International Review of Research in Open and Distributed Learning

Meaningful Learner Information for MOOC Instructors Examined Through a Contextualized Evaluation Framework

Kerrie A Douglas, Mitchell W. Zielinski, Hillary Merzdorf, Heidi A Diefes-Dux et Peter Bermel

Citer cet article

Douglas, K., Zielinski, M., Merzdorf, H., Diefes-Dux, H. & Bermel, P. (2019). Meaningful Learner Information for MOOC Instructors Examined Through a Contextualized Evaluation Framework. *International Review of Research in Open and Distributed Learning*, *20*(1). https://doi.org/10.19173/irrodl.v20i1.3717

Résumé de l'article

Improving STEM MOOC evaluation requires an understanding of the current state of STEM MOOC evaluation, as perceived by all stakeholders. To this end, we investigated what kinds of information STEM MOOC instructors currently use to evaluate their courses and what kinds of information they feel would be valuable for that purpose. We conducted semi-structured interviews with 14 faculty members from a variety of fields and research institutions who had taught STEM MOOCs on edX, Coursera, or Udacity. Four major themes emerged related to instructors' desires: (1) to informally assess learners as an instructor might in a traditional classroom, (2) to assess learners' attainment of personal learning goals, (3) to obtain in-depth qualitative feedback from learners, and (4) to access more detailed learner analytics regarding the use of course materials. These four themes contribute to a broader sentiment expressed by the instructors that they have access to a wide variety of quantitative data for use in evaluation, but are largely missing the qualitative information that plays a significant role in traditional evaluation. Finally, we provide our recommendations for MOOC evaluation criteria, based on these findings.

Meaningful Learner Information for MOOC Instructors Examined Through a Contextualized Evaluation Framework

Kerrie Douglas, Mitch Zielinski, Hillary Merzdorf, Heidi Diefes-Dux, and Peter Bermel
Purdue University, West Lafayette, Indiana, USA

Abstract

Improving STEM MOOC evaluation requires an understanding of the current state of STEM MOOC evaluation, as perceived by all stakeholders. To this end, we investigated what kinds of information STEM MOOC instructors currently use to evaluate their courses and what kinds of information they feel would be valuable for that purpose. We conducted semi-structured interviews with 14 faculty members from a variety of fields and research institutions who had taught STEM MOOCs on edX, Coursera, or Udacity. Four major themes emerged related to instructors' desires: (1) to informally assess learners as an instructor might in a traditional classroom, (2) to assess learners' attainment of personal learning goals, (3) to obtain in-depth qualitative feedback from learners, and (4) to access more detailed learner analytics regarding the use of course materials. These four themes contribute to a broader sentiment expressed by the instructors that they have access to a wide variety of quantitative data for use in evaluation, but are largely missing the qualitative information that plays a significant role in traditional evaluation. Finally, we provide our recommendations for MOOC evaluation criteria, based on these findings.

Keywords: MOOCs, evaluation, MOOC instructors

Introduction

Massive open online courses, MOOCs, have been able to capture the investment of higher education institutions and have been accessed by millions of users worldwide. By 2016, 6,850 courses from over 700 universities had been offered as MOOCs, reaching an estimated audience of 58 million learners in 2016 alone (Hollands & Tirthali, 2014; Shah, 2016). Considering the $39,000 to $325,000 price tag for any given MOOC, these numbers reveal a significant financial investment (Hollands & Tirthali, 2014). Yet, despite these significant investments, very little evidence has been given to justify the cost expenditure or demonstrate the quality of the learning opportunities provided. Evaluation of MOOCs has been a somewhat controversial topic, as there has been much discussion concerning the inapplicability of traditional educational metrics to a MOOC environment, and a general acceptance of the low completion rates. There has been less conversation about what metrics would actually provide information on the quality of learning that could be used to further improve the pedagogical strategies employed in MOOCs. Consequently, the most commonly reported outcomes of MOOCs still primarily rely on high enrollment numbers and access to materials, rather than information that could assist one in coming to a conclusion on the quality of the learning opportunity in a particular MOOC. Criticizing evaluation metrics without providing justifiable alternatives risks preventing authentic evaluation that could lead to informed decision-making and improved courses and learner experiences. Speaking of open education resources broadly, UNESCO's 2015 *Education 2030* report states, "Access is not enough; we need a new focus on the quality of education and the relevance of learning and on what children, youth and adults are actually learning" (UNESCO, 2015, p. 4).

For MOOC platforms and institutions to justify cost expenditure and instructors to identify areas for pedagogical improvement, a comprehensive model of evaluation is needed which addresses the unique challenges of operating in an open educational environment, where learners are vastly heterogeneous and free to come and go as desired. Institutions are seeking to determine the most effective MOOC platforms; making rational choices requires establishing and applying appropriate evaluation criteria. Likewise, institutional staff and others tasked with providing evaluative information regarding institutional investments in MOOCs must establish evaluation criteria in determining the institution's merit of investment.

Although the word evaluation is used in everyday language, professional evaluation refers to a systematic determination of the merit or quality of something (Scriven, 1991). In order to determine merit or quality, one must first understand what is meaningful to stakeholders in a particular context. Therefore, principled approaches to evaluation begin with an assessment of stakeholder values (Scriven, 1983).

The *Contextualized Evaluation Framework* (Douglas et al., 2017) is based on the understanding that evaluation questions as to the overall worth of MOOCs (and any individual MOOC), can only be addressed by answering questions concerning the background and context of MOOCs, stakeholder values (specifically in terms of the basis for claims of quality or merit), MOOC learner characteristics and values, and the resources available to create MOOCs. A thorough understanding of context, stakeholder and learner values, and resources, can then be used to interpret course characteristics and learners' interactions (behavior and outcomes) within a course. This Contextualized Evaluation Framework is based upon the work of Scriven's (2015) Key Evaluation Checklist and Davidson's (2005) Genuine Evaluation. According to Scriven,

evaluation is more than simply providing data or results; it is the science of valuing, specifying what is valued, and how a judgment regarding quality will be made (Shadish, Cook, & Leviton, 1991). Although evaluation judgement can be subjective, it is not arbitrary, but rather based on stakeholder values. Different groups of stakeholders may value different things, and could therefore come to different conclusions of worth (Scriven, 1983). Under this approach, before any process or outcome evaluation information is interpreted, the evaluator must first understand the intended outcomes for all stakeholder and user groups. Evaluation metrics for MOOCs, like completion rates, would become valuable if such an outcome is important to a particular stakeholder group, such as the learners themselves. Evaluation findings are therefore interpreted through the lens of what stakeholders and learners value. Specific to MOOCs, the Contextualized Evaluation Framework, as an extension of the evaluation methodology proposed by Davidson (2005), includes a theoretical perspective that, in an open educational context, learner characteristics (e.g., intentions for learning content, level of preparedness for content, current career state, socio-economic demographics) and course characteristics (e.g., content, pedagogy, instructional design) influence learner behavior and ultimately the learning outcomes.

Researchers have begun to explore instructors' perspectives regarding the benefits provided by MOOCs, both to institutions, instructors, and the learners themselves. MOOC instructors have communicated a variety of reasons for teaching MOOCs, not all of which are directly related to the MOOC learners. For example, Najafi, Rolheiser, Harrison, & Håklev (2015) found that instructors believe teaching MOOCs would ultimately encourage better teaching practices on their campus. Instructors also have discussed the perceived benefit of show-casing their institutions "best courses" to a world audience (Evans & Myrick, 2015). Some instructors have discussed that MOOCs provide the opportunity to conduct research on student learning, behavior, and attitudes at a large scale (Zheng, Rosson, Shih, & Carroll, 2015). While researchers have found MOOC instructors to have some self-serving intended benefits from MOOCs, it is also true that many instructors are motivated by a sense of altruism and a genuine belief in the democratization of higher-education (Hew & Cheung, 2014). The literature provides much evidence to conclude that many instructors truly endorse MOOCs main value proposition: to provide high-quality education to those that could not otherwise access it (Evans & Myrick, 2015; Najafi et al., 2015; Zheng et al., 2015). How exactly "high-quality" open online education is defined has yet to be determined. Researchers have found some instructors question the quality of MOOCs in comparison to more traditional instruction, perhaps in part because they struggle with pedagogies for a massive open environment (Evans & Myrick, 2015).

One important pedagogical consideration for online distance education courses is the instructor presence (Baker, 2010). With enrollment easily in the thousands, the nature of the relationship between an individual instructor and their students in a MOOC is distinct from traditional classrooms (Haavind & Sistek-Chandler, 2015). While not all instructors, there is a group of MOOC instructors who have communicated a dislike for the often low levels of personal interaction with students (Hew & Cheung, 2014). Relatedly, instructors struggle with translating their classroom-based teaching practices to large numbers of learners (Zheng et al., 2015). Pedagogies that lend themselves to interpersonal contact have not found a place in MOOCs. There is an opportunity for both course developers and instructors to reconsider the role of the instructor and how to support MOOC students, perhaps through mechanisms to fulfill the roles of instructors or to aid instructors in effective class management. Supporting instructors with information

that will enable new strategies for increasing their impact in terms of teaching and learning will require deeper understanding of what value an individual instructor can bring to a mass of students and what instructors find valuable about the MOOC experience.

The range of educational objectives in MOOCs varies from personal health and financial choices to learning goals intended to prepare someone for highly-technical work. Instructor goals likely vary based on their educational objectives for the MOOC. Here, we focus on instructors who teach science, technology, engineering, or mathematics (STEM) MOOCs. In recent years, improving STEM education has been identified as a major goal by organizations such as the U.S. National Academy of Engineering and the National Science and Technology Council (National Academy of Engineering, 2004; National Nanotechnology Initiative, 2016). The push for STEM education has not gone unnoticed by MOOC providers. STEM MOOC initiatives include Georgia Tech's 2017 announcement that the school would offer an online Master of Science in Data Analytics in collaboration with edX (Diamond, 2017). A review of the literature found that while researchers have begun to explore MOOC instructors' goals, there is still a limited understanding of what STEM instructors hope will be the outcome of teaching a MOOC and what information would be useful to inform their teaching. Considering the foundation of evaluation is a needs assessment of the stakeholders, the purpose of this study is to explore STEM MOOC instructors' perspectives on teaching MOOCs and explore what information would be beneficial to them. Specifically, we asked what outcomes STEM MOOC instructors hope to achieve and what types of evaluation information are currently available to them. We aim to identify information that would be valuable to STEM MOOC instructors and could be used to inform their teaching and learning in open online educational contexts. In addition, administrators and members of instructional support team could use this information to guide the generation of outcome reports and to help in evaluating courses. Therefore, in this work, we consider the following two research questions: (1) What kind of course and learner information is available to STEM MOOC instructors for the purpose of evaluation? and (2) What kind of evaluative information would STEM MOOC instructors like to have available?

Methods

Participants and Data Collection

STEM MOOCs from a variety of institutions, fields, and nations were identified through a search of three large MOOC platforms: Udacity, edX, and Coursera. Emails were sent to the instructors of these MOOCs to recruit for interviews, with a $25 Amazon gift card offered as compensation. Interviews were conducted with instructors who agreed to participate until saturation was reached, indicated by a clear repetition of responses. Phone interviews were conducted with 17 instructors between April 2016 and July 2016. Of the 17 interviewees, 14 held tenure-track faculty positions, two were graduate students, and one was an industry professional and guest lecturer at an academic institution. We made the decision to exclude the interviews conducted with two graduate students from our results, as we felt that their perspective on evaluation might differ from that of the typical MOOC instructor. The fields of discipline and job titles for the remaining interviewees are listed in Table 1.

The interview protocol included an introductory statement to procure informed consent and inform interviewees that their responses were being recorded for research purposes. Interviews were conducted by two researchers using the responsive interviewing method described by Rubin and Rubin (2005). A semi-structured interview protocol consisting of open-ended questions was used for the phone interviews, allowing researchers to develop follow-up questions based on instructor responses. The recordings were transcribed by a third party. Upon completion, the transcriptions were checked and subsequently reviewed for quality.

The aim of the interviews was to capture and explore the experiences of various instructors and the design and implementation of their respective MOOCs. Interview questions were designed to focus on three areas of the relationship between instructors and their MOOCs: reasons for teaching a MOOC, information that would be useful for the instructor, and details about their experience teaching a MOOC. In the present study, we focus on the questions about information that would be useful for the instructor.

Table 1

Interviewees, Disciplinary Affiliations, and Job Titles

Instructor number	Instructor information Discipline	Job title
01	Computer Science	Full Professor
03	Nanomaterials	Lecturer
04	Industrial Engineering/ Operations Research	Professor
05	Electrical Engineering	Professor
06	Mechanics	Department Head
07	Comparative Media Studies	Visiting Lecturer
08	Computer Science	Associate Professor
09	Agricultural and Biological Engineering, Biomedical Engineering	Full Professor
10	Nanomaterials	Lecturer
11	Physics	Assistant Professor
13	Physics	Full Professor
14	Information Systems	Faculty
15	Mechanical Engineering	Professor
16	Mechanical Engineering	Senior Academic Professional
17	Information and Communication Technology	Associate Professor

[a] Missing numbers correspond to excluded participants.

Data Analysis

We followed qualitative methods based on a phenomenological perspective (Patton, 2002) to understand more about instructors' experiences with teaching MOOCs and their perspectives on what information would be beneficial to them. We followed Patton's guidelines for qualitative analysis, which include several steps. First, three of the authors explored the transcripts, writing memos, and taking notes on a line-by-line level. Next, based on the notes, a large number of initial codes were developed through consensus by two researchers, representing a wide variety of topics potentially relevant to instructor information use and pedagogical considerations. These codes were applied to the interview transcripts by segmenting and labeling text, and the resulting excerpts were grouped by code and further analyzed through a consensus process between two of the authors. Codes were tested for strength across interviews and similar codes were collapsed into larger categories that were reflective of all instructors. The authors then went back to review the transcripts to make meaning of each category and identify the themes. The remaining themes with example excerpts are provided. These themes are summarized in Table 2.

Table 2

Summaries of Themes in MOOC Instructor Feedback

Theme	Summary
Informal learner assessment	Instructors desire the kinds of personal interaction and immediate learner feedback that they use to supplement formal assessment and adjust their courses when teaching in person.
Learner audience	Instructors desire more information about who learners are, and what kinds of personal learning goals they are pursuing.
Course feedback	Instructors receive ratings and short reviews, but they desire more in-depth qualitative feedback from learners.
Learner usage data	Instructors report receiving different amounts of learner analytics. They want usage information that will enable them to improve their courses and predict learner performance.

Results

Our research team identified four major themes related to instructor information use that emerged during the coding process. These were informal learner assessment, learner information, course feedback, and learner usage data. These themes represent topics that were discussed consistently throughout the interviews and carry implications for MOOC evaluation.

Informal Learner Assessment

When asked about the differences between teaching MOOCs and teaching traditional courses, the most common topic brought up by instructors was the lack of information that could be used for informal or formative assessment of learner performance in MOOCs. Specifically, many instructors talked about desiring "learner presence," or the face-to-face interaction that learners have with instructors and teaching

assistants in a traditional classroom setting. Instructor 01 talked about how one-on-one interactions are used to assess students in the on-campus course on which their MOOC is based.

> Even if I personally don't know all those students, there is some TA who does know every student. Each TA has a section of twenty to twenty-five students that they meet with regularly. At the end of the semester, if I have any doubts about whether someone's on the borderline between an A and an A-, I can always talk to their TA and say, "Tell me about this person. Do they ask questions a lot? Do they come to office hours? Do they appear to be a solid contributor to their project group?" There's a human element of subjectivity when you're assessing student's performance at the end of the course.

Instructor 06 mentioned learner presence from the perspective of evaluating the course itself, saying that in a traditional classroom they can adjust their lectures in real-time based on the reaction of the learners.

> When you teach in a lecture room, for example, you have lots of lecture rooms where the lecture room is too big and I don't see the students. If you don't see the students, and you can hear if it's too silent or too noisy, and then you can adapt with what you are saying. But if you don't see them behaving, it's really difficult to adapt. As soon as you see them, when you say something they don't understand, you can say it again or do a summary and adapt something.

Instructors also expressed concerns about specific assessment techniques used in MOOCs, in particular those used for the assessment of open-ended assignments. Open-ended assignments can be assessed using peer grading or automated grading tools. Both approaches provide a final score for the assessment, but provide little information that would allow for instructors to evaluate whether or not learners are correctly applying their knowledge. When on-campus instructors personally grade an assignment or discuss grading with the TAs, there is an informal sense of what is going with students in the class. However, MOOC instructors discussed how their removal from the grading process can make them feel disconnected from the students. Instructor 01 explained the difficulties encountered when attempting to automate the grading of programming assignments.

> (W)hat we're finding is the automation actually is not always capturing if the students are getting it right. In particular, there are ways that you can either game the automation, or that the automation is just not perceptive enough, if you will. Sometimes the automation can measure if the student got the right result, but it's not always able to measure if the process that the student followed is the right process.

What the instructors' comments make clear is that, in a traditional classroom, informal forms of assessment are used to not only supplement formal forms of assessment, but to validate them as well. Instructor 05 described how the lack of information necessary to make informal assessments in their MOOC lowers their confidence in the effectiveness of the assessments used.

> There is no way for me to tell whether students who have successfully completed this assignment are in fact able to do some of the things that we would, for example, we would expect from our students [on campus]. The type of assessment that I'm thinking about is the type of assessment that

you see when you have the chance to have a conversation with students. I might not be able to do that with these people taking this class. You never know unless there is some clever mechanism to find that out. I will never think or know exactly how well this really works.

Learner Audience

Throughout the interviews, the instructors made it clear that they design their MOOCs to target a specific audience and to enable that audience to achieve specific learning objectives. Some MOOCs are intended for a specific set of learners, and others for more general audiences, but every instructor could clearly state for whom they had designed their MOOC. However, as Instructor 13 pointed out, they lack information that would let them know whether or not the learners who actually participate in a MOOC are a part of that intended audience.

> So I was basically planning that [the audience] would be just, you know, people with some knowledge of physics and science in general. It turns out that I was wrong. People who signed on, they were all over the place. You know, I [previously mentioned] high school students. We had some people who were retired. We had some students, people from other countries, where they just don't have access to physics. In fact, people who I thought would be interested didn't sign up. People who already go to the U.S. universities don't need my course, they can just, you know, get their own courses, real ones. Not what I thought would be the audience. I think I was just wrong.

In addition, learners who sign up for a MOOC may have learning goals that are completely different from those intended by the instructor. Even if their learning objectives do align, the learners may not choose to fully participate, as Instructor 13 continued to explain: "There are a lot of people who are just curious or interested to listen to some lectures but don't want to do any homework."

Some instructors noted that assessments in MOOCs are generally defined by alignment with the course's learning objectives, rather than also allowing learners to specify assessment of their own learning goals. Whether or not any given learner met her personal learning goals and got what she wanted out of the course is difficult to gauge using the information provided by these assessments. Unsurprisingly, many of the instructors described traditional measures of learner performance, such as completion rates and final grades, as being meaningless in the MOOC context. Instructor 15 described the gap between traditional assessments of course success and their idea of what would make their MOOC a success:

> (W)e've set this thing up as an educational opportunity. Our view, even though it wasn't measured, [is that] so long as you learn something...maybe you just watched the first 10 minutes of the first video that we've pretty carefully set up to introduce this whole field...and maybe from that you learned something about it that you didn't know before. To us, that would be a success. We're glad some people completed [the MOOC], but we weren't too hung up on the completion rates because we had a broader mission of providing education on a number of different levels.

A couple of instructors provided ideas on information that would enable problems related to meeting individual learner needs to be addressed. Instructor 09 believes that it would be helpful to have a more detailed breakdown of learner performance.

It would be interesting to know [a learner's performance] as a function of how much of the course they actually did, because I think it's possible that somebody did a quarter of the course but still got something out of it.

Instructor 03 suggested that their MOOC could accommodate a wider variety of learning objectives simply by giving learners the opportunity to provide more information about themselves and using that data to modify or add additional course material.

(H)aving a breakdown, having certain categories that people can put themselves into or their current level of qualification or categories for why they were attending the course or what they expecting to get out of it. That kind of thing would certainly be very useful.

Course Feedback

Feedback from learners is one of the primary sources of information that instructors make use of when evaluating a course in any setting. In MOOCs, feedback generally comes in the form of course ratings, short reviews, and end-of-course surveys. For example, instructors 16 and 05 summarized the feedback they received concerning their MOOC: "(T)here's a rating system. The students rate my courses out of 5.0. They give it so many stars. They can write comments, and they can write learner stories. New learners can see the ratings for my courses."

They have opportunities to provide ratings. They have opportunities to provide stories or reviews. Some of them do. Those are relatively short messages that basically stress satisfaction or some suggestion and so on. They don't reveal the level of the type of assessment that I would be interested in.

Some instructors attempt to glean course feedback from discussion board posts, but Instructor 07 provided an example of why this approach isn't always as useful as hoped.

I mean I get tons of comments in the forum, which is how I can gauge [satisfaction with the MOOC's assessments], but those are only the people who are kind of active and loud and saying things in the forums, right? It's not maybe necessarily the average student who's going to be posting in there and giving feedback.

The consensus among the instructors was that the course feedback currently received from learners, such as numerical ratings and short reviews, is sometimes useful, but it does not provide the depth of information that instructors would like to have available. Some instructors expressed a desire for more in-depth, qualitative feedback from some or all learners. Others gave specific examples of communication they had with individual learners after the MOOC was over that they found valuable. Instructor 05 expressed the belief that: "It would be nice to come up with some way of post-course interviewing students." Instructor 17 saw obtaining meaningful feedback as a major challenge in a MOOC and believed that the difficulty may be a symptom of large enrollments.

(I)t's such a huge number of people, that most people, they are completely silent, and for me, very frustrating that you don't get any feedback from [them]. Then the forum is very noisy, so it's very

difficult to...there are so many posts, that it's very difficult to find the good ones, so ways of filtering, cleaning, prioritizing all this information should be much nicer. I don't know. I don't have a specific solution, but I think it's a very common problem in these very large MOOCs, getting lost in all these big things, big numbers. Then the meaningful information is not there.

Learner Usage Data

Instructors reported receiving very different levels of learner analytics (e.g., clickstream usage data, material access counts, grading breakdowns, discussion board usage data), even when speaking about the same course platform. This is due to the different levels of data access available for purchase from major MOOC platform providers, as well as the constantly evolving capabilities of the course platforms themselves. A divide became apparent between instructors who were satisfied with the amount of learner usage data that they received from the course platform and those who were not. The instructors who were not satisfied with the amount of data that they received described two main ways that they wanted to apply learner usage data: making improvements to their course and predicting learner performance. Instructor 07 said that the data available to them limits their ability to evaluate and improve their course for the next offering.

> Learning the points where they drop off would be extremely valuable in updating the course content. Now, I just sort of get week-to-week where they drop off. By the time the course is over I have that. For the next year I could say, "Oh, we're getting a lot of people dropping off." You know, the week where they're building the prototype or something like that, so let's focus on that a little bit more. I can't get down to the level of what particular video did they drop out on, or what particular question did they drop out on?

Instructor 08 explains that having a more detailed breakdown of learner quiz performance would allow them to correlate lecture-viewing behavior with performance.

> I'd like to see for each question, what students' performance is on that question. It would be good to correlate whether students see the lecture or how much of the lecture they see, with their performance on the quiz questions themselves. That would be kind of interesting to study.

No matter how they intended to use the information, instructors who wanted more information expressed a desire to know more about the way that learners were interacting with their course. The common theme expressed by the instructors was a desire for useful, actionable information about the way that learners use course materials. Instructor 10 summarizes this desire:

> (C)an you create an instructor dashboard to monitor the students' behavior? To some extent, there's potential to do that. I don't think we fully realized that yet, but I think, ideally, what you want to see is, first of all, are our students actually keeping up with the material? Are they just using the material on a regular basis, consistently, or are they basically just skipping everything? Are they all doing binge watching, like just watch everything and do everything in the course of an hour, which is not necessarily bad, but the point is that you want to be able to at least see what are the patterns there.

Discussion

In this study, by following a process of interviewing MOOC instructors and coding their responses, we found four main instructor values that are relevant for informing criteria and metrics to evaluate MOOCs. The first is that instructors value high-quality assessment. Despite knowing that the bulk of learners are not using MOOCs as an actual course, these instructors still discussed a desire to use assessment to inform their teaching. In a traditional classroom setting, learners show evidence of understanding in a variety of ways, including assessment scores, interactions with instructors, and "showing their work" on open-ended assessments. The instructor can use this evidence as part of their pedagogy; such as posing questions to learners to assess understanding of the topic before moving on or even re-designing assessments to focus on topics with which the instructor feels learners may be struggling. However, MOOC learners are limited to the specific forms of expression defined by the course platforms. Interpersonal forms of assessment are generally limited to those that might occur on discussion boards and can be onerous for instructors to manage. In a MOOC, both content and assessments are generally static and developed well before any interaction with learners. Therefore, instructors are unable to use the assessments in a truly formative way to adjust instruction in real-time, even if they know a significant proportion of their learners did not understand a concept.

MOOC instructors complain about the feeling of speaking into a vacuum and missing "learner presence" when recording online lectures, both in the present study and in others' work (Hew & Cheung, 2014). Instructors in the present study explained that the lack of learner presence and other informal sources of information about learners forces them to rely entirely on formal assessments for evidence of learning. Unfortunately, MOOC platforms have fairly limited capabilities for assessment. Certainly, with the bulk of MOOC learners not engaged throughout the course, it would be prudent to focus higher-quality assessment on the smaller percentage of learners who actually do intend to use the materials as designed. This could also help instructors not to feel so overwhelmed by the masses, but rather have opportunities to support those few learners that want to gain the depth of information an entire course provides. The difficulty of implementing open-ended assignments in a MOOC environment precludes most forms of qualitative assessment of learner work. Attempts to implement open-ended assignments in MOOCs have generally involved peer grading or automated grading systems. Previous studies have called into question the reliability of both approaches (Hew & Cheung, 2014), a concern that was echoed by instructors in the present study. When evaluating the learning quality of MOOC platforms, institutions may consider what mechanisms are available for instructors to obtain direct and specific feedback on learners' understanding separate from graded work.

The instructors in our interviews agreed that learners in MOOCs often have personal learning goals that differ from those intended by the course designers, but differed on whether or not an attempt should be made to evaluate these goals. Currently, MOOC platforms offer limited capability to accommodate learners who have different goals, and therefore all learner assessment scores are lumped together in instructor dashboards. The instructors who wish to evaluate the attainment of personal learning goals discussed that they do not currently have opportunities to interpret outcomes based on an individual's desired goal. This points to one ongoing inconsistency in MOOCs: completion is regarded as an unimportant outcome because of the diversity of learner intent, but on the other hand, the outcomes provided to instructors are largely based on the extent to which learners met the course learning goals (i.e., performance on homework,

quizzes, and exams). Additionally, the specialized information being presented in STEM MOOCs can lead to very specific intended learning objectives. Even when the intended learning objectives are more general, the pre-requisite knowledge required to participate in some STEM MOOCs means the intended audience can be very narrow. As mentioned previously, other instructors simply do not see a need to evaluate whether or not learners have met their personal learning objectives. Instead, they see their MOOC as an open resource that learners are free to use as they wish. This view agrees with work by Liyanagunawardena, Lundqvist, & Williams (2015), which concluded that MOOCs should focus on serving their intended audience.

Instructors expressed a desire for improved feedback from learners, but providing this information presents a challenge for course platforms. Instructors agreed that in-depth post-course feedback allows them to understand the perspectives of learners in their course and make improvements to the curriculum. However, few learners are willing to provide such in-depth feedback when given the opportunity, and there is no way to guarantee that these learners form a representative sample of the course's participants. Many more learners are willing to provide feedback in the form of ratings and short comments, but instructors don't always feel that these forms of feedback are particularly helpful. Even if every learner in a MOOC could be persuaded to leave an in-depth review, how could instructors condense thousands of course reviews into usable information? Following the suggestion of one of the instructors interviewed in this study, a potential solution to the feedback problem could involve instructors contacting a sample of learners for post-course interviews. Our findings agree with work by Knox et al. (2014) which concluded that learner feedback must go beyond simple satisfaction ratings in order to be useful.

Similar to the divide between instructors who are interested in knowing their learners' learning objectives and those who are not, a split exists between instructors who want more detailed analytic information on learner behavior and those who are content with a general overview. The instructors who said that they want more detailed information on the ways in which learners are using their course made it clear that raw data itself is not necessarily useful. They need learner behavior data translated into actionable information that they can use to improve their course or predict further behavior and performance. Some instructors are doing this on their own, but currently, the process of gleaning information from the enormous sets of raw data provided by the course provider is rather cumbersome. Instructors expressed a desire for strong data visualization capabilities and real-time instructor dashboards, capabilities that major MOOC providers have been working to improve since the instructors of MOOCs in the present study were conducted.

Examining the results of the present study, an overarching theme emerges that unites the themes discussed thus far: similar to work by Stephens-Martinez, Hearst, & Fox (2014) we found that MOOCs provide an enormous amount of quantitative data for use in evaluation, but traditional evaluation also has a qualitative component that is largely missing. The focus on quantitative assessment is to be expected from the current limitations of online learning environments, but several instructors pointed out a troubling implication of that focus. In some cases, the lack of qualitative assessment capabilities in MOOCs can actually devalue the existing quantitative assessments. Without interpersonal and open-ended assessments, instructors have no way to validate the scores that learners receive on quantitative assessments. Instructors expressed concern that, while they receive exam scores and final grades for the learners who complete their course, they can't be as sure that those students are actually capable of applying the knowledge that they received

in the course as they would be in a traditional course setting. Similarly, instructors find it difficult to draw conclusions from learner analytics data such as completion and drop-out rates because they have no qualitative evidence that might explain the behavior. A learner who drops out because a course was too difficult and a learner who drops out because they achieved their personal learning objective make the same contribution to a course's completion rate despite achieving drastically different outcomes, rendering completion rate a significantly incomplete metric for instructors to use when evaluating their course. The research required to address these challenges aligns with the suggested research questions by London et al. (2016) regarding participants in MOOCs.

One limitation of the current study is that the instructors interviewed all taught on U.S.-based MOOC platforms (Udacity, edX, and Coursera). Their perspectives may be different from those who teach through other non-U.S. based platforms (e.g., Future Learn). However, the instructors themselves were from a number of different regions, including northern Europe, Asia, and India. Different course platforms employ a wide variety of instructional design techniques and emphasize different aspects of the online learning experience. Future research should have a search strategy to locate instructors that teach on other platforms with different pedagogical strategies. Additionally, the capabilities of MOOC course platforms are constantly evolving. Given the rapid pace of advancements in learning analytics, some of the concerns held by instructors in these interviews may have already been addressed by the time of this article's publication.

Conclusions

The primary aim of this research was to identify information that STEM MOOC instructors would find valuable. We found four main themes regarding instructors desire to: 1) informally assess learners, 2) assess learners' achievement of own learning goals, 3) have more representative learner feedback on course materials, and 4) have more detailed analytics regarding usage of course materials. From our findings, we recommend that evaluation criteria for MOOCs include: the quality of assessments, extent to which authentic formative assessment is possible, the capability to interpret learner outcomes based on learner goals, mechanisms for feedback, and metrics for evaluating specific course content. Instructors desire opportunities to formatively assess learners in an authentic way. One implication would be for platforms to create ways for instructors to have more authentic interaction with learners without being bogged down by the masses, perhaps employing sampling strategies. Instructors recognized that few learners participate in the end-of course surveys, limiting the type of learners from whom they receive feedback. One possibility might be for platforms to provide the end-of-course survey benchmarks that instructors could use to compare their feedback with other courses. Another would be for platforms to assist instructors in reaching out to learners who either disengaged or were sporadic to get feedback. Others have noted that learner intentions should be used to contextualize completion outcomes (Koller, Ng, & Chen, 2013), yet based on our interviews, these capabilities for individual instructors still appear to require further development. Instructors have accepted that not all learners want to fully engage with materials; however, they are often frustrated by their inability to sort out the extent to which learners did meet their goals. The quality of any learning resource is based on characteristics of both the learner and the resource. To go beyond simple reporting of access toward improved quality educational opportunities, it is imperative that outcomes are reported based on learner characteristics.

It is our intention that MOOC platforms could use these findings to inform the analytics they provide to instructors and partner institutions. In addition, administrators and members of instructional support teams could use our findings to evaluate the degree to which different platforms provide instructors with relevant information. MOOCs provide a great deal of data, but data alone is not sufficient for evaluative decision-making; more work is needed to contextualize raw data, to translate it into actionable information. Thus, future research should consider how instructors could use the analytics and dashboards currently or potentially provided by MOOC platforms now and in the future to inform and potentially improve their teaching and learning.

Acknowledgement

References

Baker, C. (2010). The impact of instructor immediacy and presence for online student affective learning, cognition, and motivation. *The Journal of Educators Online*, *7*(1).

Davidson, E. J. (2005). *Evaluating methodology basics: The nuts and bolts of sound evaluation*. Thousand Oaks, CA: SAGE Publications.

Diamond, L. (2017, January 11). Online master of science in analytics degree to be offered for less than $10,000. *Georgia Tech News Center*. Retrieved from http://www.news.gatech.edu/2017/01/11/online-master-science-analytics-degree-be-offered-less-10000

Douglas, K. A., Diefes-Dux, H. A., Bermel, P., Madhavan, K., Hicks, N. M., & Williams, T. V. (2017, June). Board# 32: NSF PRIME Project: Contextualized Evaluation of Advanced STEM MOOCs. In *2017 ASEE Annual Conference & Exposition*. https://doi.org/10.18260/1-2--27830

Evans, S., & Myrick, J. G. (2015). How MOOC instructors view the pedagogy and purposes of massive open online courses. *Distance Education*, *36*(3), 295–311. https://doi.org/10.1080/01587919.2015.1081736

Haavind, S., & Sistek-Chandler, C. (2015). The emergent role of the MOOC instructor: A qualitative study of trends toward improving future practice. *International Journal on E-Learning*, *14*(3), 331–350.

Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational Research Review*, *12*, 45–58. https://doi.org/10.1016/j.edurev.2014.05.001

Hollands, F. M., & Tirthali, D. (2014). *MOOCs: Expectations and reality. Full report*. Center for Benefit-Cost Studies of Education, Teacher's College, Columbia University, NY. Retrieved from https://files.eric.ed.gov/fulltext/ED547237.pdf

Knox, J., Ross, J., Sinclair, C., Macleod, H., & Bayen, S. (2014). MOOC feedback: Pleasing all the people. In S. D. Krause & C. Lowe (Eds.), *Invasion of the MOOCs: The promises and perils of massive open online courses* (pp. 98–104). Anderson, South Carolina: Parlor Press.

Koller, D., Ng, A., & Chen, Z. (2013). Retention and intention in massive open online courses: In depth. *Educase Review*, *48*(3), 62–63.

Liyanagunawardena, T. R., Lundqvist, K. Ø., & Williams, S. A. (2015). Who are with us: MOOC learners on a FutureLearn course. *British Journal of Educational Technology*, *46*(3), 557–569. https://doi.org/10.1111/bjet.12261

London, J., & Young, C. (2016). The role of massive open online courses (MOOCs) in engineering education: Faculty perspectives on its potential and suggested research directions. *International Journal of Engineering Education*, *32*(4), 1788–1800.

Najafi, H., Rolheiser, C., Harrison, L., & Håklev, S. (2015). University of Toronto instructors' experiences with developing MOOCs. *International Review of Research in Open and Distance Learning*, *16*(3), 233–255.

National Academy of Engineering. (2004). *The engineer of 2020: Visions of engineering in the new century*. Washington, DC: The National Academies Press. https://doi.org/https://doi.org/10.17226/10999

National Science and Technology Council, Committee on Technology, Subcommittee on Nanoscale Science, Engineering, and Technology, National Nanotechnology Initiative. (2008). *Strategy for nanotechnology-related environmental, health, and safety research. Arlington, Virginia: National Nanotechnology Coordination Office*. Retrieved from https://www.nano.gov/sites/default/files/pub_resource/nni_ehs_research_strategy.pdf

Patton, M. Q. (2002). *Qualitative evaluation and research methods* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

Rubin, H. J., & Rubin, I. S. (2005). *Qualitative interviewing: The art of hearing data* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Scriven, M. (1983). Evaluation ideologies. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 258–260). Springer Science & Business Media.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Thousand Oaks, CA: SAGE Publications.

Scriven, M. (2015, August 15). Key evaluation checklist. *Western Michigan University Evaluation Center*. Retrieved from http://www.michaelscriven.info/images/MS_KEC_8-15-15.doc

Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice* (Revised). Thousand Oaks, CA: SAGE Publications.

Shah, D. (2016, December 25). By the numbers: MOOCs in 2016. *Class Central*. Retrieved from https://www.class-central.com/report/mooc-stats-2016/

Stephens-Martinez, K., Hearst, M. A., & Fox, A. (2014). Monitoring MOOCs: Which Information Sources Do Instructors Value? In *Proceedings of the first ACM conference on Learning @ scale conference* (pp. 79–88). ACM. https://doi.org/10.1145/2556325.2566246

UNESCO. (2016). *Education 2030: Incheon Declaration and Framework for Action for the implementation of Sustainable Development Goal 4*. UNESCO Digital Library. Retrieved from

https://unesdoc.unesco.org/ark:/48223/pf0000245656.

Zheng, S., Rosson, M. B., Shih, P. C., & Carroll, J. M. (2015). Understanding student motivation, behaviors and perceptions in MOOCs. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 1882–1895). Vancouver, CA: ACM. https://doi.org/10.1145/2702613.2702628