# The Life Between Big Data Log Events: Learners' Strategies to Overcome Challenges in MOOCs

**George Veletsianos**
*Royal Roads University*
**Justin Reich**
*Massachusetts Institute of Technology*
**Laura A. Pasquini**
*University of North Texas and Royal Roads University*

*Big data from massive open online courses (MOOCs) have enabled researchers to examine learning processes at almost infinite levels of granularity. Yet, such data sets do not track every important element in the learning process. Many strategies that MOOC learners use to overcome learning challenges are not captured in clickstream and log data. In this study, we interviewed 92 MOOC learners to better understand their worlds, investigate possible mechanisms of student attrition, and extend conversations about the use of big data in education. Findings reveal three important domains of the experience of MOOC students that are absent from MOOC tracking logs: the practices at learners' workstations, learners' activities online but off-platform, and the wider social context of their lives beyond the MOOC. These findings enrich our understanding of learner agency in MOOCs, clarify the spaces in-between recorded tracking log events, and challenge the view that MOOC learners are disembodied autodidacts.*

Keywords: *massive open online course, big data, strategies, online learning, persistence*

REFLECTING a broader societal trend toward big data, one of the most exciting features of contemporary digital learning environments is the extensive data that they collect about student behavior and learning. These new data sets enable educational researchers to examine learning processes at almost infinite levels of granularity, from patterns across millions of learners to the moment-to-moment activities of individuals. Recently, large data sets consisting of the trace and log data that learners generate in online learning environments have been employed to understand dropout and attrition. One risk in this development is that researchers may mistakenly assume that the data automatically collected by learning platforms offer a comprehensive and complete representation of learner behaviors. Education researchers (e.g., DeBoer & Breslow, 2014; Perna et al., 2014) and others (boyd & Crawford, 2012) have warned that massive data sets do not necessarily track the most important elements of human participation in digital environments. In this research, we aim to illustrate more precisely what big data are missing about digital learning and how other data sources can enrich our understanding of these gaps.

This article draws on interviews with 92 students in massive open online courses (MOOCs) to examine the strategies they use to overcome the challenges they face. Our findings indicate that some of their strategies, such as rewatching videos, returning to a course after a break, and rereading content, could appear in data automatically collected by learning platforms. Other strategies, such as asking family members for assistance or searching the web for supplementary study materials, are invisible to tracking logs. We present three important domains of the experience of MOOC students that are absent from the tracking logs: the practices at learners' workstations, learners' activities online but off platform, and the wider social context of their lives beyond the MOOC. With the rise of sophisticated new methods of learning analytics (Siemens & Baker, 2012), researchers, faculty, and course developers have a growing and diverse body of quantitative data and computational research methods to guide improvements in MOOC design; our findings highlight the importance of drawing on other forms of research to design courses that support learners' needs and practices.

## Background and Context

The term *massive open online course* describes an evolving ecosystem of digital learning environments, characterized by a range of course designs. A recent analysis of the 2013–2015 empirical literature on MOOCs showed that the

topics of attrition and persistence have received the greatest attention in the literature (Veletsianos & Shepherdson, 2016). Even among those students who intend to complete a MOOC, the majority of them stop before earning a certification (Ho et al., 2015; Reich, 2014). Attrition from MOOCs follows a dependable pattern of very high early attrition in the early days of a course, followed by a much lower, but more steady, rate after the second week of a course (Evans, Baker, & Dee, 2015; Ho et al., 2015; Perna et al., 2014; Reich, 2015).

Researchers' efforts to understand why some MOOC students persist and others drop out have been only modestly successful. With a week or two of clickstream data, machine learning algorithms can predict with high degrees of precision which students will drop out of the course (Halawa, Greene, & Mitchell, 2014; Taylor, Veeramachaneni, & O'Reilly, 2014), but as is common with machine learning studies, these effective predictions have not yielded persuasive explanations. Learner surveys have identified a few important correlates of persistence, such as commitment to completion and socioeconomic status (Hansen & Reich, 2015; Kizilcec & Schneider, 2015; Reich 2015), but measures of specific motivations (e.g., advancing one's career or education) or motivation-related mind-sets have generally proven to be weak predictors of student completion (Greene, Oswald, & Pomerantz, 2015; Kizilcec & Schneider, 2015; Reich, 2015; Wang & Baker, 2015). Efforts at surveying students who dropout to understand their reasons for leaving have been plagued by low response rates (Halawa et al., 2014; Kizilcec & Schneider, 2015; Whitehill, Williams, Lopez, Coleman, & Reich, 2015), but these studies have found that external commitments and time constraints were the primary reported causes of student attrition. Kizilcec and Halawa (2015), for instance, found that 65% of respondents reported ceasing participation because they could not keep up with deadlines and 45% said that courses took too much time, as compared with 25% who found courses to be too difficult. These results both affirm and extend findings reported in the distance education literature: Students face personal and academic challenges (e.g., time constraints, isolation) that affect their success (e.g., Kember, 1989; Knapper, 1988; Sheets, 1992). While these results clarify the importance of constraints external to the courses on the subset of students who respond to the surveys, they leave many questions unanswered regarding what strategies students employ to respond to these challenges (Liyanagunawardena, Adams, & Williams, 2013; Milligan & Littlejohn, 2014).

For decades before the MOOC phenomenon, students have been learning at a distance, often through online means. The reasons for persisting in or leaving distance education represent a complex combination of individual and contextual factors (Aragon & Johnson, 2008; Wilging & Johnson, 2009). To better understand the strategies that learners employ to resolve the challenges they face, our research is informed by three theoretical frameworks: Pintrich and de Groot's (1990) theory of self-regulated learning, Rovai's (2002) model of distance learner persistence, and Cottom's (2014) critique of individualism in MOOC discourse. These theories point to the importance of understanding individual learner behaviors and how individual students are situated in broader contexts and communities.

Online learners are generally expected to be active, autonomous, independent, and self-directed (Moore, 1972, 1986), with the ability to manage their learning and employ learning strategies to achieve desired outcomes. This ability is often described as self-regulated learning (Zimmerman, 2000, 2008). Self-regulated learning refers to the degree to which learners are "metacognitively, motivationally, and behaviorally active participants in their own learning process" (Zimmerman, 1989, p. 329). According to Pintrich and de Groot (1990), the learning strategies that self-regulated learners use are varied and may be cognitive (e.g., elaboration), metacognitive (e.g., monitoring how well one's learning strategies are working), or related to resource management (e.g., interacting with others). Pintrich (2000) further notes that self-regulated learning is itself an active process that is influenced by learners' goals and environments. Rovai (2002) draws on ideas from Tinto (1975, 1993), Bean and Metzner (1985), Cole (2000), and Workman and Stenard (1996) to create a model of distance learner persistence that includes three dimensions: prior factors, internal factors and external factors. Prior factors include student demographics, preparation, and computer skills; internal factors include issues such as study skills, social integration, and access to support services. External factors include family responsibilities, financial situation, and life crises. Although the importance of the combination of internal and external factors is well established in the literatures on persistence in distance education and higher education (Tinto, 1975, 1993), Cottom (2014) argues that MOOC developers, researchers, and advocates often position students as "roaming autodidacts," idealized autodidactic personas disembodied from their communities and contexts. MOOC-focused research that uses *only* log data to model the interactions between individual students and the platform reinforces this individualistic discourse.

Using these three frameworks, we theorize that MOOC learners are likely to employ a number of cognitive, metacognitive, and resource management strategies to persist and resolve challenges and that these strategies exist on Rovai's dimensions (2002). Furthermore, we theorize that even though MOOC developers position learners as autonomous and roaming autodidacts, an investigation of their learning strategies will reveal contextual and relational aspects that are essential to their success or struggle.

Understanding these broader contexts requires research that extends beyond the data sets of platform tracking

logs. The large, continuously recorded tracking logs from online environments hold tremendous opportunity for new research, but there is a risk that data scientists and learning analytics researchers will reify the tracking logs as "the data." boyd and Crawford (2012) argue that false claims to objectivity and accuracy are inherent risks in the expansion of big data research. We add the false claim of comprehensiveness to that list. Research by Veletsianos, Collier, and Schneider (2015) from interviews with MOOC learners demonstrates that crucial learning activities (e.g., taking notes on paper) are happening off platform, where they cannot be recorded: Not all of the ways that learners respond to difficulties can be inferred from the logs.

To contextualize this argument in tacking log research, consider the analyses conducted by Guo and Reinecke (2014) regarding learners' pathways through MOOCs. These authors used tracking logs to define certain course navigation behaviors, including the "backjump." A backjump is a navigation event where students using an online environment move from one point to an earlier point in the navigational structure. Since backjumps are typically from assessment items to learning resources that can aid students in completing those assessments correctly, Guo and Reinecke presented backjumps as a type of strategy that learners use to address the challenges they face. The assumption embedded in this analysis is that a backjump represents two sequential learning activities. While the backjump may capture two adjacent entries in a tracking log, this two-part sequence might actually be interrupted by a myriad of learning activities, such as asking others for help, reading a book, or searching online. Inferences drawn exclusively from examining data logs are blind to these unrecorded events and experiences and risk over-weighting the causal impact of recorded actions over unrecorded actions.

In this study, therefore, we investigate learners' descriptions of the strategies they use to overcome the challenges they face, to better understand their worlds and the possible mechanisms for their attrition and to extend conversations about the use of big data in education. This research is important because it generates a greater understanding of learners' experiences, enables designers to develop learning designs that are sensitive to the ways that adult learners approach open courses, and furthers researchers' understanding of learners' activities and behaviors in open courses. We hope that these kinds of rich descriptions encourage computationally minded researchers (ourselves included) to recognize that even the most detailed tracking logs record only a fraction of the learners' activities and experiences. Ignoring this reality will lead to analyses that overemphasize the importance of recorded learner activities within learning environments that include digital systems.

**Method and Research Design**

We address the following research question using a basic interpretive study (Merriam, 2002): How do MOOC learners attempt to resolve the challenges they face? We were interested in understanding how learners make sense of their activities in MOOCs and, in particular, the strategies they use to address the challenges they faced. Because our primary goal was to uncover and interpret these meanings, a basic qualitative study was an appropriate method of inquiry.

*Participants*

In precourse surveys, students from four HarvardX courses on the edX platform were invited to participate in an interview study. We examined the precourse survey responses and course participation patterns from consenting individuals to identify a diverse group of participants to interview and gather as wide a range of experiences as possible. To this end, we invited participants who reported high and low levels of engagement and familiarity with course content, expressed high and low levels of commitment to complete courses, participated and did not participate on discussion forums, resided in North America as well as outside North America, and scored a passing or failing grade on assessments.

Ninety-two participants (Table 1), representing the largest number of students interviewed in a study on MOOCs, were recruited from four courses: ContractsX ($n = 21$), Statistics and R for the Life Sciences ($n = 19$), Einstein ($n = 31$), and Super-Earths and Life ($n = 21$). These courses spanned a wide range of disciplines and levels. Since these courses differed in their design (e.g., ContractsX included a heavy narrative component; Einstein was the only course that included peer-review activities; and SuperEarths was interdisciplinary), we anticipated that the experiences that we gathered would be applicable to a diverse range of MOOCs, although the generalizability of our findings may be limited by an examination of only a single platform.

*Data Collection*

Individuals were invited to participate in interviews around the midpoint of each course such that the learning experience was still fresh on their minds. Interviews were conducted over the telephone or Skype and lasted 30 to 45 minutes. They were audio recorded and transcribed verbatim. The interview protocol was open-ended and invited students to describe memorable moments, thoughts, and feelings in relation to their courses. Open-ended questions were followed with more targeted questions about challenges and the efforts to address those challenges. For example, participants were asked to "describe as many challenges or difficulties that you have faced in this course," which was followed by "What are some ways or strategies that you used to resolve these challenges?" At times, this question was

TABLE 1
*Participant demographic information*

| Participants | |
| --- | --- |
| Age Range | 22 - 81 |
| Mean Age (S.D) | 45 (14.93) |
| Reported Gender | 63 (m), 29 (f) |
| Location | North America (52), Europe (24) South America (6), Asia (8) Africa (1), Australia (1) |

posed differently (e.g., "How did you handle these difficulties?"), and learners were often probed to tell the interviewee more about the particular strategies or actions they took in response to the difficulties they faced. For instance, when learners mentioned that the material they were encountering was challenging, interviewers probed for actions to address this challenge by asking, "What did you do to understand material that was difficult?" These methods placed us in a position to listen to learners' descriptions of their lives, rather than evaluators of their online behaviors, thus allowing us to focus on gaining an understanding of the rationale behind and meaning of their behaviors.

### *Data Analysis*

We analyzed the data using an inductive and iterative process. First, the interviews were loaded in Dedoose, a qualitative analysis software application, and a single researcher read all transcripts. Whenever the researcher identified a segment of text in which participants discussed strategies aimed at addressing the challenges they faced, the researcher coded the segment as such. At the end of this process, 219 segments were identified/coded as strategies. All 92 participants identified at least one strategy. Next, the three of us independently read these data and engaged in a process of open coding by generating categories in response to the guiding question "How is the learner attempting to resolve the challenge she or he is facing?" An inductive and open coding process was used because little is known about the phenomenon in the MOOC context and the open coding process allowed us to remain open to what was available in the data. This activity was guided by the constant comparative approach (Glaser & Strauss, 1967). Specifically, each researcher read the first passage and generated a code to describe how the learner attempted to resolve the challenge. Next, the researcher read a second passage. If the already-generated code described the second passage, the passage was assigned the code, and the researcher moved to the third passage. If a new code was needed to describe the second passage, the researcher generated a new code and reread the first passage to examine whether the new code could be applied to the first passage. If the code was applicable, the

code was assigned to the first passage. If the code was not applicable, the researcher moved on to the next passage. This process of constantly comparing codes and passages allowed us to engage in in-depth, iterative, and comparative analysis. During this process, we independently generated 62, 58, and 39 unique codes, which were assigned 171, 286, and 344 times, respectively (see Table A1). After this round of independent coding, we discussed and compared notes and investigated possible themes to categorize the codes. We then reexamined the data independently and met three more times. During these discussions, codes were examined, questioned, debated, and eventually grouped into the themes presented here.

In this study, we took multiple steps to reduce the incidence of bias and establish validity as suggested by Creswell and Miller (2000) and Merriam (1995). Such steps included our independently analyzing the data, bracketing our preunderstandings of the phenomenon, searching for contradictory and disconfirming evidence in each other's codes, and presenting findings with participant descriptions intended to permit better understanding of learner strategies and evaluation of the credibility of the results.

### Results

As a preface to our investigation of student remediation strategies, we identified two key themes in student descriptions of the challenges they faced in courses. As with previous MOOC survey research and other research in the distance education literature, the primary challenge reported by students involved time constraints. MOOC students are busy with their families, work, and a variety of other commitments. Time for learning in MOOCs competes with all of these other commitments. Students also struggled with the intellectual difficulty of some material in MOOCs, especially when the prerequisites for a course were insufficiently explained or underestimated. While students expressed frustration about other features of courses—such as discussion forums that were hard to navigate or assessment items that were "tricky"—time constraints and the intellectual difficulty of course material emerged as the primary challenges.

In responding to these two challenges, learners described an array of activities aimed at resolving them. If tracking logs describe the interactions of students with the MOOC platform, our findings begin to characterize three overlapping domains beyond these interactions, as shown in Figure 1. First, learners work at workstations that include not only computers but also notebooks, paper printouts, reference books, additional devices, and other people. Second, students' online activities extend beyond the MOOC platform, to a variety of reference resources and online social networks that support student learning. Whereas many MOOCs are designed as a comprehensive learning experience,
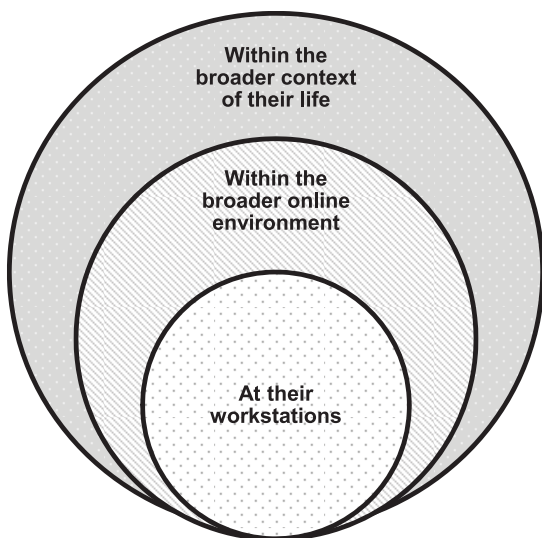
FIGURE 1.   *Three ways that learners address the challenges they face in massive open online courses.*

students appear to treat them as a single node in a broader network of learning opportunities. Finally, MOOC learning takes place in a broader learner world. This world goes beyond workstations, MOOC platforms, and online spaces, and it is a world in which students negotiate for time across multiple competing commitments. Next we describe students' strategies for persisting through challenges in each of these three domains.

### MOOC Student Workstations

To address challenges with course material, learners reported a variety of ways that they enhanced their physical learning environments, ranging from taking notes to printing course materials and engaging with people in their homes and workplaces. Several students described note-taking strategies both on paper and in other software or even on a second digital device. One student used several of these strategies:

> I have a ring-bound notebook, and I make quite neat notes if I can. Probably totally illegible, but to me they look neat. And I also have a ring file folder. And then I have some, on my hard disk, I keep some folders as well.

One student watched the lecture once with complete focus and then a second time while taking notes: "[I] watched the video through the faster speed first, [and] just really concentrated on the content. And then afterwards, [I] watch it through while taking notes." While tracking logs would record that this student watched each video twice at two speeds, they could not record this student using different learning strategies for each viewing. Another student read the lecture transcripts, took notes, and then watched the lecture video:

> So I would pretake notes, if you will, based on just reading the closed-captioning scripts and then I would listen to the lectures. So essentially, my brain had seen that information twice and three times for the stuff that I wrote down.

One student printed out whole sections of the course when the material got difficult:

> Some of them [weeks] were very difficult. And I had to, I actually printed out the lectures and read them, and sometimes I had to read them three and four times to really get so that I felt I had some real understanding of the concepts.

These reports paint a portrait of busy activity happening around a student's computer.

Learners' workspaces also include other print and human resources. Students bought books or borrowed from libraries or friends. One student described his purchase of two books (*R for Everyone* and *R for Dummies*): "I bought them because I was having a hard time in this class." She continued: "I need all the resources I can get to learn this thing [so] I'm gonna flood my desk with resources." Another student noted a similar strategy: "I've watched all the videos, and I'm doing some readings now. . . . I am reading *Einstein: The Life and Times*, by Ronald Clark. It's sort of a supplementary to the class."

In other cases, course participants received help from others external to the course, such as family members, friends, or colleagues. For example, one learner sought coding support from her ex-husband:

> My ex-husband is a computer scientist, he has a master's in computer science. He's an entrepreneur, he's a fantastic coding architect, and I just sat down with him and said, "I am stuck on this thing." It was like taking the mean of something as it iterated through a vector and he had no idea. And so I watched him try to Google how to figure out what was happening in this one line of code. And he's like "You gotta know what it's doing and how it's moving." And it was just such a basic concept but we couldn't . . . he couldn't find it anywhere. He finally interpreted it somehow and we figured it out.

While this section focuses on the learners' strategies by their workstations and physical environments, the ex-husband's turn to Google points to the next layer beyond the workstation: students' engagement with online resources beyond the MOOC platform.

### The Broader Online Environment of MOOC Learners

Researchers have argued that the shift from scarce to abundant content represents a radical change for higher education (e.g., Weller, 2011). In a society in which knowledge and expertise are scarce, a classroom—the professor, textbook, and other resources—might represent the totality of a student's opportunity to access certain learning materials or experiences. With the proliferation of learning resources

online, students are now perpetually situated in a richer online ecology of learning. Learners in this research treated MOOCs not as complete learning experiences but as important nodes in a broader network of resources for learning about a topic.

One of the signature pieces of evidence revealing connections beyond the course is the students' references to search engines and Wikipedia. For many of our interviewees, search engines (Google in particular) are a logical companion to the courses studied and a first station of departure when learners get stuck or confused. Wikipedia was another site specifically referenced by students in multiple courses as a resource that could help provide an orientation to learners finding their way in a new topic. In fact, whether assigned or not, Wikipedia emerged as a companion text to every course that we investigated. Sometimes students used such resources to bridge gaps between course prerequisites and their own background knowledge. As one student said,

> many times it was completely new to me so I had to look it up on Wikipedia or other Internet resources first . . . you were already [supposed] to have some background and basic terms of biology—which I didn't—so many times I had to look it up in Wikipedia to understand the [assignment] questions.

In other circumstances, online resources provided alternative explanations when course materials were unclear.

Students also participated in a variety of online communities and networks beyond the discussion forums provided in the courses. One student was part of an online study group that had formed in a previous class but remained a source of support and inspiration as students went their separate ways to pursue other MOOCs. Other students formed groups with peers in class over email or voiceover IP technologies such as Skype, or they joined related social networks with interests similar to the course—for example, a student in the SuperEarths course joined the Society of Modern Astronomy Facebook Group. Some students chose to pursue their studies in online communities that extended beyond the peers of a single class.

All these interactions—Google searches, Wikipedia pages, study groups, online interest groups, and other online resources—represent nodes in students' ecologies of learning that extend beyond the central node formed by the course.

### *The Lifeworld and Broader Context of MOOC Students*

Learners' strategies to address challenges also appear to revolve around their making conscious choices about their lives and course activities. Participants described abandoning courses that did not serve their needs, setting courses aside to take care of more pressing needs and returning to them as time allowed, stealing time from friends and family to complete courses, and skipping course activities that they deemed insignificant.

One of the most common narratives was that persistence and extra effort were effective antidotes to time constraints. For busy students, finding small chunks of time was critical. For instance, one learner who happened to be a higher education professional said,

> Maybe, if I'm lucky, I might get a block of two hours. But it's never predictable, I work, you know, a 60-hour work week, a 55- to 60-hour workweek. Matter of fact, I'm a little behind right now. So, I fit it in wherever I can. Sometimes even during office hours when there's no students, I'll do a module. It's real hard to predict. I just have to fit it in wherever I can. Sometimes in my car, I'll watch videos in my car while I'm driving. I'll put the thing in front of me, I'll put it on play and I'll just drive along and have a video running. Crazy.

One learner evocatively noted that "stealing time" enabled him to address course needs:

> The time . . . I have to steal from it. To steal from my kids, to steal from my work, I work at home. This is the time that I . . . in theory, I do not have it, but I have to steal it. I have to break the time and do those MOOCs. Mostly because I wanted to massage my brain and also because I'm really interested, deeply, deeply, deeply in how those MOOCs are going . . . I'm really interested in it. And that is what makes me step away from my job, mainly as a father, to do this.

In this instance, we see a learner managing external factors (i.e., family and work responsibilities) to participate in the course (cf. Rovai, 2002). When there was not enough time to steal, however, learners sometimes reported changing their levels of participation in courses. One student described how he made room to participate in one of the courses but at a slower rate than her other courses

> Student: I'm currently still participating, I'm a bit behind. I still want to finish the course, but yeah, I'm now [at] chapter eight.
> Researcher: And can you tell me the reasons why you kind of fell behind?
> Student: Yeah, of course it's a bit time related. So, I'm not only doing this course; I'm also doing some others. I was also following another massive open online course, which was more important for me to finish in time because I wanted to use the certificate later on, maybe. While with [this] course, it's purely out of interest.

Another student reported that she was having difficulties with the content and took a break to return to the course at a later time: "I felt like I was really putting my heart into it, and I still wasn't getting it . . . and that made me really upset . . . [so I] paused." A third student reported,

> I'm not finished yet. I need to spend a few hours in the course more, because I need to put all the work online. . . . But I still want to finish

the course, but at this moment I need to spend a little more time to get to the goal.

Other learners reported that they resolved the challenges they faced by skipping assigned activities and being selective in which tasks to complete. For instance, one learner said, "I went probably up to chapter six as far as being a participant in the course. And then after chapter six, I just started watching the lectures and no longer participating in the activities." Multiple learners reported watching the video lectures and trying to make sense of the content but avoiding the discussions and assessments.

When slowing or shifting participation is insufficient, students reported dropping out of courses entirely. Even in these cases, however, dropping out of one course can be a reaffirmed commitment to other courses. Several students reported signing up for multiple courses but persisting and completing only some of them. One student described that he stopped participating in one course after he realized that he enrolled in too many courses that seemed to require a lot of effort:

> I apparently didn't understand it or think about it enough, and I was trying to do three [courses]. I just thought, "Oh, yes. I'm in a candy store, this will be so great." And then I got in [the courses] and went, "Oh, yeah, collegiate level." [*chuckle*] So, it changed. I could see these are Ivy League collegiate classes.

In the data logs from a single course, this student would appear as one dropout. With a fuller picture of multiple courses across multiple platforms—which would be impossible for quantitative researchers to obtain at present—a more complex picture emerges of students refining their course commitments over time, shedding some and investing in others.

## Discussion

These findings enrich our understanding of learner agency in MOOCs, clarify the significance of the spaces in-between recorded events in the tracking logs of online learning environments, and challenge those who hold the archetypal view that all successful learners are disembodied autodidacts. Results reveal that learners are situated in particular places, in particular communities, with varying levels of commitment and engaging multiple strategies to overcome the hindrances they face. In accordance with Pintrich's (2000) research on self-regulated learning, these activities are rooted in learners' goals and environments. First, students work in actual physical spaces. They take notes, print materials, consult books, and interact with people in the same space. Second, they work in a broader ecology of online learning opportunities. They form study groups, search online for additional resources, and consult supplementary materials, such as Wikipedia. Finally, their studies take place in the broader context of their lives. They

prioritize among work, relationships, and courses, and they "steal" time from some activities to facilitate their studies. These three domains—the workstation, the ecology of learning, and the lifeworld—represent three spaces where students exist and participate in MOOCs in ways that leave no record in tracking logs. Understanding learners' experiences requires attending to their activities in these spaces. Turning these insights into actionable practices to improve student learning requires envisioning students beyond their representations as roaming autodidacts and disembodied producers of tracking log records.

Before we attend to such actionable practices, we need to acknowledge a number of limitations facing this study. First, this study depends on individuals who volunteered their time to participate in an interview, and as a result, it may be missing the perspectives and behaviors of those who did not volunteer to participate for a variety of reasons, such as lack of time. Second, this study focuses on four MOOCs offered by HarvardX, and these courses and this institution might have attracted a student sample that is unrepresentative of learners enrolled MOOCs in general. Finally, the results reported here depend on self-reported methods and, as such, may not identify all the strategies used or attempted by learners in their studies.

One productive response to these findings would be for course developers to design interventions or tools to support student strategies. Although students' commitments and time constraints may need to be addressed in alternative ways, course developers can examine the emergent behavior of learners in course platforms and support those behaviors. If there are more and less effective approaches to note taking and studying in particular disciplines, faculty can share those practices. Faculty can also facilitate sharing of high-quality notes generated by students and highlight effective study habits. Platforms can design learning materials readily printed in usable formats. If students are using references beyond the central platforms, faculty can design courses that intentionally guide students through a useful network of broader resources. Faculty can recommend additional readings, review and improve Wikipedia pages for key terms and concepts, and point students to alternative courses and sources. They can design their classes, not as hermetic containers of knowledge, but as central nodes in a network of learning resources about a topic. The best MOOCs may not be those that attempt to comprehensively introduce students to a topic but those that attempt to provide students with a map and compass of the network of online resources available for learning about such a topic. Such a design may be informed by past research in networked learning (e.g., Hodgson, de Laat, McConnell, & Ryberg, 2014). Future research in this area may explore how learners participating in open courses experience networked learning opportunities, which students succeed in networked courses, and which scaffolds may be necessary

to support learners of diverse needs and skills. Finally, if persistence, time stealing, reevaluation of needs, and reaffirmation of course commitments are central to student success, courses and platforms should explore interventions that support student plan-making and follow-through (cf. Duckworth, Peterson, Matthews, & Kelly, 2007) as well as reevaluation of needs (cf. Lucas, Gratch, Cheng, & Marsella, 2015). Such interventions are worthy of further research.

The lived experience and reported strategies of MOOC learners should also encourage us to reevaluate our interpretations of MOOC phenomena. From one perspective, students are signing up for multiple MOOCs and failing to complete some of them. From another perspective, MOOC students are inventing a much more dynamic course sampling and trial experience than what traditional residential "shopping periods" allow. Rather than committing to a set of learning experiences by an arbitrary date, participants dynamically allocate their attention and effort as they are able and in courses that prove capable of capturing their interest. Rather than ask the question "Why can't MOOC providers motivate students to make the same commitments that traditional higher education requires of course registrants?" it might be equally productive to ask, "Why can't traditional educational offerings adjust to allow students to make the kinds of flexible allocation of learning time and commitment that MOOC students are demonstrating?" While it makes sense for faculty to consider how they can reduce attrition within their own courses, it is also important to recognize that dropping a course can be a signal of a deepening commitment to another course.

In the spirit of the excitement generated by big data and learning analytics and education, our final recommendation is to use students' descriptions of their learning to identify new behaviors that can be examined via diverse research methods. One productive response to recognizing off-platform activities is to find ways to bring them onto the platform. EdX is experimenting with an internal note-taking platform that could encourage note taking and support sharing. Key Wikipedia pages might be imported into the platform so that student use of these resources can be evaluated, perhaps with an annotation tool such as the one described by Zyto, Karger, Ackerman, and Mahajan (2012). Instructors could supplement student Google searches with recommended materials and with trackable links for difficult concepts and key ideas. While rich descriptions of student experience reveal dimensions of learning that may never be trackable online, when important practices are found, platform developers could explore how these practices can be instrumented and nurtured.

Much of what we found about MOOC learners and their efforts to overcome challenges echoes and builds up what researchers such as Tinto (1975) discovered about residential college students or what Rovai (2002) found of earlier generations of distance education students. Many of our findings about how students learn in MOOCs may equally well describe students in today's traditional face-to-face or online courses. The implications of this are twofold: First, by investigating how we can support learners in MOOCs, we may be able to uncover ways to support learners in other contexts. Second, by exploring how learners in traditional contexts resolve the challenges they face, we may be able to design support structures for learners in digital contexts.

## Appendix

TABLE A1

*Codes Generated and Used More Than Two Times by Three Independent Researchers*

| Researcher A | Researcher B | Researcher C |
|---|---|---|
| Online research (22) | Time (32) | Time management (40) |
| Book (13) | External resources (26) | External resources (27) |
| Manage time (12) | Prerequisites (24) | Search (20) |
| Rewatch video (11) | Google (20) | Scheduling strategy (20) |
| Note taking (8) | Peer review (19) | Review (18) |
| Paused (7) | Rewatch (16) | Discussion board (18) |
| Review materials (6) | Quiz questions (14) | Selective with course components (17) |
| Skip (6) | Persistence (14) | Persistence (13) |
| Stopped (6) | Forums (11) | Quit (13) |
| Discussion forum (6) | Notes (11) | Rewatch (12) |
| Notes (5) | Other courses (9) | Reread (12) |
| Discuss with others (4) | Fluency (7) | Video/transcription (10) |
| Focus on items of personal | Steal time (6) | Feedback (10) |
| Significance (3) | Forums (4) | Troubleshoot (9) |
| Spend more time (3) | Wikipedia (4) | Course resource (9) |
| Tricks to reduce time spend (3) | Uneven difficulty (3) | Examples (9) |

*(continued)*

TABLE A1. (continued)

| Researcher A | Researcher B | Researcher C |
|---|---|---|
| Rereading (3) | No TA (3) | Take notes (8) |
| Documentation (2) | No video (3) | Platform suggestions (7) |
| Online (Google, Wikipedia) (2) | Online groups (3) | External support/help (7) |
| Other online resources (2) | Personal contact/relationship (3) | Study group (7) |
| Paying more attention (2) | Printing (3) | Practice exercises (6) |
| Refer to discussion board (2) | User interface (3) | TA/instructor support (6) |
| Rewatch video segment (2) | Health (2) | Prior learning experience (6) |
| | Hangouts (2) | Tutorials (5) |
| | Schedule (2) | Did nothing (5) |
| | Diverse students (2) | Visualize/draw (4) |
| | Other courses (2) | Print out materials (3) |
| | TA (2) | Syllabus (3) |
| | Expectations(2) | Summarize lesson (3) |
| | Jargon (2) | Study technique (3) |
| | Mobile (2) | Complain (3) |
| | No summaries (2) | Multiple attempts (2) |
| | Bored (2) | Translation (2) |
| | Grades/progress(2) | |
| | Video (2) | |

*Note.* Values in parentheses indicate the number of times that the corresponding code was used. TA = teaching assistant.

## References

Aragon, S. R., & Johnson, E. S. (2008). Factors influencing completion and noncompletion of community college online courses. *American Journal of Distance Education*, *22*(3), 146–158.

Bean, J. P., & Metzner, B. S. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research*, *55*(4), 485–540.

boyd, d., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, *15*(5), 662–679.

Cole, R. A. (2000). *Issues in web-based pedagogy: A critical primer*. London, UK: Greenwood.

Cottom, T. (2014, July 29). *Democratizing ideologies and inequality regimes in digital domains*. Retrieved from https://cyber.law.harvard.edu/events/luncheon/2014/07/cottom

Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory Into Practice*, *39*(3), 124–130.

DeBoer, J., & Breslow, L. (2014). Tracking progress: Predictors of students' weekly achievement during a circuits and electronics MOOC. In *Proceedings of the first ACM Conference on Learning@ Scale* (pp. 169–170). Atlanta, GA: ACM.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*(6), 1087–1101.

Evans, J. B., Baker, R. B., & Dee, T. (2015). *Persistence patterns in massive open online courses (MOOCs)*. Retrieved from http://cepa.stanford.edu/content/persistence-patterns-massive-open-online-courses-moocs

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. Chicago, IL: Aldine.

Greene, J. A., Oswald, C. A., & Pomerantz, J. (2015). Predictors of retention and achievement in a massive open online course. *American Educational Research Journal*, *52*(5), 925–955.

Guo, P. J., & Reinecke, K. (2014). Demographic differences in how students navigate through MOOCs. In *Proceedings of the first ACM Conference on Learning@ Scale* (pp. 21–30). Atlanta, GA: ACM.

Halawa, S., Greene, D., & Mitchell, J. (2014). Dropout prediction in MOOCs using learner activity features. Experiences and best practices in and around MOOCs. *eLearning Papers, 37*. Retrieved from http://www.moocsandco.com/sites/default/files/elearning%2037.pdf#page=7

Hansen, J. D., & Reich, J. (2015) Democratizing education? Examining access and usage patterns in massive open online courses. *Science*, *350*(6265), 1245–1248.

Ho, A. D., Chuang, I., Reich, J., Coleman, C. A., Whitehill, J., Northcutt, C. G., . . .Petersen, R. (2015). *HarvardX and MITx: Two years of open online courses fall 2012–summer 2014*. Retrieved from http://ssrn.com/abstract=2586847

Hodgson, V., de Laat, M., McConnell, D., & Ryberg, T. (2014). *The design, experience and practice of networked learning*. London, UK: Springer.

Kember, D. (1989). A longitudinal-process model of drop-out from distance education. *Journal of Higher Education*, *60*(3), 278–301.

Kizilcec, R., & Halawa, S. (2015). Attrition and achievement gaps in online learning. In *Proceedings of the second ACM Conference on Learning@ Scale* (pp. 14–15). Vancouver, Canada: ACM.

Kizilcec, R. F., & Schneider, E. (2015). Motivation as a lens to understand online learners: Toward data-driven design with the OLEI scale. *ACM Transactions on Computer-Human Interaction*, *22*(2), article 6.

Knapper, C. (1988). Media and adult learning: A forum. Lifelong learning and distance education. *American Journal of Distance Education*, *2*(1), 63–72.

Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: A systematic study of the published literature 2008–2012. *International Review of Research in Open and Distance Learning*, *14*(3), 202–227.

Lucas, G. M., Gratch, J., Cheng, L., & Marsella, S. (2015). When the going gets tough: Grit predicts costly perseverance. *Journal of Research in Personality*, *59*, 15–22.

Merriam, S. (1995). What can you tell from an N of 1? Issues of validity and reliability in qualitative research. *PAACE Journal of Lifelong Learning*, *4*, 51–60.

Merriam, S. (2002). *Qualitative research in practice: Examples for discussion and analysis*. San Francisco, CA: Jossey-Bass.

Milligan, C., & Littlejohn, A. (2014). Supporting professional learning in a massive open online course. *International Review of Research in Open and Distributed Learning*, *15*(5). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1855/3071

Moore, M. G. (1972). Learner autonomy: The second dimension of independent learning. *Convergence*, *5*(2), 76–88.

Moore, M. (1986). Self-directed learning and distance education. *Journal of Distance Education*, *1*(1), 7–24.

Perna, L. W., Ruby, A., Boruch, R. F., Wang, N., Scull, J., Ahmad, S., & Evans, C. (2014). Moving through MOOCs: Understanding the progression of users in massive open online courses. *Educational Researcher*, *43*(9), 421–432.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). San Diego, CA: Academic Press.

Pintrich, P., & de Groot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*, 33–40.

Reich, J. (2014, December 8). MOOC completion and retention in the context of student intent. *EDUCAUSE Review Online*. Retrieved from http://er.educause.edu/articles/2014/12/mooc-completion-and-retention-in-the-context-of-student-intent

Reich, J. (2015). Rebooting MOOC research. *Science*, *347*(6217), 34–35.

Rovai, A. P. (2002). Building sense of community at a distance. *International Review of Research in Open and Distributed Learning*, *3*(1). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/79/152

Sheets, M. F. (1992). Characteristics of adult education students and factors which determine course completion: A review. *New Horizons in Adult Education and Human Resource Development*, *6*(1), 3–18.

Siemens, G., & Baker, R. S. J. D. (2012). Learning analytics and educational data mining: Towards communication and collaboration. In *Proceedings of the Second International Conference on Learning Analytics and Knowledge* (pp. 252–254). Atlanta, GA: ACM.

Taylor, C., Veeramachaneni, K., & O'Reilly, U.-M. (2014). Likely to stop? Predicting stopout in massive open online courses. arXiv preprint. Retrieved on June 1, 2016 from http://arxiv.org/abs/1408.3382.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, *45*(1), 89–125.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition*. Chicago, IL: University of Chicago Press.

Veletsianos, G., Collier, A., & Schneider, E. (2015). Digging deeper into learners' experiences in MOOCs: Participation in social networks outside of MOOCs, notetaking and contexts surrounding content consumption. *British Journal of Educational Technology*, *46*(3), 570–587.

Veletsianos, G., & Shepherdson, P. (2016). A systematic analysis and synthesis of the empirical MOOC literature published in 2013–2015. *International Review of Research in Open and Distributed Learning*. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/2448/3655

Wang, Y., & Baker, R. (2015). Content or platform: Why do students complete MOOCs? *MERLOT Journal of Online Learning and Teaching*, *11*(1), 191–218.

Weller, M. (2011). A pedagogy of abundance. *Revista Española de Pedagogía*, *69*(249), 223–235.

Whitehill, J., Williams, J. J., Lopez, G., Coleman, C. A., & Reich, J. (2015). Beyond prediction: First steps toward automatic intervention in MOOC student stopout. Retrieved from http://ssrn.com/abstract=2611750

Willging, P. A., & Johnson, S. D. (2009). Factors that influence students' decision to dropout of online courses. *Journal of Asynchronous Learning Networks*, *13*(3), 115–127.

Workman, J. J., & Stenard, R. A. (1996). Student support services for distance learners. *Education at a Distance*, *10*(7), 18–22.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, *81*(3), 329–339.

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: Academic Press.

Zimmerman, B. J. (2008). Investigation self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*(1), 166–183.

Zyto, S., Karger, D., Ackerman, M., & Mahajan, S. (2012, May). Successful classroom deployment of a social document annotation system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1883–1892). Atlanta, GA: ACM.

## Authors

GEORGE VELETSIANOS is associate professor and Canada research chair in Innovative Learning and Technology at Royal Roads University, 2005 Sooke Road, Victoria, BC Canada V9B 5Y2; george.veletsianos@royalroads.ca. His research focuses on learner and scholar practices and experiences in digital learning environments.

JUSTIN REICH is a research scientist in the Office of Digital Learning and the executive director of the Teaching Systems Lab at the Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02319; jreich@mit.edu. His research focuses on learning at scale.

LAURA A. PASQUINI is a lecturer with the Department of Learning Technologies, University of North Texas, Denton, TX and a postdoctoral researcher with The Digital Learning and Social Media Research Group at Royal Roads University, Victoria, BC; Pasquini.Laura@gmail.com. Her scholarly work examines digital and open practices, specifically how social technologies impact and influence professional development, mentoring, learning organizations, and communities of practice.