
| Quantitative | 26 (33.3%) | 2.4 (1.2) | 16.7 (10.3) |
| Total | 78(100%) | 2.5 (1.3) | 20.3 (12.3) |

Phase 2 demographic characteristics of the authors.

With respect to the primary research areas of the submission authors, Table 14 shows that education was the primary research field of the large majority of the authors and that computer science was the distant second. In terms of the average number of authors, we can see on Table 10 that submissions related to MOOC design and curriculum (i.e., research theme 2) and motivation, attitude and success criteria (i.e., research theme 5) had on average a slightly higher number of authors than the other three research themes. In terms of their number of citations, submissions dealing with the engagement and peer assessment had on average 15 citations, while the submissions about other research themes had a bit higher number of citations ranging from 20 to 23. Similar to Phase 1, in all research themes, the field of education was found to be the main research background of submission authors. This was followed by the submissions authored by computer science and engineering researchers, and in the case of submissions about motivation, attitude and success criteria, by social scientists. Finally, similar to Phase 1, we see the strong presence of researchers from North America, followed by the much smaller number of researchers from other parts of the world (Table 15).

Table 13

Phase 2 Top 5 Research Fields

| Field | Authors |
|------------------|---------|
| Education | 106 |
| Computer Science | 21 |
| Engineering | 13 |
| Industry | 8 |
| Social Sciences | 6 |

Table 14

Phase 2 Geographic Distribution of the Authors

| Continent | Authors | Authored proposals | Accepted proposals |
|---------------|---------|--------------------|--------------------|
| Asia | 17 | 4.64 | 0.14 |
| Australia/NZ | 11 | 4.25 | 1 |
| Europe | 40 | 15.66 | 4 |
| North America | 137 | 52.44 | 22.85 |
| South America | 3 | 1 | 0 |

Phase 2 citation analysis.

We calculated a total number of citations (Table 16) for each publication, and extracted a list of the most cited authors (Figure 3). We can observe that the most cited authors were not necessarily the ones with the highest betweenness centrality, but the ones whose research focus was most relevant from the perspective of the MRI initiative and researchers from different fields and with different research objectives.

We also extracted the citation network graph which is shown on Figure 4. At the centre of the network is L. Pappano, the author of a very popular *New York Times* article “The Year of the MOOC”, as the author with the highest betweenness centrality value. The reason for this is that his article was frequently cited by a large number of researchers from a variety of academic disciplines, and thus making him essentially a bridge between them, which is clearly visible on the graph.

We also analyzed citation networks for each research theme independently and extracted common network properties such as diameter, average degree, path and density (Table 17).

Table 15

Phase 2 Most Cited Papers

| Paper name | Citation count |
|--|----------------|
| Kizilcec, R. F., Piech, C. and Schneider, E. (2013). Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. | 15 |
| Liyanagunawardena, T. R., Adams, A. A. and Williams, S. (2013). MOOCs: a Systematic Study of the Published Literature 2008-2012. | 13 |
| McAuley, A., Stewart, B., Siemens, G. and Cormier, D. (2010). The MOOC model for digital practice. | 13 |
| Breslow, L. B., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D. and Seaton, D. T. (2013). Studying learning in the worldwide classroom: Research into edX's first MOOC. | 13 |
| Siemens, G. (2005). Connectivism: A Learning Theory for the Digital Age. | 12 |
| Pappano, L. (2012). The Year of the MOOC. | 10 |
| Yuan L. and Powell S. (2013). MOOCs and Open Education: Implications for Higher Education. | 9 |
| Jordan, K. (2013). MOOC Completion Rates : The Data. | 7 |
| Belanger, Y. and Thornton, J. (2013). Bioelectricity: A Quantitative Approach. Duke University First MOOC. | 7 |
| Long, P. and Siemens, G. (2012). Penetrating the fog: analytics in learning and education. | 6 |
| Kop, R. (2011). The Challenges to Connectivist Learning on Open Online Networks: Learning Experiences during a Massive Open Online Course. | 6 |
| Daniel, J. (2012). Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility. | 6 |
| Mackness, J., Mak, S. F. J. and Williams, R. (2010). The Ideals and Reality of Participating in a MOOC. | 5 |
| Means, B., Toyama, Y., Murphy, R., Bakia, M. and Jones, K. (2010). Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies. | 5 |

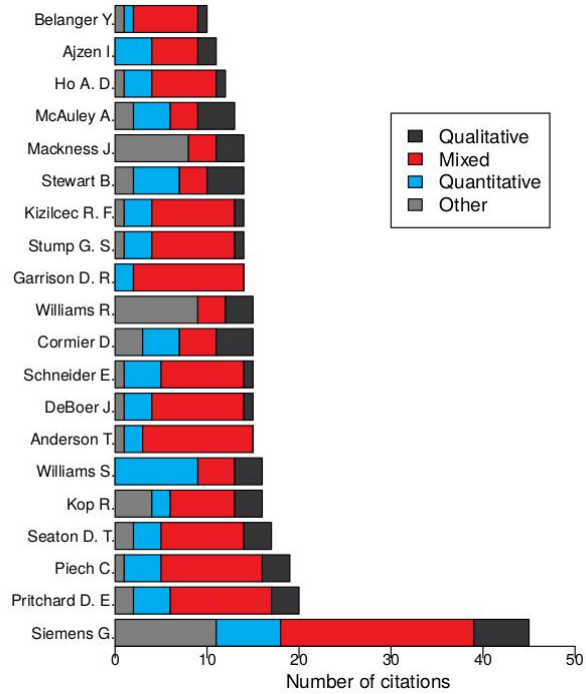


Figure 3. Phase 2 most cited authors

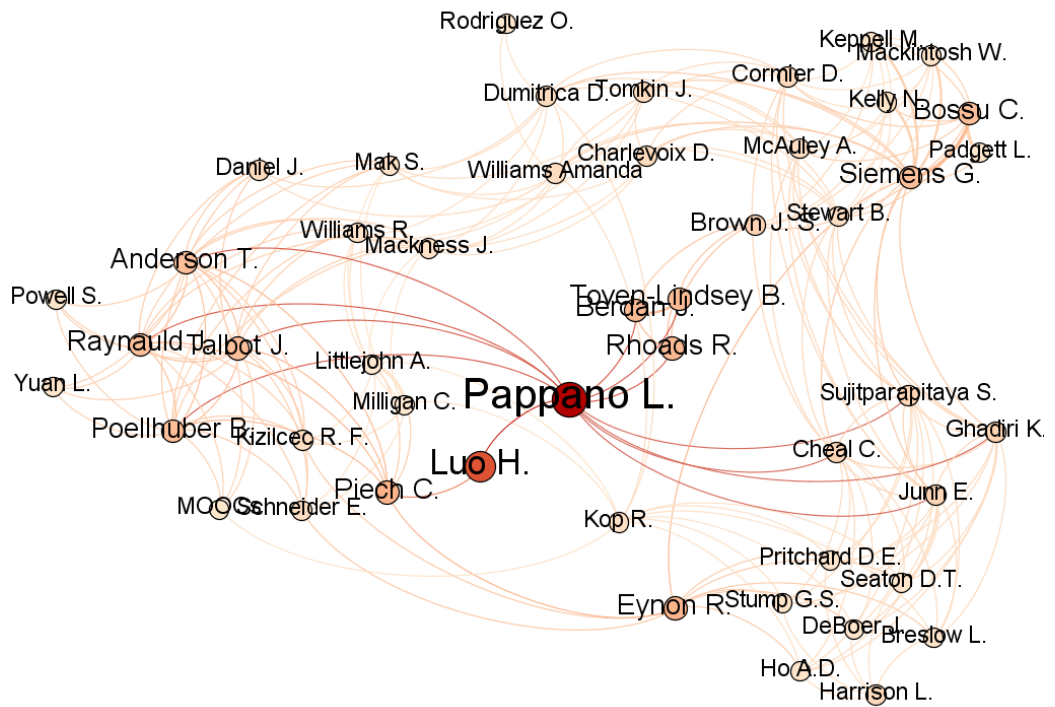


Figure 4. Phase 2 citation network.

Table 16

Phase 2 Citation Network Metrics

| Cluster | Average degree (SD) | Diameter | Average shortest path (SD) | Density |
|---|---------------------|----------|----------------------------|---------|
| Cluster 1 Engagement and Peer Assessment | 4.6 (8.4) | 8 | 4.5 (1.6) | 0.014 |
| Cluster 2 MOOC Design and Curriculum | 5.3 (10.9) | 9 | 4.3 (1.8) | 0.017 |
| Cluster 3 Learning Characteristics and Social Learning | 5.4 (8.7) | 7 | 4.1 (1.3) | 0.023 |
| Cluster 4 SNA and Networked Learning | 4.9 (9.6) | 8 | 3.9 (1.4) | 0.015 |
| Cluster 5 Motivation, Attitude and Success Criteria | 6.9 (9.0) | 8 | 3.7 (1.5) | 0.033 |
| Total | 5.1 (7.3) | 11 | 4.0 (1.3) | 0.012 |

Phase 2 success factors.

Similar to the analysis in Phase 1, we wanted to see whether there was any significant correlation between the citation network centrality measures (Table 17) and the final submission acceptance rates. However, unlike in Phase 1, Spearman's rho did not reveal any statistically significant correlation at the $\alpha=0.05$ significance level.

Discussion

Emerging Themes in MOOC Research

The results of the analysis indicated a significant attention of the researchers to the issues related to MOOCs that have received much public (media) attention. Specifically, the issue of low course completion and high degree of student attrition was often pronounced as the key challenge of MOOCs (Jordan, 2013; Koller, Ng, Do, & Chen, 2013). Not only was the topic of engagement and learning success (Cluster 1 in Phase 2) identified as a key theme in the MRI submissions, but it was also identified as a theme that was clearly cross-cutting all other research themes identified in Phase 2, including

motivation, attitudes and success criteria in Cluster 5, course design in Cluster 2, and learning strategies, social interaction, and interaction with learning resources in Cluster 3. With the aim to understand the factors affecting student engagement and success in MOOCs, the proposals had suggested a rich set of data collection methods – for example, surveys, physiological brain activity, knowledge tests, and demographic variables (see Table 12). The theory of planned behavior (TBP) (Ajzen, 1991) was found (see Cluster 1 in Table 11) as the main theoretical foundation for research of student engagement and learning success. While TBP is a well-known framework for studying behavioral change – in this case changing students intention to complete a MOOC and thus, increase their likelihood of course completion – it remains to be seen to what extent a student's intention can be changed if the student did not have an intention to complete a MOOC in the first place. What would be a reason that could motivate a student to change their intention in cases when she/he only enrolled into a MOOC to access information provided without intentions to take any formal assessments? In that sense, it seems necessary first to understand students' intentions for taking a MOOC, before trying to study the effects of interventions (e.g., motivational messages) on the students with different initial intentions.

The results also confirmed that social aspects of learning in MOOCs were the most successful theme in the MRI initiative (see Table 9). A total of 15 out of the 28 accepted proposals (Clusters 3 and 4) were related to different factors of social learning in MOOCs. Not only has it become evident recently that students require socialization in MOOCs through different forms of self-organization, such as local meet-ups (Coughlan, 2014)⁷ and that social factors contribute to attribution in MOOCs (Rosé et al., 2014), educational research is also very clear about numerous educational benefits of socialization. The Vygotskian approach to learning posits that higher levels of internalization can be achieved through social interaction most effectively (Vygotsky, 1980). These benefits have been shown to lead to deeper approaches to learning and consequently to higher learning outcome (Akyol & Garrison, 2011). Moreover, students' positions in social networks have been found in the existing literature to have a significant positive effect on many important learning outcomes such as creative potential (Dawson et al., 2011), sense of belonging (Dawson et al., 2011), and academic achievement (Gašević, Zouaq, & Janzen, 2013). Yet, the lack of social interaction can easily lead to the sense of social isolation which is well documented as one of the main barriers in distance and online education (Muilenburg & Berge, 2001; Rovai, 2002). Finally, Tinto's (1997) influential theory recognizes social and academic integration as the most important factors of student retention in higher education.

⁷ It is important to acknowledge that the importance of a "face-to-face contact with other students" was found in the Lou et al. meta-analysis (2006) of the literature – published in the period from 1985 to 2002 – about the effects of different aspects of distance and open education on academic success.

Research Methods in MOOC Research

The high use of mixed methods is a good indicator of sound research plans that recognized the magnitude of complexity of the issues related to MOOCs (Greene, Caracelli, & Graham, 1989). The common use of design-based research is likely a reflection of MOOC research goals aiming to address practical problems, and at the same time, attempting to build and/or inform theory (Design-Based Research Collective, 2003; Reeves, Herrington, & Oliver, 2005). This assumes that research is performed in purely naturalistic settings of MOOC offering (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003), always involves some intervention (Brown, 1992), and typically has several iterations (Anderson & Shattuck, 2012). According to Anderson and Shattuck (2012), there are two types of interventions – instructional and technological – commonly applied in online education research. Our results revealed that the focus of the proposals submitted to the MRI initiative was primarily on the instructional interventions. However, it is reasonable to demand from MOOC research to study the extent to which different technological affordances, instructional scaffolds and the combinations of the two can affect various aspects of online learning in MOOCs. This objective was set a long time ago in online learning research, led to the Great Media debate (Clark, 1994; Kozma, 1994), and the empirical evidence that supports either position (affordances vs. instruction) of the debate (Bernard et al., 2009; Lou, Bernard, & Abrami, 2006). Given the scale of MOOCs, a wide spectrum of learners' goals, differences in roles of learners, instructors and other stakeholders, and a broad scope of learning outcomes, research of the effects of affordances versus instruction requires much research attention and should produce numerous important practical and theoretical implications. For example, an important question is related to the effectiveness of the use of centralized learning platforms (commonly used in xMOOCs) to facilitate social interactions among students and formation of learning networks that promote effective flow of information (Thoms & Eryilmaz, 2014).

Our analysis revealed that the issue of the number of iterations in design-based research was not spelled out in the proposals of the MRI initiative (Anderson & Shattuck, 2012). It was probably unrealistic to expect to see proposals with more than one edition of a course offering given the timeline of the MRI initiative. This meant that the MRI proposals, which aimed to follow design-based research, were focused on the next iteration of existing courses. However, given the nature of MOOCs, which are not necessarily offered many times and in regular cycles, what is reasonable to expect from conventional design-based methods that require several iterations? Given the scale of the courses, can the same MOOC afford for testing out several interventions that can be offered to different subpopulations of the enrolled students in order to compensate for the lack of opportunity of several iterations? If so, what are the learning, organizational, and ethical consequences of such an approach and how and whether at all they can be mitigated effectively?

The data collection methods were another important feature of the proposal submissions to the MRI initiative. Our results revealed that most of the proposals planned to use conventional data sources and data collection methods such as grades, surveys on assessments, and interviews. Of course, it was commending to see many of those proposals being based on the well-established theories and methods. However, it was surprising to see a low number of proposals that had planned to make use of the techniques and methods of learning analytic and educational data mining (LA/EDM) (Baker & Yacef, 2009; Siemens & Gašević, 2012). With the use of LA/EDM approaches, the authors of the MRI proposals would be able to analyze trace data about learning activities, which are today commonly collected by MOOC platforms. The use of LA/EDM methods could offer some direct research benefits such as absence and/or reduction of self-selection and being some less unobtrusive, more dynamic, and more reflective of actual learning activities than conventional methods (e.g., surveys) can measure (Winne, 2006; Zhou & Winne, 2012).

Interestingly, the most successful themes (Clusters 3-4 in Phase 2) in the MRI initiative had a higher tendency to use the LA/EDM methods than other themes. Our results indicate that the MRI review panel expressed a strong preference towards the use of the LA/EDM methods. As Table 12 shows, the data types and analysis methods in Clusters 3-4 were also mixed by combining the use of trace data with conventional data sources and collection methods (surveys, interviews, and focus groups). This result provided a strong indicator of the direction in which research methods in the MOOC arena should be going. It will be important however to see the extent to which the use of LA/EDM can be used to advance understanding of learning and learning environments. For example, it is not clear whether an extensive activity in a MOOC platform is indicative of high motivation, straggling and confusion with the problem under study, or the use of poor study strategies (Clarebout, Elen, Collazo, Lust, & Jiang, 2013; Lust, Juarez Collazo, Elen, & Clarebout, 2012; Zhou & Winne, 2012). Therefore, we recommend a strong alignment of the LA/EDM methods with educational theory in order to obtain meaningful interpretation of the results that can be analyzed across different contexts and that can be translated to practice of learning and teaching.

Importance of Interdisciplinarity in MOOC Research

The analysis of the research background of the authors who submitted their proposals to the MRI initiative revealed an overwhelmingly low balance between different disciplines. Contrary to the common conceptions of the MOOC phenomena to be driven by computer scientists, our results showed that about 53% in Phase 1, 67% in Phase 2, about 67% of the finally accepted proposals were the authors from the discipline of education. It is not clear the reason for this domination of the authors from the education discipline. Could this be a sign of the networks to which the leaders of the MRI initiative were able to reach out? Or, is this a sign of fragmentation in the community? Although not conclusive, some signs of fragmentation could be traced. Preliminary and somewhat anecdotal results of the new ACM international conference

on learning at scale indicate that the conference was dominated by computer scientists⁸. It is not possible to have a definite answer if the fragmentation is actually happening or not based on only these two events. However, the observed trend is worrying. A fragmentation would be unfortunate for advancing understanding of a phenomenon such as MOOCs in particular and education and learning in general, which require strong interdisciplinary teams (Dawson et al., 2014). Just as an illustration of possible negative consequences of the lack of disciplinary balance could be the theme of educational technology innovation (Cluster 1 in Phase 1) in the MRI initiative. As results showed, this theme resulted in no proposal approved for funding. One could argue that the underrepresentation of computer scientists and engineers in the author base was a possible reason for the lack of technological argumentation. Could a similar argument be made for Learning @ Scale regarding learning science and educational research contribution remains to be carefully interrogated through a similar analysis of the Learning @ Scale conference's community and topics represented in the papers presented at and originally submitted to the conference.

The positive association observed between the success of individual themes of the MRI submissions and citation network structure (i.e., diameter and average network path) warrants research attention. This significance of this positive correlation indicates that the themes of the submitted proposals, which managed to reach out to broader and more diverse citation networks, were more likely to be selected for funding in the MRI initiative. Being able to access information in different social networks is already shown to be positively associated with achievement, creativity, and innovation (Burt et al., 2013). Moreover, the increased length of network diameter – as shown in this study – was found to boost spread of behavior (Centola, 2010). In the context of the results of this study, this could mean that the increased diameters of citation networks in successful MRI themes were assessed by the MRI review panel as more likely to spread educational technology innovation in MOOCs. If that is the case, it would be a sound indicator of quality assurance followed by the MRI peer-review process. On the other hand, for the authors of research proposals, this would mean that trying to cite broader networks of authors would increase their chances of success to receive research funding. However, future research in other different situations and domains is needed in order to be able to validate these claims.

Conclusions and Recommendations

Research needs to come up with theoretical underpinnings that will explain factors related to social aspects in MOOCs that have a completely new context and offer practical guidance of course design and instruction (e.g., Clusters 2, 4, and 5 in Phase 2). The scale of MOOCs does limit the extent to which existing frameworks for social

⁸ <http://learningatscale.acm.org>

learning proven in (online) education can be applied. For example, the community of inquiry (CoI) framework posits that social presence needs to be established and sustained in order for students to build trust that will allow them to comfortably engage into deeper levels of social knowledge construction and group-based problem solving (Garrison, Anderson, & Archer, 1999; Garrison, 2011). The scale of and (often) shorter duration of MOOCs than in traditional courses limits opportunities for establishing sense of trust between learners, which likely leads to much more utilitarian relationships. Furthermore, teaching presence – established through different scaffolding strategies either embedded into course design, direct instruction, or course facilitation – has been confirmed as an essential antecedent of effective cognitive processing in both communities of inquiry and computer-supported collaborative learning (CSCL) (Fischer, Kollar, Stegmann, & Wecker, 2013; Garrison, Cleveland-Innes, & Fung, 2010; Gašević, Adesope, Joksimović, & Kovanović, 2014). However, some of the pedagogical strategies proven in CoI and CSCL research – such as role assignment – may not fit to the MOOC context due to common assumptions that the collaboration and/or group inquiry will happen in small groups (6-10 students) or smaller class communities (30-40 students) (Anderson & Dron, 2011; De Wever, Keer, Schellens, & Valcke, 2010). When this is combined with different goals with which students enroll into MOOCs compared to those in conventional (online) courses, it becomes clear that novel theoretical and practical frameworks of understanding and organizing social learning in MOOCs are necessary. This research direction has been reflected in the topics identified in Cluster 4 of Phase 2 such as network formulation and peer-to-peer, online, learners and asynchronous interaction (Table 11). However, novel theoretical goals have not been so clearly voiced in the results of the analyses performed in this study.

The connection with learning theory has also been recognized as another important feature of the research proposals submitted to MRI (e.g., Clusters 3-5 in Phase 2). Likely responding to the criticism often attributed to the MOOC wave throughout 2012 not to be driven by rigorous research and theoretical underpinnings, the researchers submitting to the MRI initiative used frameworks well-established in educational research and the learning sciences. Of special interest were topics related to self-regulated learning (Winne & Hadwin, 1998; Zimmerman & Schunk, 2011; Zimmerman, 2000). Consideration of self-regulated learning in design of online education has been already recognized. To study effectively in online learning environments, learners need to be additionally motivated and have an enhanced level of metacognitive awareness, knowledge and skills (Abrami, Bernard, Bures, Borokhovski, & Tamim, 2011). Such learning conditions may not have the same level of structure and support as students have typically experienced in traditional learning environments. Therefore, understanding of student motivation, metacognitive skills, learning strategies, and attitudes is of paramount importance for research and practice of learning and teaching in MOOCs.

The new educational context of MOOCs triggered research for novel course and curriculum design principles as reflected in Cluster 2 of Phase 2. Through the increased attention to social learning, it becomes clear that MOOC design should incorporate factors of knowledge construction (especially in group activities), authentic learning, and personalized learning experience that is much closer to the connectivist principles underlying cMOOCs (Siemens, 2005), rather than knowledge transmission as commonly associated with xMOOCs (Smith & Eng, 2013). By triggering the growing recognition of online learning world-wide, MOOCs are also interrogated from the perspective of their place in higher education and how they can influence blended learning strategies of institutions in the post-secondary education sector (Porter, Graham, Spring, & Welch, 2014). Although the notion of flipped classrooms is being adopted by many in the higher education sector (Martin, 2012; Tucker, 2012), the role of MOOCs begs many questions such as those related to effective pedagogical and design principles, copyright, and quality assurance.

Finally, it is important to note that the majority of the authors of the proposals submitted to the MRI were from North America, followed by the authors from Europe, Asia, and Australia. This clearly indicates a strong population bias. However, this was expected given the time when the MRI initiative happened – proposals submitted in mid-2013. At that time, MOOCs were predominately offered by the North American institutions through the major MOOC providers to a much lesser extent in the rest of the world. Although the MOOC has become a global phenomenon and attracted much mainstream media attention – especially in some regions such as Australia, China and India as reported by Kovanovic et al. (2014) – it seems the first wave of research activities is dominated by researchers from North America. In the future studies, it would be important to investigate whether this trend still holds and to what extent other continents, cultures, and economies are represented in the MOOC research.

References

- Abrami, P. C., Bernard, R. M., Bures, E. M., Borokhovski, E., & Tamim, R. M. (2011). Interaction in distance education and online learning: Using evidence and theory to improve practice. *Journal of Computing in Higher Education*, 23(2-3), 82–103. doi:10.1007/s12528-011-9043-x
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. doi:10.1016/0749-5978(91)90020-T
- Akyol, Z., & Garrison, D. R. (2011). Understanding cognitive presence in an online and blended community of inquiry: Assessing outcomes and processes for deep approaches to learning. *British Journal of Educational Technology*, 42(2), 233–250. doi:10.1111/j.1467-8535.2009.01029.x
- Allen, I. E., & Seaman, J. (2013). *Changing course: Ten years of tracking online education in the United States*. Newburyport, MA: Sloan Consortium. Retrieved from <http://www.onlinelearningsurvey.com/reports/changingcourse.pdf>
- Anderson, T., & Dron, J. (2011). Three generations of distance education pedagogy. *The International Review of Research in Open and Distance Learning*, 12(3), 80–97.
- Anderson, T., & Shattuck, J. (2012). Design-based research a decade of progress in education research? *Educational Researcher*, 41(1), 16–25. doi:10.3102/0013189X11428813
- Baker, R. S. J. d., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1(1), 3–17.
- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. In *Proceedings of the Third International AAAI Conference on Weblogs and Social Media* (pp. 361–362).
- Belanger, Y., & Thornton, J. (2013). *Bioelectricity: A quantitative approach*. Duke University. Retrieved from <http://dukespace.lib.duke.edu/dspace/handle/10161/6216>
- Bernard, R. M., Abrami, P. C., Borokhovski, E., Wade, C. A., Tamim, R. M., Surkes, M. A., & Bethel, E. C. (2009). A meta-analysis of three types of interaction treatments in distance education. *Review of Educational Research*, 79(3), 1243–1289. doi:10.3102/0034654309333844
- Brent, E. E. (1984). The computer-assisted literature review. *Social Science Computer Review*, 2(3), 137–151.

- Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., & Seaton, D. T. (2013). Studying learning in the worldwide classroom: Research into edX's first MOOC. *Journal of Research & Practice in Assessment, 8*, 13–25.
- Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2*(2), 141–178. doi:10.1207/s15327809jls0202_2
- Burt, R. S., Kilduff, M., & Tasselli, S. (2013). Social network analysis: Foundations and frontiers on advantage. *Annual Review of Psychology, 64*(1), 527–547. doi:10.1146/annurev-psych-113011-143828
- Centola, D. (2010). The spread of behavior in an online social network experiment. *Science, 329*(5996), 1194–1197. doi:10.1126/science.1185231
- Chang, Y.-H., Chang, C.-Y., & Tseng, Y.-H. (2010). Trends of science education research: An automatic content analysis. *Journal of Science Education and Technology, 19*(4), 315–331. doi:10.1007/s10956-009-9202-2
- Cheng, B., Wang, M., Mørch, A. I., Chen, N.-S., Kinshuk, & Spector, J. M. (2014). Research on e-learning in the workplace 2000–2012: A bibliometric analysis of the literature. *Educational Research Review, 11*, 56–72. doi:10.1016/j.edurev.2014.01.001
- Clarebout, G., Elen, J., Collazo, N. A. J., Lust, G., & Jiang, L. (2013). Metacognition and the use of tools. in R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 187–195). Springer New York.
- Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development, 42*(2), 21–29. doi:10.1007/BF02299088
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13. doi:10.3102/0013189X032001009
- Coughlan, S. (2014, April 8). The irresistible urge for students to talk. Retrieved from <http://www.bbc.com/news/business-26925463>
- Cribb, A., & Gewirtz, S. (2013). The hollowed-out university? A critical analysis of changing institutional and academic norms in UK higher education. *Discourse: Studies in the Cultural Politics of Education, 34*(3), 338–350. doi:10.1080/01596306.2012.717188
- Dawson, S., Gašević, D., Siemens, G., & Joksimovic, S. (2014). Current state and future trends: A citation network analysis of the learning analytics field. In *Proceedings of the Fourth International Conference on Learning Analytics*

And Knowledge (pp. 231–240). New York, NY, USA: ACM.
doi:10.1145/2567574.2567585

Dawson, S., Tan, J. P. L., & McWilliam, E. (2011). Measuring creative potential: Using social network analysis to monitor a learners' creative capacity. *Australasian Journal of Educational Technology*, 27(6), 924–942.

De Wever, B., Keer, H. V., Schellens, T., & Valcke, M. (2010). Roles as a structuring tool in online discussion groups: The differential impact of different roles on social knowledge construction. *Computers in Human Behavior*, 26(4), 516–523.
doi:10.1016/j.chb.2009.08.008

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5–8.
doi:10.3102/0013189X032001005

Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist*, 48(1), 56–66. doi:10.1080/00461520.2012.748005

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215–239. doi:10.1016/0378-8733(78)90021-7

Garrison, D. R. (2011). *E-learning in the 21st century a framework for research and practice*. New York: Routledge.

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

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

Gašević, D., Adesope, O., Joksimović, S., & Kovanović, V. (2014). Externally-facilitated regulation scaffolding and role assignment to develop cognitive presence in asynchronous online discussions. *Submitted for Publication to The Internet and Higher Education*.

Gašević, D., Zouaq, A., & Janzen, R. (2013). Choose your classmates, your GPA is at stake! The association of cross-class social ties and academic performance. *American Behavioral Scientist*, 57(10), 1460–1479.
doi:10.1177/0002764213479362

- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*(3), 255–274. doi:10.3102/01623737011003255
- GSV Advisors. (2012). *Fall of the wall: Capital flows to education innovation*. Chicago, IL: Global Silicon Valley (GSV) Advisors. Retrieved from <http://goo.gl/NMtKZ4>
- Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Boston, MA: Addison-Wesley Pub. Co.
- Hoonlor, A., Szymanski, B. K., & Zaki, M. J. (2013). Trends in computer science research. *Commun. ACM, 56*(10), 74–83. doi:10.1145/2500892
- Jordan, K. (2013). MOOC completion rates: The data. Retrieved from <http://www.katyjordan.com/MOOCproject.html>
- Kinshuk, Huang, H.-W., Sampson, D. G., & Chen, N.-S. (2013). Trends in educational technology through the lens of the highly cited articles published in the Journal of Educational Technology And Society. *Educational Technology & Society, 3*–20.
- Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 170–179). New York, NY, USA: ACM. doi:10.1145/2460296.2460330
- Koller, D., Ng, A., Do, C., & Chen, Z. (2013). Retention and intention in massive open online courses: In depth. *Educause Review, 48*(3), 62–63.
- Kovanović, V., Joksimović, S., Gašević, D., Siemens, G., & Hatala, M. (2014). What public media reveals about MOOCs? *Submitted for Publication to British Journal of Educational Technology*.
- Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research and Development, 42*(2), 7–19. doi:10.1007/BF02299087
- Leckart, S. (2012, March 20). The Stanford education experiment could change higher learning forever. *Wired*. Retrieved from http://www.wired.com/2012/03/ff_aiclass/
- Li, F. (2010). Textual analysis of corporate disclosures: A survey of the literature. *Journal of Accounting Literature*.

- Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: A systematic study of the published literature 2008-2012. *The International Review of Research in Open and Distance Learning*, 14(3), 202–227.
- Lou, Y., Bernard, R. M., & Abrami, P. C. (2006). Media and pedagogy in undergraduate distance education: A theory-based meta-analysis of empirical literature. *Educational Technology Research and Development*, 54(2), 141–176. doi:10.1007/s11423-006-8252-x
- Lust, G., Juarez Collazo, N. A., Elen, J., & Clarebout, G. (2012). Content management systems: Enriched learning opportunities for all? *Computers in Human Behavior*, 28(3), 795–808. doi:10.1016/j.chb.2011.12.009
- Martin, F. G. (2012). Will massive open online courses change how we teach? *Commun. ACM*, 55(8), 26–28. doi:10.1145/2240236.2240246
- Maslen, G. (2014, May 20). Australia debates major changes to its higher-education system. *The Chronicle of Higher Education*. Retrieved from <http://chronicle.com/article/Australia-Debates-Major/146693/>
- McAuley, A., Stewart, B., Siemens, G., & Cormier, D. (2010). *The MOOC model for digital practice*. elearnspace.org. Retrieved from http://www.elearnspace.org/Articles/MOOC_Final.pdf
- Muilenburg, L., & Berge, Z. L. (2001). Barriers to distance education: A factor-analytic study. *American Journal of Distance Education*, 15(2), 7–22. doi:10.1080/08923640109527081
- OECD Publishing. (2013). *Education at a glance 2013: OECD indicators*. doi:10.1787/eag-2013-en
- Pappano, L. (2012, November 2). The year of the MOOC. *The New York Times*. Retrieved from <http://goo.gl/6QUBEK>
- Porter, W. W., Graham, C. R., Spring, K. A., & Welch, K. R. (2014). Blended learning in higher education: Institutional adoption and implementation. *Computers & Education*, 75, 185–195. doi:10.1016/j.compedu.2014.02.011
- Reeves, T. C., Herrington, J., & Oliver, R. (2005). Design research: A socially responsible approach to instructional technology research in higher education. *Journal of Computing in Higher Education*, 16(2), 96–115. doi:10.1007/BF02961476
- Rosé, C. P., Carlson, R., Yang, D., Wen, M., Resnick, L., Goldman, P., & Sherer, J. (2014). Social factors that contribute to attrition in MOOCs. In *Proceedings of*

the First ACM Conference on Learning @ Scale Conference (pp. 197–198). New York, NY, USA: ACM. doi:10.1145/2556325.2567879

- Rovai, A. P. (2002). Building sense of community at a distance. *The International Review of Research in Open and Distance Learning*, 3(1).
- Sari, N., Suharjito, & Widodo, A. (2012). Trend prediction for computer science research topics using extreme learning machine. *International Conference on Advances Science and Contemporary Engineering 2012*, 50(0), 871–881. doi:10.1016/j.proeng.2012.10.095
- Selwyn, N., & Bulfin, S. (2014). *The discursive construction of MOOCs as educational opportunity and educational threat*. MOOC Research Initiative (MRI) - Final Report. Retrieved from <http://goo.gl/ok43eT>
- Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, 2(1), 3–10.
- Siemens, G., & Gašević, D. (2012). Special Issue on learning and knowledge analytics. *Educational Technology & Society*, 15(3), 1–163.
- Smith, B., & Eng, M. (2013). MOOCs: A learning journey. In S. K. S. Cheung, J. Fong, W. Fong, F. L. Wang, & L. F. Kwok (Eds.), *Hybrid learning and continuing education* (pp. 244–255). Springer Berlin Heidelberg.
- Thoms, B., & Eryilmaz, E. (2014). How media choice affects learner interactions in distance learning classes. *Computers & Education*, 75, 112–126. doi:10.1016/j.compedu.2014.02.002
- Tight, M. (2008). Higher education research as tribe, territory and/or community: A co-citation analysis. *Higher Education*, 55(5), 593–605. doi:10.1007/s10734-007-9077-1
- Tinto, V. (1997). Classrooms as communities: Exploring the educational character of student persistence. *The Journal of Higher Education*, 68(6), 599–623. doi:10.2307/2959965
- Tucker, B. (2012). The flipped classroom. *Education Next*, 12(1), 82–83.
- Vygotsky, L. S. (1980). *Mind in society: Development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.
- Waltman, L., van Eck, N. J., & Wouters, P. (2013). Counting publications and citations: Is more always better? *Journal of Informetrics*, 7(3), 635–641. doi:10.1016/j.joi.2013.04.001

- Wasserman, S. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.
- Winne, P. H. (2006). How software technologies can improve research on learning and bolster school reform. *Educational Psychologist, 41*(1), 5–17.
doi:10.1207/s15326985ep4101_3
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Zhou, M., & Winne, P. H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction, 22*(6), 413–419.
doi:10.1016/j.learninstruc.2012.03.004
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego, CA, US: Academic Press.
- Zimmerman, B. J., & Schunk, D. H. (2011). *Handbook of self-regulation of learning and performance*. New York: Routledge.

© Gašević, Kovanović, Joksimović, Siemens

Athabasca University 



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).