Developing a MOOC experimentation platform: Insights from a user study

Vitomir Kovanović
School of Informatics
The University of Edinburgh
Edinburgh, UK
v.kovanovic@ed.ac.uk

Srećko Joksimović
Moray House School of Education
The University of Edinburgh
Edinburgh, UK
s.joksimovic@ed.ac.uk

Philip Katerinopoulos
School of Informatics
The University of Edinburgh
Edinburgh, UK
f.katerinopoulos@sms.ed.ac.uk

Charalampos Michail
School of Informatics
The University of Edinburgh
Edinburgh, UK
c.michail@sms.ed.ac.uk

George Siemens
LINK Research Lab
University of Texas at Arlington
Arlington, USA
gsiemens@uta.edu

Dragan Gašević
Moray House School of Education
and School of Informatics
The University of Edinburgh
Edinburgh, UK
dgasevic@acm.org

ABSTRACT

In 2011, the phenomenon of MOOCs had swept the world of education and put online education in the focus of the public discourse around the world. Although researchers were excited with the vast amounts of MOOC data being collected, the benefits of this data did not stand to the expectations due to several challenges. The analyses of MOOC data are very time-consuming and labor-intensive, and require and require a highly advanced set of technical skills, often not available to the education researchers. Because of this MOOC data analyses are rarely done before the courses end, limiting the potential of data to impact the student learning outcomes and experience.

In this paper we introduce MOOCito (MOOC intervention tool), a user-friendly software platform for the analysis of MOOC data, that focuses on conducting data-informed instructional interventions and course experimentations. We cover important design principles behind MOOCito and provide an overview of the trends in MOOC research leading to its development. Although a work-in-progress, in this paper, we outline the prototype of MOOCito and the results of a user evaluation study that focused on system's perceived usability and ease-of-use. The results of the study are discussed, as well as their practical implications.

CCS Concepts

•**Human-centered computing** → **User studies;** *User interface design;* •**Applied computing** → **Distance learning;** *E-learning; Education;* •**Information systems** → **Clustering;** *Data mining;*

Keywords

MOOCs, A/B testing, controlled experiments, analysis platform, user study, technology acceptance model, Coursera

1. INTRODUCTION

The introduction of Massive Open Online Courses (MOOCs) to the landscape of online learning was welcomed with great enthusiasm. With MOOC reaching the unprecedented number of students, they have been seen as a panacea for a broad range of issues, such as increasing access to higher education, student debt crisis, providing means for lifelong learning, and the overall democratization of learning [15]. In addition to the potential of MOOCs to solve a broad range of practical challenges, they also offer great opportunities for improving the understanding of learning processes [6, 9], given the vast amounts of data being collected and made available to the researchers [21]. Although significant quantities of data are being collected, there are several issues related to its use.

One of the main challenges of MOOC research is the format of the data delivered by major MOOC providers. Such data usually requires extensive pre-processing before it can be used for the analysis. Due to the lack of MOOC-specific data analysis tools, studies are typically conducted using traditional software packages (e.g., R, Weka, SPSS). Aside from technical challenges, MOOC data analyses are generally correlational in nature, which – due to inability to eliminate the effect of confounding factors – has a significant impact on the external validity of their results. As with the rest of the educational research, the findings from MOOC studies are also very localized to a particular context (e.g., particular course design, pedagogy, subject domain, student population), and data analysis procedures (e.g., construct operationalization and the analysis process) [13] making it hard to generalize across a variety of educational settings. Due to the high costs of MOOC data analysis, they are also typically conducted after courses are over, limiting the potential of MOOC instructors to alter their instructional approach based on the data generated during the course. Finally, at present, popular MOOC platforms provide very limited insights into student learning activities [2, 8], focusing primarily on demographic data, cumulative statistics related to course content, and student satisfaction with the course (i.e., upvotes and downvotes, star-ratings, and stories about student learning experiences)

This paper is the first step towards resolving some of the issues identified above. We present our work-in-progress on a novel MOOC analytics platform, that explicitly focuses on enabling instructors with little technical background to conduct MOOC analyses, to gain better insight into student learning. Also, instructors should be able to *act* and *experiment* based on the analysis results during the course, leading to the improved student learning experience and the better understanding of the learning processes. In this paper, we present a prototype of the proposed system and outline the design principles guiding its design. We also present a user evaluation study which sought to examine the perceived usefulness and ease-of-use by the target user group. Results and the implications of our findings are further discussed.

2. BACKGROUND WORK

With the growing popularity of MOOCs, it became apparent that there is a need for a common platform for the analysis of the MOOC data [25]. One of such efforts is MOOCdb project [25], which focuses on developing a standardized format for MOOC data so that analytics and visualizations projects (e.g., MOOCviz project[1]) built on top of it can be used for the analysis of data

[1]moocviz.csail.mit.edu

coming from different MOOC platforms. However, one significant challenge of MOOCdb is that it requires great technical competence, limiting its use for the majority of MOOC instructors.

Another major trend in MOOC research is the growing interest in controlled experiments and A/B testing, which has supported by several MOOC platforms and used by an increasing number of researchers [3, 8, 21]. Currently, experimental studies in MOOC domain primarily focus on examining the differences in instructional practices, such as serving random sub-populations of students different learning materials [16, 5, 27], instructional approach [12, 24], or platform interfaces [1] without the regard for their individual differences. As pointed out by Lamb et al. [16], this raises several issues related to the estimation of the actual effectiveness of the particular instructional measure, in large part by the significant student attrition. As shown by Winne [28], individual differences are also limiting the validity of findings across different settings. Instead of observing students as homogeneous groups and randomly assigning them to groups, there is a need to account for the specificity of different individuals need to be considered.

An additional challenge with the existing support for MOOC experimentation is that they do not enable instructors to alter their instructional approach during the course. For example, it is not possible to identify students who are inactive in the discussions and investigate the effectiveness of different intervention emails *on this particular group of students*. Due to the high costs of MOOC data analysis, the intervention approaches have to be planned in advance and assigned to students at random, without examination of their learning activities. Although some systems, such as MOOClet [27] or Bazaar [22], can be used to dynamically assign students to groups, they require substantial technical expertise and infrastructural planning to be successfully utilized in the course.

Given the need to adjust instruction by understanding student behavior during the course, the proposed system focuses on cluster analysis of students based on the different indicators of their engagement. Cluster analysis is commonly adopted method in MOOC research, with a significant number studies using it to examine the student behavior (e.g., [7, 19, 10, 14, 17, 4, 23]). For example, a study by Kizilcec et al. [10] used clustering to identify four groups of students in MOOCs based on their course engagement (i.e., completing students, auditing students, disengaged students, and sampling students). In a similar manner, MOOC instructors should be able to identify different subgroups of students and then conduct various instructional interventions on those subgroups. Through identification of groups of students based on individual differences in their behavior, we can examine the effects these differences have on the success of different intervention strategies. From the practical perspective, identification of student sub-populations in the real time enables the provision of personalized and focused interventions to students in various subgroups which can improve their course success and learning experience.

3. METHOD

3.1 Design principles

As the first step in our process, we focused on developing a prototype of an envisioned analytics system. The process was guided by the following set of design criteria which were based on the important issues identified above.

- **No technical prerequisites**. Given that many of MOOC instructors and researchers do not have advanced IT training, the system should not require substantial technical knowledge from its target users.
- **Incremental data import**. As MOOC platforms do not typically provide data in real-time but rather in batches (e.g., daily or weekly), the system should support incremental import of the data when the new batch becomes available, as this is necessary to enable data analysis during the course.
- **Support cluster analysis**. Given the need to understand different patterns of student engagement, the system should support clustering of students based on their interaction with the course materials and other learners.
- **Support class interventions**. With the need for the more proactive use of the MOOC data, the system should support quick analysis during the course. Hence, the data preprocessing and the extraction of important indicators of student engagement should be automated as much as possible.
- **Enable follow-up data analysis**. After an instructor implements an intervention, the system should enable the analysis of the effects of that intervention. As such, the system should store information about conducted interventions in a format which is suitable for follow-up analysis by the popular statistical packages (e.g., SPSS, R, Matlab).
- **Support study replication**. With the goal of the platform to advance the state of MOOC research, the system should support "pre-configured" analyses based on the previously published research. Not only would this limit the need for a manual analysis by the researchers, but it would also help validate previous research findings and examine their generalizability. By supporting "analysis templates", it would be possible to examine the replicability of the published research and the effect of study context on its findings.

Based on the defined design principles, we developed a prototype of the user interface following an iterative design cycle. Using paper-based prototypes and Axure RP platform[2], we incrementally developed a final version of the system which was then used in our user study. At present, the focus of the platform development is on supporting cluster analysis and the Coursera platform, with the plan on supporting additional platforms and other types of analysis (e.g., social network analysis).

3.2 Prototype overview

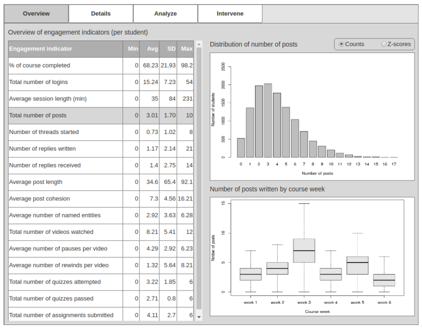
The final version of the developed prototype is shown in Fig. 1. The system consists of four main application tabs corresponding to the main steps of the analysis process.

The Overview tab (Fig. 1a) provides details of all engagement indicators which are available in the system. The particular list of indicators is defined following the review of MOOC literature on predicting student learning and persistence [9]. For each indicator, minimum, maximum, average and standard deviation are provided, alongside its score distribution and weekly mean values. Due to space limitations, we will not be going in depth over the list of extracted indicators, as they are not the focal point of this report.

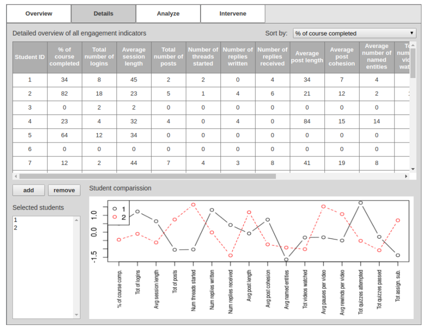
The details tab (Fig. 1b) enables instructors to examine and compare engagement indicators between specific students. The upper part provides a table with the values of all indicators for all students so that values of engagement indicators for each student can be seen. Given that the table is sortable, it is easy to identify students with extreme values for all engagement indicators. The lower portion of the window enables instructors to compare values of all engagement indicators for a selected list of students (e.g., compare students with messages posted and most submitted assignments).

The analysis tab represents a central component of the system which can be used to identify different clusters of students. The first step (Fig. 1c) is a selection (and limited pre-processing) of engagement indicators which are used for clustering (i.e., feature selection in data mining terminology), together with a predefined indicator configurations based on the published literature [18, 11]. Next, users can conduct a cluster analysis (Fig. 1d) using some of the popular algorithms (i.e., k-means, hierarchical clustering, em-clustering) which can be configured with very basic parameters

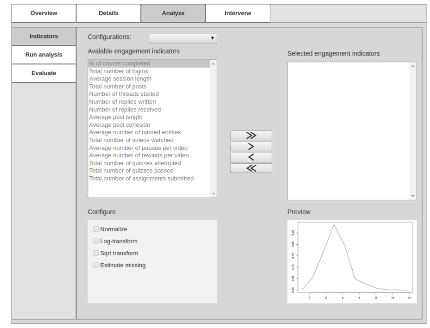
[2] www.axure.com



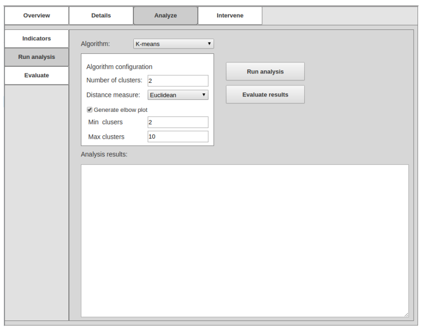
(a) Overview tab



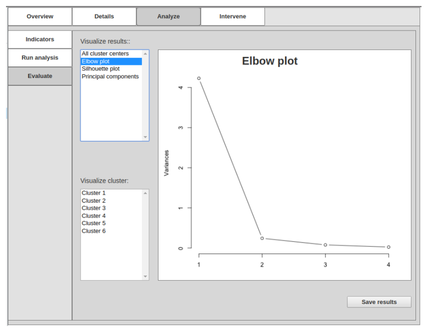
(b) Details tab



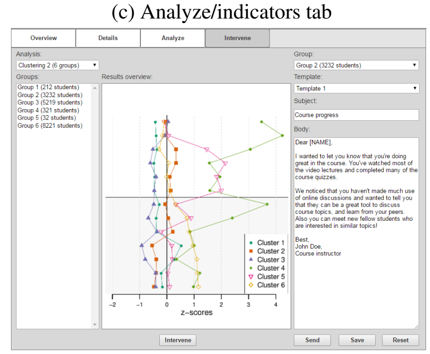
(c) Analyze/indicators tab



(d) Analyze/Run analysis tab



(e) Analyze/evaluate tab



(f) Intervene tab

Figure 1: MOOC Analytics prototype

(e.g., distance measure and K, the number of clusters). From the analysis results, several of visualizations are produced (Fig. 1e): visualization of cluster centroids, silhouette evaluation plot, and projections of clusters in two principal component space. For K-means and hierarchical clustering, it a range of values for K is selected, an elbow evaluation plot is also produced.

Finally, after analysis (or several analyses) are performed, instructors can send intervention messages to a desired group of students (Fig. 1f). Those can be whole clusters of students or some parts of the identified student groups. For instance, it is possible to send one message to 50% of cluster one, and another message to remaining 50% of that cluster (or no message at all, for a "control group" case). Instructional messages can be saved as templates and contain several "variables" in the email body, as commonly done in mail-merge applications. Finally, the instructor can save a CSV file with the intervention summary which can be later used for comparing the effectiveness of different interventions.

3.3 Evaluation study design

The study was conducted with eleven experienced MOOC instructors/researchers/designers from the three large research-intensive universities from the UK, USA, and South America. All participants had several years of experience with online learning, MOOCs and Coursera platform (Table 1). The focus of the examination was on obtaining rich qualitative insights about the developed prototype to evaluate the impressions of its targeted end-user population.

After signing the informed consent form, all participants were shown a fifteen-minute presentation of the study scenario which focused on a hypothetical MOOC instructor halfway through an introductory programming course delivered on the Coursera platform. The first part of the presentation provided an overview of the current Coursera dashboard, with the actual data from the "Code Yourself" MOOC offered by the University of Edinburgh, followed then by the detailed description and overview of the functionalities of the proposed MOOCito system. Finally, participants were presented a Web-based survey[3] with twenty five-item Likert-scale (from 1: strongly disagree to 5: strongly agree) and nine open-ended questions on the perceived usability and ease-of-use of the proposed system, and general impressions of the proposed system. The design of the study instrument was based on the technology acceptance model (TAM) [26] which is a psychometric tool commonly used for assessing the suitability of software systems. As shown in Fig 2, TAM can be used to evaluate the determinants of the perceived usefulness (i.e., perceived ease-of-use, subjective norm, image, job relevance, output quality, and result demonstrability) and perceived ease-of-use (i.e., computer self-efficacy, perception of external control, computer anxiety, computer playfulness, perceived enjoyment, and objective usability) which influence the adoption of a new software system. As our participants were presented the system prototype, we focused on the qualitative investigation of the system's perceived usefulness and ease-of-use and used only the relevant subset of dimensions defined by TAM.

Table 1: Study participant experience.

Question	Mean	SD	Mdn	IQR
Years of experience with online learning	9.10	4.53	9.50	7.75
Years of experience with MOOCs	3.38	1.12	4	0.75
Years of experience with Coursera	3.28	1.11	4	1.00

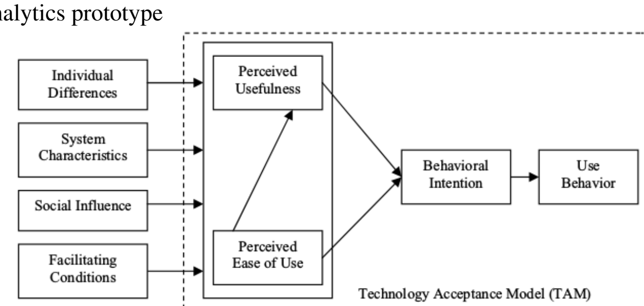


Figure 2: Technology Acceptance Model

4. RESULTS AND DISCUSSION

The summary of the Likert scale questions is given in Table 2. Overall, this indicates that participants were satisfied with the proposed system and its perceived usability and ease-of-use. The most positive responses were related to the ability to generate interesting questions and hypotheses (4.73), the ability of the tool to provide insights into students' engagement with the learning environment

[3] goo.gl/EjLc8o

Table 2: Five-item Likert-scale user study responses.

Q#	Question	M	SD	Mdn	IQR
Dashboard: *Perceived usefulness*					
D1	The tool enables me to get an insight into the students' engagement within the learning environment.	4.64	0.67	5.00	0.50
D2	The information the tool provides helps me identify students that might need assistance.	4.18	0.75	4.00	1.00
D3	The tool helps me to generate interesting questions related to course design/students worth exploring in more detail.	4.73	0.47	5.00	0.50
Dashboard: *Perceived ease-of-use*					
D4	Conducting analyses with the tool is easy and intuitive.	3.55	0.93	4.00	1.00
D5	Graphical user interface of the tool prototype is intuitive enough.	4.18	0.75	4.00	0.59
D6	Graphical user interface of the tool prototype is overburdened with information.	2.18	1.08	2.00	1.09
D7	Information about indicators of student engagement are adequately presented.	3.82	0.98	4.00	1.50
Intervention: *Perceived usefulness*					
I1	The tool enables me to intervene on the student learning during the course.	4.18	0.87	4.00	1.50
I2	It is important to be able to provide different interventions/instructions/guidances to different groups.	4.73	0.47	5.00	0.50
I3	It is important to be able to provide several different interventions/instructions/guidances to the same student group.	4.55	0.82	5.00	0.50
Intervention: *Perceived ease-of-use*					
I4	Selecting students for intervention is intuitive.	3.82	1.08	4.00	0.50
I5	Information on the Intervene tab is adequately laid out.	4.09	0.83	4.00	1.04
I6	Specifying different intervention/instruction/guidance messages for different groups of students is intuitive.	3.82	1.08	4.00	0.50
General opinion					
G1	I would like to be able to use the tool in my courses.	4.36	0.81	5.00	1.00
G2	I intend to use the tool in my future courses.	4.00	0.89	4.00	1.50
G3	I am willing to use the tool in my future course.	4.36	0.81	5.00	1.00
G4	I would recommend using the tool to my colleagues.	4.40	0.52	4.00	0.85
G5	I would use the tool frequently.	3.91	1.04	4.00	1.50
G6	I would use the tool for research purposes.	4.55	0.82	5.00	0.50
G7	I would use the tool for improving course design & teaching, and improving student learning.	4.30	0.67	4.00	1.00

(4.64), and the capacity to provide different interventions to subgroups of students (4.55). A similar sentiment was also reflected in the open-ended responses where participants emphasized the selection of engagement indicators, weekly plots, and the clear presentation of the interface, despite the vast amounts of data being presented. Regarding the ability to conduct instructional interventions, participants highlighted the flexibility and ease of performing meaningful interventions without "the dangers of providing blanket information inappropriately" to a particular target subset of students, and the use of templates and "variables" in the message body.

Although the overall impressions were positive, participants indicated several important issues which should be resolved before the system is fully developed. Despite our best efforts, there were several concerns regarding the accessibility of the tool to the general population of MOOC instructors. First of all, the participants stressed that instructors, users of MOOCito, might not be familiar with some of the adopted terminology, extracted engagement indicators, statistical analyses, or studies used as "study templates". Thus, additional information should be provided, for example in the form of short summaries and video tutorials. Also, full bibliographical information and summaries of the reference studies should be provided. Some participants indicated that they would not know what would be the best way to analyze the data, and in this regard, additional information on the "standard approaches" should be provided. The need for simplification of the analysis steps is also seen in responses to Question D4, which measured how much the tool was intuitive and how easy was to conduct the analysis. As such, one of the future directions is related to enabling more streamlined analysis, with better guidelines and support for instructors with little background and experience in statistics.

Another set of concerns relates to the ability to provide meaningful insights and instructional interventions. First of all, the participants indicated that instructors should be able to go from indicators of engagement to the actual student content. For instance, if a student's message cohesion is low, instructors should be able to easily see messages that have low cohesion and through their examination get a better understanding of the student's engagement. Moreover, some participants indicated the challenge of knowing *when or how to intervene*, highlighting the need for more "interpretable" results, summarized in a form which is easy to act upon. For example, besides the detailed description of cluster centers, a simpler description using "low", "moderate", and "high" for describing engagement indicators could be employed (e.g., cluster one characterizes low cohesion, high volume of message postings, and moderate viewing of online lecture videos). In this regard, some of the data-to-text techniques [20] could be used to provide a one-paragraph description of each cluster which are easier to interpret by the MOOC instructors.

Overall, the participants indicated a strong willingness (4.36) to use MOOCito and would like to be able to use it in their courses (4.00). Although the participants reported the eagerness to use it for improving both teaching and research, the willingness to use it for research was higher (4.55) than for course design and teaching purposes (4.30). As nicely summarized by one study participant *"This has the potential to be a powerful tool in educational research as well as a means of tracking learner engagement day-to-day."*. This and similar comments indicate the real need for an MOOC analysis platform and we hope that MOOCito can fill this gap which hinders the advancement in the MOOC teaching and research.

4.1 Limitations and future work

There are several limitations of our current approach. As the focus of our efforts is on supporting MOOC instructors on the Coursera platform, the results of our evaluation study might be significantly affected by its current analytics capabilities. Hence, despite Coursera being the MOOC platform with the highest user base, it might not be indicative of the MOOC domain as a whole. As such, one important area of future work is to support additional MOOC platforms. This is especially important given the Coursera's recent shift towards corporate training[4], which might impact its future adoption in the higher education sector. Secondly, although we provided a detailed prototype of the system graphical interface, it is still work-in-progress, and with its development underway, the final usability of MOOCito is yet to be seen when it is finally used in the real-world. Finally, in our study we had eleven participants in total which – although common in prototype testing – might be a small number to reliably measure the usability of the proposed system. Still, their qualitative feedback offers much helpful guidance that can inform future development of MOOCito.

[4] video.cnbc.com/gallery/?video=3000547421

Based on the results of our user study, we identified several important directions for our future work on the MOOCito platform, besides the development work. There is a need for more guidance and support in using the system (through detailed descriptions, summaries, and tutorial materials) which will be included in the next system prototype. Also, there is a need for better connection to the MOOC platform itself (e.g., study materials), and ability to go from engagement indicators to the associated learning content (e.g., discussion messages, video lectures, quizzes). Finally and most importantly, a more interpretable and actionable cluster descriptions must be provided, and this a critical advancement which will be the center of the future work on the MOOCito platform.

5. CONCLUSIONS

While it is true that MOOCs generated vast amounts of data about student learning, the lack of analytics support for MOOC instructors and researchers significantly reduced the potential of this data to improve student learning. Not only this, but the complexities of analyzing MOOC data leading to its disconnect to the teaching practice also affects our ability to use the same data to understand the learning phenomena better.

In this paper, we introduced MOOCito, a novel MOOC analytics platform which focuses on enabling MOOC instructors to gain insights about student learning and perform instructional interventions based on the collected student data. With the goal of helping "ordinary" MOOC instructors with little or no advanced technical knowledge, the system focuses on conducting in-course analyses and interventions based on those analyses so that students can be given proper instructional support. Also, the tool provides the ability to experiment with the instructional approach, with the goal of providing more reliable evidence on the success of different instructional interventions. Although experimentation in MOOCs is already happening [21], the lack of dedicated analytics support and pre-experimentation analysis results in experiments that have to be pre-planned and have significant methodological challenges [16].

Although still a work-in-progress, in this paper we present results of the user study which investigated perceived usefulness and ease-of-use of the proposed MOOC platform. The study findings confirmed our intuition on the necessity of providing analytical support for MOOC instructors beyond what is currently available. As one participant concluded: *"I believe this is an excellent tool that will significantly improve teaching and learning within the MOOCs. It will also expand the way the MOOCs data are used and enable the instructors to conduct research and publish data"*.

While the results are unequivocally positive, we identified several alleys for future work. The most important are related to i) better support and help materials in conducting analyses and interventions, ii) deeper examination of engagement indicator scores by looking at the source data, and iii) improved interpretability of analysis results through the simplification of cluster descriptions (e.g., using text summarization techniques). By focusing on issues identified by the actual MOOC instructors, we hope to provide analytics toolkit which will significantly advance the state of MOOC teaching and research and enhance student learning experience.

REFERENCES

[1] A. Anderson, D. Huttenlocher, J. Kleinberg, and J. Leskovec. Engaging with Massive Online Courses. In *Proceedings of the 23rd International Conference on World Wide Web*, pages 687–698. ACM, 2014.

[2] C. Coffrin, L. Corrin, P. d. Barba, and G. Kennedy. Visualizing Patterns of Student Engagement and Performance in MOOCs. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, pages 83–92. ACM, 2014.

[3] P. Diver and I. Martinez. MOOCs as a massive research laboratory: opportunities and challenges. *Distance Education*, 36(1):5–25, 2015.

[4] R. Ferguson and D. Clow. Examining Engagement: Analysing Learner Subpopulations in Massive Open Online Courses (MOOCs). In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, pages 51–58. ACM, 2015.

[5] W. W. Fisher. HLS1x: CopyrightX course report, 2014.

[6] D. Gašević, V. Kovanović, S. Joksimović, and G. Siemens. Where is research on massive open online courses headed? A data analysis of the MOOC Research Initiative. *The International Review of Research in Open and Distributed Learning*, 15(5), 2014.

[7] T. Hecking, S. Ziebarth, and H. U. Hoppe. Analysis of Dynamic Resource Access Patterns in Online Courses. *Journal of Learning Analytics*, 1(3):34–60, 2014.

[8] F. M. Hollands and D. Tirthali. MOOCs: Expectations and Reality. Full Report, Center for Benefit-Cost Studies of Education, Teachers College, Columbia University, 2014.

[9] S. Joksimović, O. Skrypnyk, V. Kovanović, N. Dowell, C. Mills, D. Gašević, S. Dawson, A. C. Graesser, and C. Brooks. How do we Model Learning at Scale? A Systematic Review of the Literature. submitted.

[10] R. F. Kizilcec, C. Piech, and E. Schneider. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pages 170–179. ACM, 2013.

[11] R. F. Kizilcec, C. Piech, and E. Schneider. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pages 170–179. ACM, 2013.

[12] R. Kizilcec, E Schneider, G. Cohen, and D. McFarland. Encouraging Forum Participation in Online Courses with Collectivist. *eLearning Papers*, 37:13–22, 2014.

[13] V. Kovanović, D. Gašević, S. Dawson, S. Joksimović, and R. Baker. Does time-on-task estimation matter? Implications on validity of learning analytics findings. *Journal of Learning Analytics*, 2(3):81–110, 2016.

[14] V. Kovanović, S. Joksimović, D. Gašević, J. Owers, A.-M. Scott, and A. Woodgate. Profiling MOOC Course Returners: How Does Student Behavior Change Between Two Course Enrollments? In *Proceedings of the Third (2016) ACM Conference on Learning @ Scale*, pages 269–272. ACM, 2016.

[15] V. Kovanović, S. Joksimović, D. Gašević, G. Siemens, and M. Hatala. What public media reveals about MOOCs: A systematic analysis of news reports. *British Journal of Educational Technology*, 46(3):510–527, 2015.

[16] A. Lamb, J. Smilack, A. Ho, and J. Reich. Addressing Common Analytic Challenges to Randomized Experiments in MOOCs: Attrition and Zero-Inflation. In *Proceedings of the Second (2015) ACM Conference on Learning @ Scale*, pages 21–30. ACM, 2015.

[17] N. Li, L. Kidzinski, P. Jermann, and P. Dillenbourg. MOOC Video Interaction Patterns: What Do They Tell Us? In *Proceedings of the 10th European Conference on Technology Enhanced Learning*, pages 197–210. Springer International Publishing, 2015.

[18] N. Li, Ł. Kidziński, P. Jermann, and P. Dillenbourg. Mooc video interaction patterns: what do they tell us? In *Design for Teaching and Learning in a Networked World*, pages 197–210. Springer, 2015.

[19] S. F. J. Mak, R. Williams, and J. Mackness. Blogs and Forums as Communication and Learning Tools in a MOOC. In pages 275–284, 2010.

[20] A Ramos-Soto, B Vázquez-Barreiros, A Bugarín, A Gewerc, and S Barro. Evaluation of a data-to-text system for verbalizing a learning analytics dashboard. *International Journal of Intelligent Systems*, 2016.

[21] J. Reich. Rebooting MOOC Research. *Science*, 347(6217):34–35, 2015.

[22] C. P. Rosé, O. Ferschke, G. Tomar, D. Yang, I. Howley, V. Aleven, G. Siemens, M. Crosslin, D. Gasevic, and R. Baker. Challenges and opportunities of dual-layer moocs: reflections from an edx deployment study. In *Proceedings of the 11th International Conference on Computer Supported Collaborative Learning (CSCL 2015)*, volume 2, 2015.

[23] T. Sinha. "Your click decides your fate": Leveraging clickstream patterns from MOOC videos to infer students' information processing & attrition behavior. *arXiv:1407.7143 [cs]*, 2014. arXiv: 1407.7143.

[24] J. H. Tomkin and D. Charlevoix. Do Professors Matter?: Using an a/B Test to Evaluate the Impact of Instructor Involvement on MOOC Student Outcomes. In *Proceedings of the First ACM Conference on Learning @ Scale Conference*, pages 71–78. ACM, 2014.

[25] K. Veeramachaneni, S. Halawa, F. Dernoncourt, U.-M. O'Reilly, C. Taylor, and C. Do. MOOCdb: Developing Standards and Systems to Support MOOC Data Science. *arXiv:1406.2015 [cs]*, 2014.

[26] V. Venkatesh and H. Bala. Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39(2):273–315, 2008.

[27] J. J. Williams, J. Kim, and B. Keegan. Supporting Instructors in Collaborating with Researchers Using MOOClets. In *Proceedings of the Second (2015) ACM Conference on Learning @ Scale*, pages 413–416. ACM, 2015.

[28] P. H. Winne. Minimizing the black box problem to enhance the validity of theories about instructional effects. *Instructional Science*, 11(1):13–28, 1982.