Computers & Education 127 (2018) 41–65

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

The different relationships between engagement and outcomes across participant subgroups in Massive Open Online Courses

Qiujie Li*, Rachel Baker

School of Education, University of California, Irvine 3200 Education, Irvine, CA, 92697-5500, USA

ARTICLE INFO

Keywords:
Adult learning
Engagement
Achievement
Subgroups
MOOCs

ABSTRACT

Previous research has found that early engagement in MOOCs (e.g., watching lectures, contributing to discussion forums, and submitting assignments) can be used to predict course completion and course grade, which may help instructors and administrators to identify at-risk participants and to target interventions. However, most of these analyses have only focused on the average relationships between engagement and achievement, which may mask important heterogeneity among participant subgroups in MOOCs. This study examines how the relationship between engagement and achievement may vary across the four common behaviorally identified participant subgroups ("disengagers," "auditors," "quiz-takers," and "all-rounders") in three MOOC courses offered on the Coursera platform. For each of these subgroups, we used measures of behavioral and cognitive engagement from the first half of the ten-week courses to predict two outcomes: course grade and overall lecture coverage. Results indicate that the same engagement measure may be oppositely associated with achievement for different subgroups and that some engagement measures predict achievement for one subgroup but not another. These findings provide insight into both the benefits and the complexity of studying patterns of engagement from behavioral data and provide suggestions on the improvement of identification of at-risk participants in MOOCs.

1. Introduction

As higher education is becoming both more essential and less affordable for students in the United States (Stiglitz, 2014), technology-based instruction has been emerging as a cheaper and more accessible model (Pursel, Zhang, Jablokow, Choi, & Velegol, 2016). Among the offerings, Massive Open Online Courses (MOOCs) are seen as especially promising since they are open-access, offered at little to no cost, and often taught by established instructors at respected universities. Thousands of participants can enroll in a course at once, and many MOOCs allow participants to earn college credits or even obtain credentials or certificates (Barba, Kennedy, & Ainley, 2016; Siemens, 2013). In 2017, more than 800 universities offered 9400 unique MOOCs and 78 million people signed up for at least one course (Shah, 2017). However, the high enrollment has been paired with low rates of completion in most MOOCs. On average, fewer than 10% of registrants earn a passing grade and complete MOOCs (Ho et al., 2015; Jordan, 2014).

These low completion rates are associated with a number of factors at both the course and participant level: the structure and features of the course (Evans, Baker, & Dee, 2016; Nawrot & Doucet, 2014; Reilly, 2013; Zhang, Allon, & Van Mieghem, 2015), participants' background knowledge, preparation, and self-regulatory skills (Hansen & Reich, 2015; Kizilcec & Halawa, 2015; Nawrot & Doucet, 2014; Zheng, Rosson, Shih, & Carroll, 2015), participant' goals and motivation for taking a course (Barba et al., 2016; Evans et al., 2016), and participants' academic and social experiences in and out of class (Cottom, 2014; Tinto, 1997). Some

* Corresponding author.
E-mail addresses: qiujiel@uci.edu (Q. Li), rachelbb@uci.edu (R. Baker).

https://doi.org/10.1016/j.compedu.2018.08.005
Received 12 April 2017; Received in revised form 3 August 2018; Accepted 8 August 2018
Available online 11 August 2018

participants decide not to continue taking a MOOC because they have learned more about themselves, their own interests, or the course (e.g., Littlejohn, Hood, Milligan, & Mustain, 2016; Zheng et al., 2015). Some departures, however, are not the result of rational calculation and indicate a failure on the part of the participant or the class. Preventing such failures may be possible through inexpensive and simple interventions, such as sending automatic email reminders about the course to participants identified as at-risk. Previous studies show that these interventions can increase average MOOC completion rates (e.g., Davis et al., 2017; Dominguez, Bernacki, & Uesbeck, 2016; Whitehill, Williams, Lopez, Coleman, & Reich, 2015).

However, interventions only work if they are targeted at the appropriate participants; the efficiency and effectiveness of any intervention is dependent on early and correct identification of the participants who could, and want to, benefit from it (Baker, Evans, & Dee, 2016; Whitehill et al., 2015). MOOCs provide a promising, but also challenging, venue for such identification. Although typically MOOCs cannot provide the type of student data available in traditional education settings (e.g., past academic performance and transcripts, student surveys, and teacher ratings), participants in MOOCs leave behind an easily accessible but somewhat superficial trail of log behaviors, including data on every time a participant starts, pauses, or rewinds a video and every submission of an assignment or quiz. Past work in MOOCs has attempted to leverage these large swaths of behavioral data to measure engagement, predict achievement, and identify at-risk participants early in a course (e.g., Barba et al., 2016; Halawa, Greene, & Mitchell, 2014; Yang, Sinha, Adamson, & Rose, 2013).

Yet most of the prior work on identifying at-risk participants in MOOCs has assumed that the relationships between behavioral measures of engagement and course outcomes are the same across all participants in a MOOC and has not adequately examined whether these average relationships are masking important heterogeneity between participant subgroups (e.g., Balakrishnan & Coetzee, 2013; Barba et al., 2016; Brinton, Buccapatnam, Chiang, & Poor, 2015; Halawa et al., 2014; Pursel et al., 2016; Sinha, Jermann, Li, & Dillenbourg, 2014). Such heterogeneous relationship between engagement and outcomes might exist in MOOCs as people with very diverse backgrounds and motivations can enroll in the same MOOC (DeBoer, Stump, Seaton, & Breslow, 2013a; Kizilcec & Schneider, 2015). The rich behavioral data available in MOOCs have revealed consistent subgroups based on how participants interact with the course materials (e.g., Anderson, Huttenlocher, Kleinberg, & Leskovec, 2014; Arora, Goel, Sabitha, & Mehrotra, 2017; Chen, Haklev, Harrison, Najafi, & Rolheiser, 2015; Khalil & Ebner, 2017; Kizilcec, Piech, & Schneider, 2013; Ramesh, Goldwasser, Huang, Daume III, & Getoor, 2014; Sharma, Jermann, & Dillenbourg, 2015), and survey work has suggested that there are significant differences in background characteristics, such as prior knowledge and intentions, between these behaviorally defined MOOC subgroups (Chen et al., 2015; Kizilcec et al., 2013).

Thus, in these classes that enroll groups of participants that are diverse both in engagement patterns and in prior knowledge and motivation, relying on average relationships between engagement behaviors and achievement to identify at-risk participants may produce significantly flawed predictions. In this study we aim to address this concern by marrying two large areas of research on MOOCs: studies of the relationships between early MOOC behaviors and eventual outcomes (e.g., Balakrishnan & Coetzee, 2013; Brinton et al., 2015; Crossley, Paquette, Dascalu, McNamara, & Baker, 2016; Sinha et al., 2014) and examinations of the diversity of participant subgroups in MOOCs (e.g., Anderson et al., 2014; Chen et al., 2015; Kizilcec et al., 2013).

We motivate the need for such a study by noting that the average relationship between measures of engagement and outcomes within a group of participants is not only a result of the causal effect of the behavior but also a result of selection – which participants within that group choose to engage and why. When there is significant heterogeneity in participant background characteristics in a course, there might be multiple mechanisms that relate to which participants choose to be engaged, and these mechanisms might vary meaningfully across subgroups. For example, within a group of participants with little prior knowledge, those with relatively higher levels of prior knowledge may choose to be more engaged in the lecture videos, as they have a stronger sense of competence and expect their effort to pay off in learning gains (Metcalfe, 2009; Ryan & Deci, 2000). However, within a group of participants with deep prior knowledge, those with the most experience may choose to allocate the least time to the materials, as they are already familiar with the content and expect to gain little from the tasks (Metcalfe, 2009). This could be particularly true in the courses we examine (algebra and pre-calculus), as the topics covered in these courses build in a sequential and linear fashion.

In this exploratory study, we classified participants into subgroups based on their course interactions (watching lectures and submitting assignments) from the first half of three ten-week STEM courses and then tested whether the relationships between engagement and outcomes vary across these subgroups. Our results substantiate the hypothesis that engagement may predict achievement differently for different subgroups. These results suggest that predicting achievement in MOOCs based on early course activities without attending to this diversity might lead to over- or under- identification of at-risk participants and ineffective or even harmful interventions. These findings not only highlight the benefits and complexity of studying patterns of engagement from behavioral data, but may also aid in the improvement of interventions targeting at-risk participants.

2. Literature review

In this study, we examine whether measured engagement in MOOCs is differentially related to outcomes for behaviorally defined groups of participants. Such examination requires drawing upon three separate strands of related research: (a) defining and measuring engagement, (b) understanding observed relationships between engagement and learning, which requires examining both the causal effects of engagement and what predicts selection into engagement, and (c) establishing participant subgroups in MOOCs.

2.1. Engagement in educational settings

Defining Engagement. Educational engagement is usually defined as the investment of time, energy, and effort in learning activities (e.g., Fredricks, Blumenfeld, & Paris, 2004; Henrie, Halverson, & Graham, 2015; Pekrun & Linnenbrink-Garcia, 2012). Pulling from numerous studies and mirroring seminal work on engagement (e.g., Linnenbrink & Pintrich, 2003; Skinner & Belmont, 1993), the theoretical framework proposed by Fredricks et al. (2004) distills engagement into three distinct domains: behavioral, cognitive, and emotional engagement.[1] This framework has been commonly used to differentiate types of engagement and identify measures of engagement in both face-to-face classrooms and online learning environments (e.g., Appleton, Christenson, Kim, & Reschly, 2006; Henrie et al., 2015; Hew, 2015; Pellas, 2014). Prior work in online contexts has mainly used log data to examine two types of engagement—behavioral and cognitive engagement (e.g., Balakrishnan & Coetzee, 2013; Crossley et al., 2016; Cutumisu, Blair, Chin, & Schwartz, 2015; Sinha et al., 2014).

Behavioral engagement refers to participation in course activities and completion of course requirements (Archambault, Janosz, Morizot, & Pagani, 2009; Fredricks et al., 2004). In traditional school environments, typical indicators of behavioral engagement are positive behaviors, such as attending school regularly, paying attention in class, and completing homework (e.g., Archambault et al., 2009; Wellborn & Connell, 1987). Cognitive engagement is usually conceptualized as both the willingness to engage with difficult activities and the use of deep cognitive strategies to process, integrate, and generate new knowledge (Archambault et al., 2009; Dupeyrat & Mariné, 2005; Greene, Miller, Crowson, Duke, & Akey, 2004). Cognitive engagement is typically thought to require significant mental effort, which may produce deep understanding of target concepts, skills, and ideas (Biggs, 1987; Brunken, Plass, & Leutner, 2003; Pintrich & De Groot, 1990; Sweller, 1994; Weinstein & Mayer, 1986).

Measuring engagement. Typical methods for measuring engagement in traditional educational settings include questionnaires and observation (Fredricks et al., 2004; Helme & Clarke, 2001; Karpicke, Butler, & Roediger, 2009; Lee & Anderson, 1993; Lee & Brophy, 1996). Behavioral engagement has usually been measured by teacher or parent ratings and student self-reporting of the levels of various behavioral indicators such as class attendance, homework completion, and time spent studying (e.g., Carini, Kuh, & Klein, 2006; Finn, Folger, & Cox, 1991). Survey instruments, such as the Study Process Questionnaire (SPQ) and Approaches to Study Inventory (ASI), are often used to measure student-perceived cognitive engagement by examining how often students report using specific cognitive strategies that have been shown to promote integration, sense-making, and reflection (Biggs, 1987; Ramsden & Entwistle, 1981).

Studies in online learning environments, where self-reports and rich observational data are not readily available, have generally used the detailed log data from online learning platforms to measure both behavioral and cognitive engagement (e.g., Balakrishnan & Coetzee, 2013; Crossley et al., 2016; Macfadyen & Dawson, 2010; Sinha et al., 2014). While such measures are unable to directly capture either participants' mental effort or their perceptions on strategy use, measuring engagement with log data presents some advantages: these measures can be applied to a large number of participants in ways that would be impossible or inefficient with surveys or observations, and since they can be obtained quickly and repeatedly, these behavioral measures can help to identify at-risk participants while there are still ample opportunities to provide support to those participants.

A variety of frequency-type indicators have been used to operationalize behavioral engagement in online contexts (Henrie et al., 2015). Commonly used indicators include the number of clicks and pageviews (e.g., Cocea & Weibelzahl, 2011; Macfadyen & Dawson, 2010), the number of assignments or assessments completed (e.g., Thompson, Klass, & Fulk, 2012), and the number of discussion forum views and posts (e.g., Morris, Finnegan, & Wu, 2005; Peters, Shmerling, & Karren, 2011). In MOOCs, the most commonly used indicators of behavioral engagement are measures of the number or the percentage of lectures watched and assignments submitted by participants, as these are the most frequent activities in MOOCs (e.g., Crossley et al., 2016; He, Bailey, Rubinstein, & Zhang, 2015).

Cognitive engagement has also been inferred through behavioral data from online environments such as MOOCs (e.g., Brinton et al., 2015; Käser, Hallinen, & Schwartz, 2017; Sinha et al., 2014). As cognitive engagement involves processes that are not visible, much work has been done to create theoretically motivated sets of behaviors that can be used to infer cognitive engagement. A number of studies have based their measures of cognitive engagement on theoretical frameworks of information processing, such as Chi's Interactive-Constructive-Active-Passive (ICAP) framework (Chi & Wylie, 2014) and the levels of processing framework proposed by Craik and Lockhart (1972) (e.g., Brinton et al., 2015; Sinha et al., 2014; Wang, Yang, Wen, Koedinger, & Rosé, 2015; Zhang, Lin, Zhan, & Ren, 2016). In this paper, we draw on Chi's ICAP framework as it provides specific guidance for inferring cognitive engagement from the granular log data of video interaction available in MOOCs and other online courseware (Chi & Wylie, 2014).

Chi and Wylie (2014) differentiate students' overt behaviors into four ordered engagement levels—interactive, constructive, active, and passive—with interactive being the highest. Chi and Wylie's passive engagement, defined as receiving information without doing anything else explicitly related to learning, maps onto Frederick et al.'s conception of behavioral engagement, while active engagement and the two highest categories (interactive and constructive) fit Frederick et al.'s definition of cognitive engagement. Active engagement is defined as students actively manipulating learning materials to integrate new information with prior knowledge (Chi & Wylie, 2014). This matches the key feature of cognitive engagement: students strategically exerting mental effort in an attempt to better master the learning materials (Fredricks et al., 2004).

Chi and Wylie (2014) provide behavioral measures of active engagement specific to online educational contexts: pausing, repeating parts, and changing the speed of course videos. Such actions allow students to control the pace, order, and density of the information presented in order to adapt the information presented to their own cognitive processing need (Brinton et al., 2015; Mayer & Chandler,

[1] Emotional engagement refers to students' affective feelings for teachers, classmates, and coursework (Fredricks et al., 2004; Skinner & Belmont, 1993). Although researchers have started to examine emotional engagement in MOOCs, these studies have mainly relied on survey and forum data. Thus the target population of these studies has been limited to participants who responded to the course survey or participated in the discussion forum (e.g., Dillon et al., 2016; Wen, Yang, & Rose, 2014). In order to include a wider range of participants, we focus on behavioral and cognitive engagement in this study.

2001; Moreno & Mayer, 2007; Schwan & Riempp, 2004; Sinha et al., 2014). Following past research of engagement in MOOCs, this study measures pausing, seeking backward, and slowing down lecture videos to infer cognitive engagement (e.g., Brinton et al., 2015; Mayer & Chandler, 2001; Sinha et al., 2014). Just as is the case in traditional educational settings, these behavioral data cannot be considered pure measures of cognitive engagement but are instead measures that allow us to infer probable cognitive engagement.

2.2. The relationship between engagement and achievement in educational settings

Correlational patterns. Prior research has demonstrated a consistent positive relationship between behavioral engagement and achievement across various samples and settings. In both K-12 and higher education, studies on school engagement have found that students with higher levels of behavioral engagement, such as attending class regularly and spending more time studying, are more likely to persist and to have better school performance (e.g., Carini et al., 2006; Finn & Rock, 1997; Kuh, Cruce, Shoup, Kinzie, & Gonyea, 2008). Results from online learning environments and MOOCs have also demonstrated that participants who exhibited higher levels of behavior indicators, such as more page views (e.g., Cocea & Weibelzahl, 2011; Macfadyen & Dawson, 2010), video watching (e.g., Barba et al., 2016; Crossley et al., 2016), completion of assignments or assessments (e.g., Barba et al., 2016; Thompson et al., 2012), and discussion forum views and posts (e.g., Morris et al., 2005; Peters et al., 2011), were more likely to persist and achieve higher scores.

The association between cognitive engagement and outcomes, however, does not appear to be as clear. Some work has found a positive relationship between cognitive engagement and outcomes. In both traditional and online classes, prior studies have found a positive relationship between students' self-reported cognitive strategy use and course grades (e.g., Greene et al., 2004; Pintrich & De Groot, 1990; Puzziferro, 2008). A number of studies have also documented positive relationships between behavioral measures of cognitive engagement (mainly video interaction events) and persistence and performance in online learning environments including MOOCs (e.g., Brinton et al., 2015; Li & Baker, 2016; Mayer & Chandler, 2001; Sinha et al., 2014). In particular, studies in MOOCs show that pausing and backward seeking events are predictive of better quiz performance (Brinton et al., 2015; Li & Baker, 2016) and that slow watching and backward seeking events are positively associated with in-video persistence as well as course persistence (Sinha et al., 2014).

However, other studies have shown that cognitive engagement, measured both by self-report and by behavioral data, predicts lower achievement. For instance, Pintrich and De Groot (1990) found that reported cognitive strategy use was associated with lower performance for some students in a 7th-grade classroom. In higher education settings, several studies have found that students with lower grades reported higher use of cognitive strategies in both face-to-face and online classrooms (e.g., Boyer & Usinger, 2015; Wuellner, 2015). Similar negative relationships have been identified in studies of MOOCs: participants who seek backward and re-watch parts of the video are more likely have problems answering in-video quizzes (e.g., Giannakos, Chorianopoulos, & Chrisochoides, 2015; Kovacs, 2016) and seeking backward has been shown to be negatively associated with quiz performance (Dissanayake et al., 2018). Other work that has attempted to explain these relationships has shown that frequency of both pausing and slow watching in MOOCs is positively associated with participant-perceived video difficulty (Li, Kidzinski, Jermann, & Dillenbourg, 2015).

In sum, while the relationship between behavioral engagement and achievement is consistently positive, the link between cognitive engagement and achievement is complex and not unidirectional. As correlational evidence is the result of both causal influence and selection bias, the contradictory patterns for cognitive engagement may reflect the complicated relationship caused by both selection into, and the causal effects of, engagement.

Causal effects of engagement on learning. Experimental work can help to disentangle the causal and selection effects of engagement on learning outcomes. Robust evidence indicates that controlling for selection, both behavioral and cognitive engagement has positive effects on persistence and achievement (Fredricks et al., 2004). Students who are behaviorally engaged with course activities are more likely to learn and develop a sense of school belonging, which may lead to longer persistence and higher achievement (Fredricks et al., 2004). Although empirical evidence on the causal effect of behavioral engagement is comparatively thin, indirect but compelling evidence can be found from correlational studies that adjust for a large set of controls across diverse populations and educational settings (DeBoer et al., 2013b; Finn & Rock, 1997; Finn, 1993; Whitmer, 2013). Several large-sample studies of students in traditional classrooms have found large and significant differences in measures of behavioral engagement, such as attendance and active involvement in class activities, between high performing and low performing students and between students who dropped out and those who did not, after controlling for gender, race/ethnicity, socioeconomic status, psychological characteristics, and prior achievement (Carini et al., 2006; Finn & Rock, 1997; Finn, 1993; Kuh et al., 2008). In online settings, Whitmer (2013) found that after controlling for high school GPA, race/ethnicity, family background, and enrollment status, there is still a small but significant positive relationship between students' use of the learning management platform and final course grade.

Research in traditional education and online learning has similarly shown the positive effect of cognitive engagement on knowledge acquisition (e.g., Craik & Tulving, 1975; Gamino, Chapman, Hull, & Lyon, 2010; Mayer & Chandler, 2001). Learning activities that involve higher levels of cognitive engagement promote students' in-depth understanding of target concepts, skills, and ideas (Pintrich & De Groot, 1990; Weinstein & Mayer, 1986) and this comports with theories of information processing and learning approaches (Biggs, 1987; Craik & Lockhart, 1972; Marton & Säljö, 1976). Random-assignment experiments provide consistent causal evidence that learning activities that induce higher levels of cognitive engagement, lead to higher levels of performance on achievement measures over the material studied (e.g., Biggs & Rihn, 1984; Craik & Tulving, 1975; Scevak & Moore, 1998). Recent studies have demonstrated the positive effects of cognitive strategy instruction on math and reading scores and problem-solving performance (e.g., Gamino et al., 2010; Khezrlou, 2011; Montague, Krawec, Enders, & Dietz, 2014). In online learning, experimental evidence from early studies on interactive instructional videos shows that allowing students to interact with the videos by stopping, re-watching, and changing the speed, significantly improves learning (e.g., Mayer & Chandler, 2001; Schwan & Riempp, 2004; Zahn, Barquero, & Schwan, 2004; Zhang, Zhou, Briggs, & Nunamaker, 2006).

Selection into Engagement. Unlike in laboratory studies, in which engagement can be experimentally manipulated, students in traditional and online classes actively make choices about their own learning. A variety of environmental and individual factors are related to the extent to which students engage in learning activities and what learning strategies they use. Prior studies on engagement have suggested that students with different levels of prior knowledge may select into engagement through different or even opposite mechanisms (Metcalfe & Finn, 2008; Pintrich, 2002). Considering this, selection into engagement is important when examining the observed relationship between engagement and outcomes in participant populations with diverse backgrounds, such as MOOCs.

On one hand, students with more prior knowledge tend to have higher perceived competence and expectancy for success. According to motivation theories, such as expectancy theory (Eccles, 1983) and self-determination theory (Ryan & Deci, 2000), students with stronger perceived competence are more motivated to engage in a task and are more willing to use cognitive strategies to enhance learning (Pintrich & De Groot, 1990; Pintrich & Garcia, 1991; Zimmerman & Martinez-Ponz, 1992, pp. 185–207). On the other hand, people's beliefs about their prior knowledge may also be negatively related to behavioral and cognitive engagement (Metcalfe & Finn, 2008). First, people tend to focus specifically on the information that they do not yet know but that is in their region of proximal learning; they allocate less effort to material that they believe they already know (Metcalfe, 2009). Moreover, people tend to selectively choose the most effective engagement strategies (e.g., re-reading, rehearsal, and elaboration) for their goals (Rohrer & Pashler, 2010) and to use more cognitive strategies when they perceive higher levels of challenges in the learning tasks (Fredricks, Blumenfeld, Friedel, & Paris, 2002; Meece, Blumenfeld, & Hoyle, 1988).

Given the varied factors associated with students' decisions to engage with course material, selection into engagement could be either a positive or negative signal – it could indicate a high level of competence and motivation, or it could indicate that the student is unfamiliar with the material and is experiencing a high level of challenge. Thus, while both cognitive and behavioral engagement are associated with increased performance at the individual level, differential selection into engagement might affect the relationships between engagement and outcomes across a group of students. Selection may boost the positive relationship between engagement and achievement, attenuate the observed relationship toward zero, or even produce a negative relationship between engagement and achievement. This implies that, within a diverse student population, predicting outcomes based solely on the use of engagement behaviors could be challenging. In this paper, we examine the relationships between behavioral indicators of engagement and course outcomes for different participant subgroups in MOOCs. By allowing for varied relationships for participant subgroups, this study may help to illustrate the existence and impact of the different underlying processes that these behaviors may stand for.

2.3. Subgroups in MOOCs

The fact that selection into engagement is dependent on participant characteristics motivates the need to better understand the different types of participants present in MOOCs. Unlike traditional educational research, in which demographic, survey, and past academic data can be used to classify participants into meaningful groups, MOOC researchers and instructors must rely primarily, if not exclusively, on behavioral data. Past work has consistently shown that the great diversity in the behavioral patterns of participant engagement in MOOCs can be distilled down into a few distinct categories (e.g., Anderson et al., 2014; Arora et al., 2017; Chen et al., 2015; Ferguson & Clow, 2015; Kahan, Soffer, & Nachmias, 2017; Khalil & Ebner, 2017; Kizilcec et al., 2013; Pursel et al., 2016; Ramesh, Goldwasser, Huang, Daume, & Getoor, 2014; Rodrigues et al., 2016; Sharma et al., 2015). These studies have attempted to classify participants based on their interaction with lectures and assignments using a variety of methods and approaches, including bottom-up approaches to identify potential subgroups (such as K-means clustering) (e.g., Kizilcec et al., 2013) and top-down approaches to classify participants into pre-defined groups (e.g., Anderson et al., 2014).

Across all methods, four common subgroups have been identified: disengagers, auditors, quiz-takers, and all-rounders (e.g., Anderson et al., 2014; Arora et al., 2017; Chen et al., 2015; Khalil & Ebner, 2017; Kizilcec et al., 2013; Pursel et al., 2016; Ramesh et al., 2014; Sharma et al., 2015). Other subgroups, such as “social engagers”, “offline engagers” have also been identified in prior research based data that is relatively less frequently used in MOOCs, such as forum discussion or video downloading.

Disengagers are participants who register for a course but do not intend to complete it. Their total interaction with lectures and assignments is very limited. Studies have used different terms to label these participants, such as “bystander”, “disengagers”, “taster”, and “all-rounders with low intensity” (e.g., Anderson et al., 2014; Chen et al., 2015; Kahan et al., 2017; Kizilcec et al., 2013). In most of the courses studied, disengagers were the largest subgroup. *Auditors* are participants who interact with a course primarily by watching lectures; they submit assignments infrequently, if at all. Different terms used to label these participants include “viewers,” “auditors,” and “casual students” (e.g., Anderson et al., 2014; Chen et al., 2015; DeBoer, Ho, Stump, & Breslow, 2014; Kizilcec et al., 2013). On the contrary, *quiz-takers* are participants who interact with a course primarily by submitting assignments; they watch very few, if any, lectures. Different terms used to label these participants are “solvers,” “performers,” and “quiz-takers” (e.g., Anderson et al., 2014; Arora et al., 2017; Chen et al., 2015). *All-rounders* are participants who are most similar to students in traditional courses; they watch lectures and submit assignments. Studies have used different terms to label these participants, including “completing students,” “achiever,” “all-rounders,” and “all-rounders with media and high intensity” (e.g., Anderson et al., 2014; Arora et al., 2017; Chen et al., 2015; Kizilcec et al., 2013). All-rounders reported the highest levels of course satisfaction (Kizilcec et al., 2013) and many of them earned relatively high course grades (Anderson et al., 2014).

Prior research, which has leveraged self-reported data available from surveys in some MOOCs, has shown that these participant subgroups are likely to differ in their intentions and backgrounds. Chen et al. (2015) identified significant differences in participants' self-reported intentions across subgroups. They found that participants who enroll for career and academic reasons are more likely to be quiz-takers and participants who enroll for fun and enjoyment are more likely to be all-rounders. Similarly, Kizilcec et al. (2013) found that all-rounders were most likely to say that they enrolled in the course for fun and challenge than other subgroups. They also found that although many auditors achieved a grade of zero, their reported satisfaction was as high as completing participants. These results suggest that auditing could be an alternative engagement pathway and that focusing only on specific academic outcomes such as course grade may overlook the success of these participants who are engaged in a course with no intention to earn a certificate. In addition, there is evidence that these subgroups have different prior knowledge and education background. Anderson et al. (2014) found that some quiz-takers achieved high course grades, suggesting that they might have learned the content previously or they were learning it elsewhere. Similarly, Khalil and Ebner (2017) found that quiz-takers are more likely than any other subgroup to be currently enrolled undergraduate students.

2.4. Research questions

The current literature in these two areas (participant subgroups in MOOCs and the relationship between participant engagement and outcomes) provides the starting point for this study, which uses log data to examine differences in the relationship between engagement and achievement in MOOCs for various subgroups of participants. Prior studies have found that measures of both behavioral and cognitive engagement can predict achievement. While behavioral engagement is consistently predictive of longer persistence and higher course grade, cognitive engagement may predict either higher or lower course performance.

Past research in MOOCs has either assumed common relationships between engagement and achievement in the whole participant population (e.g., Balakrishnan & Coetzee, 2013) or has only examined a relatively homogeneous participant population, such as participant who posted in the forum or participant who persisted to the end (e.g., Brinton et al., 2015; Crossley et al., 2016; Li et al., 2015). These general findings may mask more nuanced differences between participants with different intentions and backgrounds. Previous research in MOOCs have identified consistent subgroups who are likely to differ in their intentions and backgrounds and these behaviorally-defined subgroups in MOOCs can serve as a starting point for examining the heterogeneity of the relationships between engagement and achievement in MOOCs.

The exploration of the different relationships between engagement and achievement for each subgroup also requires diverse measures of achievement, as some subgroups may not have the intention to complete a course or earn a certificate (e.g., disengagers and auditors). While prior research has tended to either ignore differences in intentions among participants and only use course completion or course grade as an outcome for all participants (e.g., Balakrishnan & Coetzee, 2013; Crossley et al., 2016; Jiang, Williams, Schenke, Warschauer, & O'dowd, 2014) or exclude from the analysis participants who show no intention to pass the course (e.g., Barba et al., 2016), we argue that identifying at-risk participants in these subgroups and implementing interventions that could provide supports to help them achieve their self-defined goals could maximize the benefits of MOOCs. However, to identify at-risk participants in these subgroups, achievement must be measured in a way that is better aligned with their intentions. Therefore, in this study, we used both overall lecture coverage and course grade to allow for flexible definition of achievement and to better capture the success of a wider range of participants.

We therefore pose the following research questions: Do behavioral and cognitive engagement indicators predict achievement differently across the four subgroups previously identified in MOOCs? If so, how?

3. Method

3.1. Sample and data

Courses. Instructional design and pedagogical practices vary greatly across MOOCs. Most MOOCs contain some combination of the most common features (video lectures, auto-graded or peer-reviewed assessments such as multiple choice quizzes and essays, discussion forums, readings, and activities) though there is great variation in the number of each feature offered and the quality of offerings (Baturay, 2015; Grainger, 2013; Margaryan, Bianco, & Littlejohn, 2015). MOOCs are typically paced using an asynchronous weekly structure, in which participants can access weekly resources on their own timeline, though some courses also offer synchronous activities (such as live lectures or discussions) (Baturay, 2015). Given the great variety of pedagogical practices and instructional design present in MOOCs, and the importance of these factors for predicting participant performance, we limited our sample in this study to MOOCs with relatively similar structure and instructional features.

We used data from three courses that were offered by a selective public university on the Coursera platform. All three courses were academic math courses (Algebra and Pre-Calculus) that aimed to prepare participants for a college-level Calculus course. In each course, participants who completed the course with a passing grade earned a free electronic certificate, which did not represent official academic credit from the institution offering the course.

The three ten-week courses offered participants 104–117 short lectures (3–15 min each) and 32–52 small quizzes (see Table 1). Each lecture and quiz covered one or more topics about algebra or trigonometry. Lectures and quizzes were released week-by-week. Participants had the flexibility to choose what materials or activities to engage with and could adopt different engagement patterns based on their own intentions and backgrounds. Participants were also allowed to submit the quizzes as many times as they wanted.

Our decision to investigate MOOCs that have multiple lecture videos and quizzes available each week is important for a few reasons. First, it allowed us to use similar methods across multiple classes to ensure that our results are not an artifact of the particularities of one class. Second, by using classes that have a number of both lectures and quizzes (some MOOCs have very few quizzes or only one lecture each week), we were able to pick up variation across weeks and examine multiple dimensions of

Table 1
Course design and participation.

Course features	Algebra	Pre-calculus 1	Pre-calculus 2
Course			
Length	10 weeks	10 weeks	10 weeks
Number of lectures	104	115	117
Number of quizzes	32	52	32
Participants			
Enrollment	63,202	50,676	46,524
Active participants	24,171	20,288	26,998
Achievement of active participants			
Overall Quiz coverage	0.20	0.21	0.13
Overall lecture coverage	0.13	0.15	0.11
Final grade	6.70	6.68	4.63

Note. Precalculus 2 is very similar to Precalculus 1except for some small changes in the order and number of lectures. Overall quiz coverage/overall lecture coverage refers to among all the quizzes/lectures available, the proportion a participant accessed.

engagement. Third, classes with this structure allowed us to examine participant subgroups using features and methods that are similar to past studies (e.g., Anderson et al., 2014; Chen et al., 2015; Kizilcec et al., 2013).

Overall, more than 160,000 participants enrolled in the three courses. However, less than half of these participants ever watched a lecture or submitted a quiz. Participants who watched at least one lecture or submitted at least one quiz were considered as active participants (see Table 1, N = 71,457) and were used in the subgroup classification. As shown in Table 1, on average, active participants watched around 11%–15% of the available lectures and submitted about 13%–21% of the available quizzes in each course. Although the overall engagement level of the sample courses was very low, it is comparable to the engagement levels in STEM MOOC courses (Evans et al., 2016), which have been commonly used in MOOC subgroup studies. For the active participants, the average course grade, a weighted score of quizzes, homework assignments, and final exam on a 100-point scale, was very low in the three courses, ranging from 4.63 to 6.70 (see Table 1).

3.2. Data source

The data used in this study were collected through the Coursera platform in 2013 and 2014. The data include information about: 1) Course design: The Coursera database contains unique identification and names of the lectures and quizzes, the week in which they were released, and the order of lectures and quizzes within each week; 2) Quiz record: The Coursera database also provides records of quiz submission details, including when a participant submitted a quiz and what score the participant earned; and 3) Log data: the log of video interaction, in which participants' every click event of the video player (i.e., play, pause, seek forward, seek backward, and rate-change) was recorded.

While there are rich data detailing participants' behaviors in MOOCs, the information about participant backgrounds is very limited due to low survey response rates. In each course, around 5% of the registrants submitted the pre-course survey. Using such data would restrict the sample to a small subset of participants and significantly reduce the diversity of the participant population. Thus, we did not use the survey data in this study.

3.3. Subgroup classification

We classified participants into the four subgroups using behavioral indicators. In previous studies, the two main academic activities—lecture watching and assignment submission—were used for characterizing engagement patterns (Anderson et al., 2014; Chen et al., 2015; Kizilcec et al., 2013). Similar to these studies, we generated and used three indicators, lecture coverage (i.e., of all the lectures available, the proportion a given participant has played, which is an indication of participants watching the lecture), quiz coverage (i.e., of all the quizzes available, the proportion a given participant has attempted), and quiz-lecture ratio (i.e., the ratio of quiz coverage to lecture coverage), to place participants into subgroups.

Based on the conceptual definitions in the previous literature, we created operational definitions for the four subgroups using specific cutoffs of the three indicators. Cutoffs for lecture coverage (l_0) and quiz coverage (q_0) were mainly used to differentiate disengagers from the other three subgroups—auditors, quiz-takers, and all-rounders—who were all engaged with the courses though in different ways. Quiz-lecture ratio was used as an indicator to see if a participant exhibited any preference between quizzes and lectures. Two cutoffs of this indicator (r_0 *and* r_1) were used to differentiate auditors, quiz-takers, and all-rounders. Accordingly, a participant whose lecture coverage and quiz coverage were L and Q was assigned to: 1) Disengagers if $L < l_0$ and $Q < q_0$; 2) Auditors if $L \geq l_0$ and $Q/L < r_0$; 3) Quiz-takers if $Q \geq q_0$ and $Q/L >= r_1$; and 4) All-rounders if $L \geq l_0$ and $r_0 =< Q/L < r_1$.

Since there is little consistency in the operational definitions for each subgroup in the literature, we chose to employ relatively strict cutoffs to make sure that participants who were assigned to a certain subgroup matched the behavioral features of that group. Accordingly, both l_0 and q_0 were set to be the course average and r_0 and r_1 were set to be 0.5 (i.e., if 1% of the available lectures were watched, 0.5% of the available quizzes were submitted) and 2 (i.e., if 1% of available lectures were watched, 2% of the available quizzes were submitted), respectively.

Moreover, although most prior studies used cumulative data from entire courses to identify potential subgroups, we used data from the first five weeks of the course for subgroup classification. In deciding how many weeks of data to use, we were balancing identification early in a course with correct identification, aiming at both timely and valid diagnoses of at-risk participants in each subgroup. In extensive testing of the data from early weeks of the course, we found that five weeks was the earliest point at which we could identify more than 70% of participants in all groups. We therefore used these first five weeks of data to identify participant subgroups. We return to these decisions regarding subgroup classification in the robustness analysis below and Appendix B includes a full explanation using different data and different cut-offs for the three indicators.

3.4. Measures

Engagement. We measured behavioral engagement during the first five weeks using lecture coverage and quiz coverage, two commonly used variables in MOOC studies. On average, participants in the analysis sample had watched about 14% of the available lectures and submitted around 19% of the available quizzes at the end of week 5. We standardized both halfway-course lecture coverage and halfway-course quiz coverage in the regression analysis to facilitate direct comparison across the two predictors. While a measure of time-on-task would be another helpful measure of behavioral engagement, these data were not saved by the platform we used.

We measured cognitive engagement by how often participants conducted each of the three video interaction events (i.e., pausing, backward seeking, and slow watching) while watching lectures. We defined a pausing event as when a participant stopped a lecture by clicking the pause button. To account for the fact that some participants may be cognitively engaged in the lectures they watched, but watched very few lectures in total, the average number of pausing events (i.e., the frequency of pausing events *per lecture watched*) was used. A backward seeking event occurred when a participant moved the playhead of a video to a new position and the new position was before the old one (e.g., moving the playhead from 08:16 marker to 05:37 marker). Multiple backward seeking events in a row in one video were only counted as one because it was very likely that participants were looking for a certain position in the lecture. Again, the average number of backward seeking events (i.e., the frequency of backward seeking events *per lecture watched*) was used. A slow watching event occurred when a participant changed the playing speed of a video and the current playing speed was slower than before it was changed. Multiple slow watching events in a row in one video were also only counted as one because it was very likely that participants were trying to fine-tune the playback speed. Again, the average number of slow watching events (i.e., the frequency of slow watching events *per lecture watched*) was used.

In the analysis, we used a non-parametric approach for the cognitive engagement variables (recoding them as categorical variables) for two reasons. First, a large percentage of participants engaged in zero or very few video interaction events, which made the distribution of cognitive engagement right-skewed. In addition, past research has found that cognitive engagement, measured by video interaction events, has a nonlinear relationship with perceived video difficulty (Li et al., 2015); thus, it is reasonable to hypothesize that cognitive engagement also has a nonlinear relationship with achievement (Indeed, our results using this non-parametric approach indicate that this relationship is non-linear). To ensure enough observations in every category of the cognitive engagement variables, the average number of pausing events ($M = 1.36$, $SD = 2.45$) was divided into five levels while the average number of backward seeking events ($M = 0.16$, $SD = 0.39$) and the average number of slow watching events ($M = 0.03$, $SD = 0.11$) were each divided into two levels (see Table 2).[2] For each type of video interaction event, participants who did not conduct that event at all in the first five weeks formed the reference groups.

Achievement. To better capture a wider range of self-defined goals for participants and account for potential differences in participants' intentions, we used two variables to measure achievement. First, like many previous studies involving both traditional education settings and MOOCs, we used course grade to measure participant achievement. However, course grade may not be an appropriate measurement of the achievement for auditors and disengagers, who may not intend to complete the course or earn a certificate. In order to better measure the achievement of these types of participants, we included overall lecture coverage throughout the whole course as a learning outcome, which has also been used in a number of MOOC studies (e.g., Evans, et al., 2016; Kizilcec & Schneider, 2015). For auditors, watching video lectures seems more aligned with their self-defined goals as compared to earning a certificate, since in the first half of the class they showed a strong interest in watching lecture videos and little interest in completing assignments. For disengagers, watching more lectures appears to be a much more feasible goal as compared to earning a certificate.

While these achievement measures might still not capture whether all participants are meeting their goals (for example, some participants may aim to only complete one section of the course), they provide relatively flexible definitions of achievement. All achievement variables were standardized in the regression analysis, which allowed for easier comparison of the magnitudes of the relationships between predictors and outcomes.

3.5. Regression analysis

To examine the relationships between engagement and achievement for each of the four subgroups, we conducted regression analyses using halfway-course engagement (i.e., engagement from week 1 to week 5) to predict course grade and overall lecture coverage for the four subgroups.

[2] Since participants in general have low levels of average backward seeking and average slow watching, these two measures are both re-coded as dummy variables with 1 indicating having ever conducted backward/slow watching and 0 indicating having never conducted backward/slow watching.

Table 2
Descriptive statistics for predictors and outcome variables.

Subgroups	Whole sample		All-rounders		Quiz-takers		Auditors		Disengagers	
	M	*SD*	*M*	SD	*M*	*SD*	*M*	*SD*	*M*	SD
Final grade	5.92	19.91	23.62	33.6	16.83	31.84	1.17	7.7	0.59	6.2
Overall engagement throughout the whole course										
Quiz coverage	0.18	0.3	0.62	0.31	0.52	0.29	0.05	0.11	0.03	0.1
Lecture coverage	0.13	0.24	0.54	0.31	0.10	0.13	0.32	0.25	0.03	0.08
Halfway-Course Engagement										
Quiz coverage	0.19	0.32	0.69	0.3	0.61	0.27	0.05	0.08	0.02	0.05
Lecture coverage	0.14	0.25	0.61	0.29	0.11	0.11	0.38	0.25	0.03	0.03
Average pausing	1.36	2.45	2.88	2.97	1.45	2.46	1.89	2.45	0.98	2.18
Average pausing = 0	0.42		0.01		0.32		0.03		0.55	
0 < Average pausing ≤ 1	0.24		0.23		0.28		0.42		0.22	
1 < Average pausing ≤ 2	0.14		0.26		0.18		0.26		0.1	
2 < Average pausing ≤ 3	0.07		0.17		0.09		0.11		0.05	
3 < Average pausing	0.13		0.33		0.13		0.18		0.08	
Average backward seeking	0.16	0.39	0.39	0.49	0.18	0.37	0.27	0.42	0.11	0.34
Average backward seeking = 0	0.68		0.1		0.59		0.27		0.85	
0 < Average backward seeking	0.32		0.9		0.41		0.73		0.15	
Average slow watching	0.03	0.11	0.07	0.15	0.04	0.12	0.06	0.13	0.02	0.1
Average slow watching = 0	0.88		0.64		0.85		0.72		0.94	
0 < Average slow watching	0.12		0.36		0.15		0.28		0.06	
N	71,266		10,226 (14%)		8792 (12%)		3607 (5%)		48,641 (68%)	

Note. The whole sample included disengagers, auditors, quiz-takers and all-rounders. All predictors were measured based on participants' cumulative behavior in the first five weeks. Overall lecture and quiz coverage, as learning outcomes, were measured based on participants' cumulative behavior in the ten weeks.

We conducted the analyses in two steps. We started by predicting course grade and overall lecture coverage using our measures of halfway-course cognitive and behavioral engagement (i.e., average pausing, average backward seeking, average slow watching, lecture coverage, and quiz coverage from week 1 to week 5) with course fixed effects.[3] This analysis measured the average relationship between engagement and achievement for all participants. We then tested if engagement differentially predicted achievement for participants in different subgroups by adding dummy variables for subgroup membership and interaction terms between subgroup membership and halfway-course engagement variables.[4] All-rounders were used as the reference group as their engagement pattern represented the learning pathway intended by the instructors. The second model thus assessed the relationship between engagement and achievement within each subgroup and allowed us to compare such relationships among subgroups.

We conducted joint F-tests to address concerns of type I error and examine whether, for a given engagement variable, there was evidence of significant differences between groups in the relationships between outcomes and the engagement variable. Specifically, we conducted F-tests for the joint significance of all the interaction terms between subgroup membership and each engagement variable. For instance, for halfway-course quiz coverage, the F-test examined the joint significance of the three interaction terms between subgroup membership (i.e., dummy variables for quiz-takers, auditors, and disengagers) and halfway-course quiz coverage, to test if there was strong evidence that the relationship between halfway-course quiz coverage and the outcome variable (e.g., course grade) differed among the four subgroups.

3.6. Robustness analysis

Appendix B presents robustness analysis for subgroup classification. The choice of which data to use for subgroup classification was based on the following two goals: (a) to classify participant subgroups as early as possible to allow for more timely diagnosis of at-risk participants in each subgroup; and (b) to classify participants as correctly as possible (as if all data from the full course were used). More specifically, we classified participants into subgroups using data from the beginning of the course up to the end of week X (where X = 1 through 9). We then compared the results from each of these analyses to the results using all data from the full course to determine the extent to which data from early weeks can be used to correctly detect participants in each subgroup. In doing so, we balanced earlier classification against how well the data were able to classify participants.

[3] Note that lecture coverage measured at different time points (i.e., in the first five weeks and at the end of the course) was used as both the independent and dependent variables in our estimation models. This allowed us to examine the relationship between cognitive engagement and overall lecture coverage controlling for halfway-course behavioral engagement. That is, we compared this relationship only within participants who watched the same number of lectures in first half of the course.

[4] The subgroup membership dummy variables are highly correlated with the behavioral engagement variables and may attenuate the relationships between engagement and achievement overall. However, as we are focusing on the relationship between engagement and achievement within groups, these correlations should not affect our primary results.

We used two commonly used statistics in classification research, precision and sensitivity, to measure the performance of subgroup classification (Attewell, Monaghan, & Kwong, 2015, p. 52). Precision is defined as the percentage of participants who were classified in a subgroup who were indeed in that subgroup, where correct group membership is defined as the classification based on the full ten weeks of data (e.g., $\frac{N_{Auditor\_week1 \cap Auditor\_week10}}{N_{Auditor\_week1}}$). Sensitivity refers to the percentage of participants in a given group who were correctly classified as such (e.g., $\frac{N_{Auditor\_week1 \cap Auditor\_week10}}{N_{Auditor\_week10}}$).

Results show that later weeks provided more accurate classifications in terms of precision and sensitivity (see Fig. B1). In this study, we chose to use data through week 5 to because it allowed us to correctly classify more than 70% of participants in every subgroup. However, the question remains regarding to what extent the results from our regression analysis are similar using data from before week 5. We tested this by using cumulative data from week 1 to week 4 to measure engagement, generate participant subgroups, and examine the outcomes of interest.

Moreover, to assess if the results were sensitive to our subgroup definitions and choice of cutoffs, we also tested a range of cutoffs for the three indicators (see Table B1 for a summary of the cutoffs). For l_0 and q_0, we tested cutoffs that were half a standard deviation or one standard deviation above the course average. In addition to using average lecture and quiz coverage to create relative cutoffs, we also examined using absolute numbers, 10 lectures and 6 quizzes, as cutoffs. Using absolute, rather than relative, numbers ensures consistency across courses. These two numbers were based on previous empirical work in MOOCs that examined the average number of lectures watched and quizzes attempted across a wide variety of courses (Crossley et al., 2016; Evans et al., 2016). We also tested stricter cutoffs for auditors and quiz-takers with the third indicator, quiz-lecture ratio (r_0 and r_1 set to be 0.25 and 4) and a more restricted sample of all-rounders, setting the two cutoffs of quiz-lecture ratio to be 0.75 and 1.5. These stricter cutoffs allowed us to examine a similar and wider range of cutoffs for quiz-lecture ratio as compared to prior research (Anderson et al., 2014).

4. Results

4.1. Subgroups classification

Using the classification rules, we were able to classify more than 99% of the active participants in each of the three courses into one of the four participant subgroups. Because the classification rules were not exhaustive, in each course there was a small proportion of participants, ranging from 0.1% to 0.4%, who were not classified. In general, the fractions of subgroups in the participant population were similar to that in prior studies in which the same subgroups were examined (e.g., Anderson et al., 2014). In Table 2, we provide profiles of participant behavioral features and performance for each subgroup (see Table A1 for more details about subgroup profiles in each course). We compare the behaviors and performance of the subgroups in the following section.

All-rounders. All-rounders accounted for 12.7%–15.4% of the total population in each course (shown in Table A1). Since participants' behavioral engagement in the first five weeks was used in the subgroup classification, it is unsurprising to find that all-rounders watched the most lectures and submitted the most quizzes during that period and in the full ten weeks. All-rounders also had the highest cognitive engagement, as measured by video interactions, in the first half of the courses. They interacted with the lectures more frequently by pausing, seeking backward, and slowing down the lectures than the other three subgroups. In terms of performance, they had the highest course grade among the four subgroups.

Quiz-takers. Quiz-takers made up 9.7%–15.7% of the total population in each course. They interacted more with the quizzes than with the lectures both in the first half of the courses and throughout the whole courses. Throughout the whole course, they submitted an average of 52% of the quizzes while watching only around 10% of the lectures. In addition, they interacted with the lectures less often than all-rounders and auditors. Although they watched very few lectures, quiz-takers had much higher course grades than auditors and disengagers. Since quiz scores counted for a large part of course grade in the sample courses, the relatively high course grade of quiz-takers may suggest that quizzes in these courses were so easy that participants could complete without watching lectures or that quiz-takers could achieve high quiz scores by simply revising and resubmitting their answers. However, it may also suggest that quiz-takers either had prior knowledge about the content or that they were learning the same content somewhere else. The large number of quiz-takers, combined with their relatively high performance, may indicate that failing to watch many lectures does not necessarily indicate being disengaged or at risk of low performance. Accordingly, it may be inappropriate to only use overall lecture coverage to measure the achievement of all participants.

Auditors. Auditors made up the smallest group in all the three courses (3.8%–6.2%). Overall, auditors had high engagement with the lectures and low engagement with the quizzes. In the first half of the courses, as well as throughout the whole courses, they watched more than 30% of the available lectures (many more than the quiz-takers and disengagers) while submitting around 5% of the available quizzes (many fewer than quiz-takers and all-rounders). In addition, they interacted with the lectures more often than quiz-takers and disengagers. Unsurprisingly, auditors had low performance in terms of course grades. This indicates that a small but significant proportion of participants in these courses were engaged mainly with lectures, which is in line with previous studies (e.g., DeBoer et al., 2014).

Disengagers. Disengagers made up the largest group in all the courses (65.4%–69.1%). Overall, disengagers had the lowest behavioral and cognitive engagement and lowest performance among all the subgroups.

4.2. Course grade as a learning outcome

Table 3 presents the results of regressions predicting course grade. Model 1 demonstrates that halfway-course quiz and lecture coverage were positively related to course grade. Conversely, all video interaction events (pausing, backward seeking, and slow

Table 3
Regression of halfway-course engagement on course grades.

	Model 1		Model 2	
	B	SE	B	SE
Halfway-Course Engagement				
Quiz coverage	0.572***	0.004	0.389***	0.016
Lecture coverage	0.157***	0.005	0.546***	0.012
0 < average pausing ≤ 1	−0.076***	0.008	0.112	0.099
1 < average pausing ≤ 2	−0.127***	0.010	0.009	0.100
2 < average pausing ≤ 3	−0.180***	0.013	−0.109	0.100
3 < average pausing	−0.218***	0.012	−0.211*	0.100
0 < Average backward seeking	−0.161***	0.009	−0.145***	0.026
0 < Average slow watching	−0.036***	0.010	0.071***	0.015
Halfway-Course Engagement × Quiz-takers				
Quiz coverage × Quiz-takers			0.761***	0.018
Lecture coverage × Quiz-takers			−0.743***	0.028
(0 < average pausing ≤ 1) × Quiz-takers			−0.113	0.102
(1 < average pausing ≤ 2) × Quiz-takers			−0.073	0.104
(2 < average pausing ≤ 3) × Quiz-takers			0.021	0.106
(3 < average pausing) × Quiz-takers			0.186	0.105
(0 < Average backward seeking) × Quiz-takers			0.121***	0.034
(0 < Average slow watching) × Quiz-takers			−0.142***	0.028
Halfway-Course Engagement × Auditors				X
Halfway-Course Engagement × Disengagers				X
Subgroup fixed effects				X
Course fixed effects			X	X
N			71,266	71,266
R^2			0.397	0.455

Note. All predictors (i.e., halfway-course engagement) were measured based on participants' cumulative behavior in the first five weeks. For each type of video interaction event, participants who did not conduct that event were used as the reference group. Model 2 tested the interaction between subgroup membership and halfway-course engagement. In Model 2, all-rounders and quiz-takers are the focal groups. However, auditors and disengagers are still kept in the model to reduce standard errors, results of auditors and disengagers were omitted from the table. All-rounders were used as the reference group. All coefficients are standardized across courses. *$p < .05$, **$p < .01$, ***$p < .001$.

watching) were negative predictors of course grade. In Model 2 we added dummy variables indicating subgroup membership (with all-rounders omitted as the reference category) and interaction terms between these subgroup dummies and our engagement predictors. For each engagement predictor, we conducted a joint F tests for this set of three interaction terms to determine if they were jointly significant. Results revealed that the relationship between engagement and course grade differed between these four subgroups for halfway-course quiz coverage, F (3, 71,228) = 902.89, $p < .001$, halfway-course lecture coverage, F (3, 71,228) = 415.81, $p < .001$, backward seeking, F (3, 71,228) = 8.21, $p < .001$, and slow watching, F (3, 71,228) = 9.74, $p < .001$, though not for pausing. We explore these relationships in more depth below.

As discussed above, course grade may not be an appropriate or meaningful measurement of achievement for some groups of participants (disengagers and auditors), so our presentation of findings for this outcome focuses only on all-rounders (the reference group) and quiz-takers.[5] For all-rounders, the reference group, the coefficient on each of the engagement variables represents the estimated relationship between engagement and course score. For quiz takers, the relationship between each engagement variable and course grade was calculated by the sum of the coefficient on the engagement variable and the coefficient on the interaction term corresponding to the same engagement variable.

Behavioral engagement. The comparison between all-rounders and quiz-takers revealed that behavioral engagement predicted course grade differently for these two groups (see Table 3, Model 2). First, the link between halfway-course quiz coverage and course grade was positive for both all-rounders ($\beta = 0.389$, $p < .001$) and quiz-takers ($\beta = 1.150$, $p < .001$). A one standard deviation increase in halfway-course quiz coverage was associated with a larger increase in course grade for quiz-takers than for all-rounders (as shown by the positive interaction term between early quiz coverage and the dummy variable of quiz-taker ($\beta = 0.761$, $p < .001$), see Table 3, Model 2). Second, unlike halfway-course quiz coverage, halfway-course lecture coverage predicted course grade oppositely for all-rounders and quiz-takers. While there was a positive relationship between halfway-course lecture coverage and course grade for all-rounders ($\beta = .546$, $p < .001$), there was a negative relationship for quiz-takers ($\beta = -0.197$, $p < .001$).

The *negative* relationship between halfway-course lecture coverage and course grade for quiz-takers is not surprising if the characteristics of this subgroup are taken into account. As proposed by previous studies, quiz-takers tend to be participants who are

[5] Results of the relationships between engagement and course grade for auditors and disengagers are presented in Fig. C1. For both auditors and disengagers, there was a positive relationship between halfway-course quiz coverage and course grade. For both groups, coefficients on other engagement behaviors were, in general, small and insignificant.

already familiar with the content and feel there is no need to watch all the lectures (Anderson et al., 2014). Accordingly, among quiz-takers, participants who watch more lectures may be those who are relatively less familiar with the content or have more trouble understanding or remembering the content.

Cognitive engagement. No significant differences were found between all-rounders and quiz-takers in terms of the relationship between pausing and course grade. However, there were significant differences between these two groups regarding how backward seeking and slow watching predicted course grade, indicating that subgroup membership may moderate the relationship between cognitive engagement variables and achievement (see Table 3, Model 2). First, while the relationship between backward seeking and course grade was negative and significant for all-rounders ($\beta = -.145$, $p < .001$), it was insignificant for quiz-takers ($\beta = -0.024$, $p = .278$). The difference between the two coefficients was significant (see Table 3, Model 2). Second, slow watching predicted course grades reversely for all-rounders and quiz-takers. While slow watching was associated with higher course grade for all-rounders ($\beta = 0.071$, $p < .001$), it was associated with lower course grade for quiz-takers ($\beta = -0.071$, $p < .01$).

4.3. Overall lecture coverage as a learning outcome

Table 4 presents the regression analysis with overall lecture coverage as the dependent variable. Model 1 examines the average relationship between engagement and overall lecture coverage for all participants. Again early quiz and lecture coverage were positively related to the outcome, and all of the video interaction events except slow watching were significant negative predictors of overall lecture coverage (see Table 4, Model 1).

In Model 2 we examine if the relationships between engagement and overall lecture coverage differed across subgroups by adding dummy variables indicating subgroup membership and interaction terms between subgroup membership and engagement predictors. Joint F-tests on the sets of interaction terms between subgroup membership and engagement predictors revealed differences between groups for halfway-course quiz coverage, F (3, 71,228) = 31.75, $p < .001$, pausing (statistical details provided in the footnote below), and slow watching, F (3, 71,228) = 2.44, $p < .1$, though not for backward seeking, F (3, 71,228) = 1.03, $p = .376$.[6] We next examine for which groups there are meaningful differences. We compare the results of all-rounders, auditors, and disengagers (with all-rounders as the reference group). We do not compare these relationships for quiz-takers, as measuring participant achievement using overall lecture coverage is less appropriate for quiz-takers, who seem to be more interested in testing and applying knowledge they already know than in learning new concepts.[7]

Behavioral engagement. We first focus on the estimated relationships between early quiz coverage and overall lecture coverage for our three subgroups of interest (see Table 4, Model 2). We included early lecture coverage as a control variable thought we do not report the results here given the construct overlap between the dependent and independent measures. In line with prior research, we found that early quiz coverage was predictive of higher overall lecture coverage for all-rounders ($\beta = .017$, $p < .05$), though the standardized coefficient was very small. On the contrary, it was predictive of lower overall lecture coverage for disengagers ($\beta = -0.074$, $p < .001$). This could indicate that among disengagers, those who were heavily engaged at the beginning more quickly realized that the course was not appropriate for them (though again we note that the coefficient was very small). There was no significant relationship between early quiz coverage and overall lecture coverage for auditors ($\beta = -.009$, $p = .741$).

Cognitive engagement. Pausing and slow watching predicted overall lecture coverage differently among all-rounders, auditors, and disengagers (see Table 4, Model 2). First, the relationship between pausing and overall lecture coverage was significant and negative for disengagers, while it was insignificant and positive for all-rounders. For slow watching, there was a significant and positive relationship for all-rounders ($\beta = .02$, $p < .01$), and an insignificant and negative relationship for auditors ($\beta = -0.021$, $p = .158$).

4.4. Robustness analysis

We conducted robustness analysis to examine if these results were sensitive to the choices of which data (cumulative data from weeks 1–5) and which cutoffs to use for subgroup classification. Overall, the results were qualitatively similar (in magnitude and direction) to our main analysis; changing the data or cutoffs used did not result in meaningful differences. First, we found that the relationships between engagement and achievement based on cumulative data from week 2–4 were similar to those based on cumulative data from week 5. Figs B2 and B3 present the estimated relationships between the eight engagement variables and our two outcomes: course grade (Fig. B2) and overall lecture coverage (Fig. B3). Across nearly all of these relationships, the sign and significance were the same when using cumulative data through weeks 2, 3, 4 or 5. Estimated relationships using data from only week 1 were significantly different in a number of cases.

Moreover, the results from regression analyses using subgroups generated by a variety of cutoffs also aligned with the results using the cutoffs from our main analysis (see Tables B2 and B3). The direction of the associations between most of the engagement variables and the outcomes of interest remained consistent for all subgroups across all classification rules, although we found changes

[6] For average number of pausing between zero and one, $F(3, 71{,}228) = 12.39$, $p < .001$. For average number of pausing between one and two, $F(3, 71{,}228) = 8.12$, $p < .001$. For average number of pausing between two and three, $F(3, 71{,}228) = 3.43$, $p < .05$. For average number of pausing larger than three, $F(3, 71{,}228) = 4.05$, $p < .01$.

[7] Results of the relationships between engagement and lecture coverage for quiz-takers are presented in Fig. C2. While halfway-course quiz coverage was positively associated with overall lecture coverage, video interaction events did not predict overall lecture coverage.

Table 4
Regression of halfway-course engagement on overall lecture coverage.

	Model 1		Model 2	
	B	SE	B	SE
Halfway-Course Engagement				
Quiz coverage	0.017***	0.002	0.017*	0.008
Lecture coverage	0.941***	0.002	1.009***	0.006
0 < average pausing ≤ 1	−0.108***	0.004	0.067	0.052
1 < average pausing ≤ 2	−0.103***	0.005	0.068	0.052
2 < average pausing ≤ 3	−0.102***	0.007	0.074	0.052
3 < average pausing	−0.107***	0.006	0.062	0.052
0 < Average backward seeking	−0.035***	0.005	−0.010	0.014
0 < Average slow watching	−0.006	0.005	0.020**	0.008
Halfway-Course Engagement × Auditors				
Quiz coverage × Auditors			−0.026	0.030
Lecture coverage × Auditors			−0.102***	0.010
(0 < average pausing ≤ 1) × Auditors			−0.090	0.062
(1 < average pausing ≤ 2) × Auditors			−0.051	0.063
(2 < average pausing ≤ 3) × Auditors			−0.121	0.066
(3 < average pausing) × Auditors			−0.045	0.064
(0 < Average backward seeking) × Auditors			0.016	0.021
(0 < Average slow watching) × Auditors			−0.041*	0.017
Halfway-Course Engagement × Disengagers				
Quiz coverage × Disengager			−0.091***	0.016
Lecture coverage × Disengager			−0.404***	0.020
(0 < average pausing ≤ 1) × Disengager			−0.133*	0.052
(1 < average pausing ≤ 2) × Disengager			−0.126*	0.052
(2 < average pausing ≤ 3) × Disengager			−0.121*	0.053
(3 < average pausing) × Disengager			−0.116*	0.053
(0 < Average backward seeking) × Disengager			0.025	0.015
(0 < Average slow watching) × Disengager			−0.020	0.011
Halfway-Course Engagement × Quiz-takers				X
Subgroup fixed effects				X
Course fixed effects			X	X
N			71,266	71,266
R^2			0.850	0.853

Note. All predictors (i.e., halfway-course engagement) were measured based on participants' cumulative behavior in the first five weeks. For each type of video interaction event, participants who did not conduct that event were used as the reference group. Overall lecture coverage refers to lecture coverage throughout the whole course. Model 2 tested the interaction between subgroup membership and halfway-course engagement. All-rounders, auditors, disengagers are the focus group in model 2. All-rounders were used as the reference group. However, quiz-takers are still kept in the model to reduce standard error and the results of quiz-takers are omitted from the table. All coefficients are standardized across courses. *p < .05, **p < .01, ***p < .001.

in the statistical significance of some coefficients. We found that only one relationship (the relationship between early quiz coverage and overall lecture coverage for all-rounders) changed direction (see Table B3).

5. Discussion

5.1. Key findings

Better understanding how early engagement is related to achievement is important for effectively identifying at-risk participants and targeting simple and low-cost interventions. The overarching goal of this study is to unmask the heterogeneity in the relationships between engagement and achievement across defined participant subgroups in MOOCs. Unlike previous studies that examine the average relationships between engagement and achievement among all participants in a MOOC, this study used behavioral data to first define four participant subgroups in MOOCs—all-rounders, quiz-takers, auditors, and disengagers—and then to examine levels of behavioral and cognitive engagement within groups. We then tested if the relationships between behavioral and cognitive engagement and course outcomes varied by participant subgroup. For both behavioral and cognitive engagement, our data provides evidence that the relationships between engagement and achievement were different for various participant subgroups (see Table 5 for an overview of the findings). Two broad types of differences were identified.

First, we found that the same engagement variable may be oppositely associated with achievement for different subgroups for both behavioral and cognitive engagement. For instance, there was a large, negative relationship between early lecture coverage and course grade for quiz-takers as compared to a positive relationship for all-rounders. Khalil and Ebner (2017) found that quiz-takers were likely to be people who might be learning the same content somewhere else, such as university students, and thus planned to skip the content in the MOOC course with which they were already familiar. This suggests that quiz-takers who showed higher levels

Table 5
Summary of relationships between halfway-course engagement and achievement by subgroups.

Course grade	All-rounders	Quiz-takers	
	Coefficient	Coefficient	Differences
Halfway-course quiz coverage	(+)	(+)	***
Halfway-course lecture coverage	(+)	(−)	***
Pausing	(−)	(−)	
Backward seeking	(−)	–	***
Slow watching	(+)	(−)	***

Overall lecture coverage	All-rounders	Auditors		Disengagers	
	Coefficient	Coefficient	Differences	Coefficient	Differences
Halfway-course quiz coverage	(+)	–		(−)	***
Pausing	+			(−)	*
Seeking backward	–	+		(+)	
Slow watching	(+)	–	*	–	

Note. Each cell under the subgroups represents the direction of the association between the engagement variable and achievement for that subgroup. Results enclosed in parentheses are significant at 0.05 level. The difference columns indicate whether the association was significantly different between a subgroup and the reference subgroup, all-rounders. The associations between pausing and overall lecture coverage for auditors are not presented since there is not clear pattern in the relationships. *$p < .05$, **$p < .01$, ***$p < .001$.

of engagement may perceive higher task challenges in the quizzes and need to go back to watch lectures. We found similar phenomena for cognitive engagement. For example, slow watching was predictive of *higher* course grades for all-rounders and *lower* course grades for quiz-takers. These contradictory findings may be explained in part by how participants select into cognitive engagement. Past work on information processing and learning approaches has demonstrated the positive effects of cognitive engagement on achievement (e.g., Biggs, 1987; Craik & Lockhart, 1972; Craik & Tulving, 1975; Marton & Säljö, 1976; Zhang et al., 2006). However, when participants are actively making choices about their own learning, their engagement may vary based on their background characteristics and other environmental factors. Not controlling for the potential determinants of cognitive engagement that are negatively correlated with achievement may bias the estimated relationship between engagement and outcomes.

It is worth noting that this pattern is not found for pausing; pausing predicted lower course grades for both all-rounders and quiz-takers. These consistent negative relationships may indicate that, rather than measuring cognitive engagement, pausing is an indication of the increased cognitive load (Sweller, 1994; Van Merrienboer & Sweller, 2005). It may also suggest that pausing is not a conceptually clear measure of cognitive engagement since participants may pause for reasons unrelated to learning, such as to take a break to do something else.

The second type of difference is that some engagement measures predict achievement for one subgroup but not another. For example, quiz coverage from the first half of the course was the strongest predictor of quiz-takers' course grade and the relationship was stronger for quiz-takers than it was for all-rounders. Also, the relationship between backward seeking and course grade was relatively large and significant for all-rounders but small and insignificant for quiz-takers. This may suggest that while backward seeking is positively associated with investing mental effort and using cognitive strategies for all-rounders, it is not a valid measurement of cognitive engagement for quiz-takers, who interact with the course mainly by completing quizzes rather than watching lectures. Additionally, it could be that the cognitive engagement behaviors of quiz-takers are not captured by these commonly used measures of video interaction and that more investigation is needed to examine cognitive engagement involved in quiz completion for this subgroup. For instance, the behavior of re-attempting quizzes, which may require mental effort and may indicate that participants are cognitively engaged with the course, could be used to measure cognitive engagement of quiz-takers (Do, Chen, Brandman, & Koller, 2013).

5.2. Implications

These findings make several contributions to the research and practice around engagement and achievement in the context of MOOCs. First, this study highlights the complexity of applying theories of engagement to interpret participants' behaviors and to predict their future outcomes. In particular, we provide evidence that the relationship between engagement and achievement in a class may be determined both by the effects of engagement and the reasons why participants decide to engage. Though largely ignored in previous research in online learning, this selection bias may play an important role in contexts with diverse participant populations, such as MOOCs. As researchers and practitioners increasingly gain access to rich behavioral data from online environments (that is often not paired with demographic or self-reported data), we must pay attention to the diversity in these environments and the different ways in which behaviors predict outcomes.

Second, we identified negative relationships between engagement and achievement, which contradict results from prior research in MOOCs (e.g., Balakrishnan & Coetzee, 2013; Brinton et al., 2015; Crossley et al., 2016; Sinha et al., 2014; Wieling & Hofman, 2010; Williams, Birch, & Hancock, 2012). We note that our study differs from these studies in two important ways: prior studies either only examined a relatively homogeneous participant population, such as college students in blended courses (e.g., Wieling & Hofman, 2010; Williams et al., 2012) or did not test the relationships between engagement and course performance for different subgroups of participants in MOOCs (e.g., Balakrishnan & Coetzee, 2013; Brinton et al., 2015). Our results suggest that examinations of participant behaviors and outcomes based on a whole class of MOOC participants may have masked subgroup variation. Ignoring subgroup differences in MOOCs could lead instructors to fail to identify at-risk participants in certain subgroups (DeBoer et al., 2014).

Finally, we found evidence that subgroups in MOOCs may engage differently and that certain engagement behaviors may be differentially important across groups. One important question raised by these results is how to define engagement in MOOCs. Current work in MOOCs defines engagement and achievement using the traditional conceptions that are aligned with institutional goals (e.g., Brinton et al., 2015; Crossley et al., 2016; Taylor, Veeramachaneni, & O'Reilly, 2014), while, as proposed by DeBoer et al. (2014), engagement in MOOCs should be conceptualized in a way that is more aligned with the learning pathways intended by participants. Although no causal inferences can be made from this study, it is possible that interventions targeting certain behaviors in MOOCs may work better for some subgroups than others. For example, interventions aiming at improving participant video watching, such as encouraging participants to watch the first video as early as possible, may be more effective for all-rounders and auditors than for quiz-takers (e.g., Baker et al., 2016). When designing, applying, and analyzing the results of interventions, it is important for practitioners and researchers to keep in mind which subgroups they are targeting and how large that subgroup is in the course.

5.3. Limitations and future research

There are a few limitations to our study. First, although previous research provides evidence that justifies our approaches to measuring cognitive engagement, the majority of these studies do not explicitly examine the actual cognitive processes that underlie video interaction events. By using behavioral data to infer cognitive engagement, we run the risk of picking up different signals than we intend. Such interactions might not always be evidence of cognitive engagement. Future research using self-reports or observations is needed to provide direct evidence of the connections between video interaction events and cognitive engagement.

Second, the courses that we focused on were unique in that they were relatively long courses in which participants were engaged with a large number of lectures and quizzes. The benefits of using courses like these to classify participant subgroups certainly come with disadvantages. It is unclear whether these results would extend to MOOCs with very different course designs. Most courses that have this type of design are in STEM fields because the content more readily lends itself to frequent, small assessments. These results might not extend to other disciplines. Moreover, we are only able to examine the relationships between certain kinds of engagement and course outcomes. By including MOOCs with a greater variety of course features (such as MOOCs where discussion forum participation is promoted), we could examine a broader set of various kinds of engagement (e.g., Do et al., 2013; Wang et al., 2015).

Finally, the conceptualization and operationalization of achievement in MOOCs should depend on many factors, including the unique characteristics of learning contexts, the goal of the providers, and the goals of the participants. One approach is to conceptualize achievement as the attainment of participants' self-defined goals. However, adopting such a conceptualization makes researching MOOCs particularly challenging. Due to lack of self-reported data on participants' intentions, there may be a mismatch between achievement measures and participants' actual goals, which may lead to under- or over-identification of at-risk participants. For instance, a participant classified as disengager may have initially had the intention to engage in and complete the course. Including only lecture coverage, and not course grade, might not actually match the participant's goals. To address this concern, we also provide the results of how engagement predicts course grade for disengagers and auditors in Appendix C. Moreover, although watching videos is a more realistic outcome than passing a course for disengagers, some disengagers may have not even intended to watch any videos. Interventions targeting these participants, who do not intend to engage with the course beyond potentially perusing the syllabus and list of topics, may have limited effects. For that reason, researchers and practitioners should always consider the cost and potential benefits of specific interventions. Future studies are needed to examine how self-report surveys or interviews can provide a more solid understanding of the goals and backgrounds of participants in each subgroup and to explicitly examine the relationship between outcome measures (e.g., lecture coverage and course grade) and self-defined goals across subgroups.

6. Conclusion

This study contributes to a more nuanced and complex understanding of how engagement predicts achievement in MOOCs, in which participant populations can be extremely diverse. Empirical evidence from this study supports the idea that early engagement is predictive of achievement. However, the relationship between early engagement and achievement varies by participant subgroups. These results highlight the importance of examining subgroup differences to improve the effectiveness of the identification of at-risk participants. Furthermore, this study offers insights and recommendations about how engagement and achievement in MOOCs can be reconceptualized in a way that aligns with participant self-defined learning pathways.

Acknowledgments

This work was supported by a grant from the National Science Foundation, USA (DUE #1535300).

Appendix A. Subgroup classification

Table A1
Descriptive Statistics for Predictors and Outcome Variables by Courses and Subgroups

	Algebra				Pre-calculus 1				Pre-calculus 2			
	All-rounder	Quiz-taker	Auditor	Disengager	All-rounder	Quiz-taker	Auditor	Disengager	All-rounder	Quiz-taker	Auditor	Disengager
Final grade	23.01	17.21	0.56	0.64	25.35	19.81	2.39	0.61	22.73	13.57	0.77	0.53
	(32.76)	(32.07)	(5.04)	(6.61)	(35.18)	(34.53)	(12.25)	(6.33)	(32.98)	(28.51)	(4.72)	(5.75)
Overall engagement throughout the whole course												
Quiz coverage	0.61	0.5	0.05	0.04	0.72	0.65	0.07	0.04	0.53	0.43	0.04	0.02
	(0.32)	(0.28)	(0.10)	(0.10)	(0.26)	(0.26)	(0.15)	(0.11)	(0.31)	(0.28)	(0.09)	(0.09)
Lecture coverage	0.53	0.1	0.29	0.04	0.58	0.11	0.39	0.04	0.51	0.09	0.29	0.03
	(0.32)	(0.13)	(0.24)	(0.08)	(0.29)	(0.14)	(0.25)	(0.08)	(0.32)	(0.12)	(0.24)	(0.09)
Halfway-Course Engagement												
Quiz coverage	0.71	0.62	0.05	0.04	0.77	0.71	0.05	0.02	0.59	0.5	0.04	0.01
	(0.29)	(0.25)	(0.09)	(0.06)	(0.26)	(0.26)	(0.08)	(0.04)	(0.31)	(0.28)	(0.08)	(0.03)
Lecture coverage	0.62	0.11	0.38	0.04	0.64	0.12	0.44	0.03	0.56	0.09	0.34	0.02
	(0.30)	(0.11)	(0.26)	(0.04)	(0.26)	(0.13)	(0.25)	(0.03)	(0.30)	(0.10)	(0.24)	(0.02)
Average pausing	3.27	1.63	2.05	1.32	2.63	1.2	1.66	1.06	2.69	1.41	1.94	0.65
	(3.39)	(2.75)	(2.78)	(2.58)	(2.72)	(2.00)	(2.15)	(2.06)	(2.62)	(2.37)	(2.43)	(1.83)
Average backward seeking	0.39	0.19	0.26	0.14	0.36	0.15	0.24	0.11	0.41	0.19	0.3	0.08
	(0.52)	(0.40)	(0.44)	(0.40)	(0.46)	(0.33)	(0.37)	(0.33)	(0.50)	(0.38)	(0.43)	(0.30)
Average slow watching	0.05	0.03	0.04	0.02	0.06	0.04	0.05	0.02	0.1	0.06	0.07	0.02
	(0.13)	(0.11)	(0.11)	(0.10)	(0.14)	(0.11)	(0.13)	(0.10)	(0.17)	(0.14)	(0.14)	(0.10)
N	3732	3801	923	15,695	3074	2367	1005	13,774	3420	2624	1679	19,172

Appendix B. Robustness checks

Table B1
Different Cutoffs of Lecture Watching and Quiz Submission Used in the Main Analysis and Robustness Checks

	Main analysis	Robustness check		
	Course average	Course average + 0.5SD	Course average + 1SD	Absolute number
Halfway-course lecture coverage				
Algebra	0.15	0.28	0.41	10 lectures
Pre-calculus 1	0.15	0.28	0.41	10 lectures
Pre-calculus 2	0.11	0.23	0.34	10 lectures
Halfway-course quiz coverage				
Algebra	0.24	0.40	0.57	6 quizzes
Pre-calculus 1	0.22	0.39	0.57	6 quizzes
Pre-calculus 2	0.13	0.26	0.40	6 quizzes

Note. In addition to using lecture and quiz coverage to create relative cutoffs, we also examined using absolute number, 10 lectures and 6 quizzes, as cutoffs to differentiate disengagers from the other three subgroups.

Table B2
Robustness Check for Course Grade Using Different Cutoffs for Group Classification

Halfway-Course Engagement	Main analysis		Using different cutoffs for quiz-lecture ratio	
	All-rounder	Quiz-taker	All-rounder	Quiz-taker
Quiz coverage	0.389***	1.150***	0.172***	1.202***
Lecture coverage	0.546***	−0.197***	0.756***	−0.238***
0 < Average pause ≤ 1	0.112	−0.001	0.099	−0.015
1 < Average pause ≤ 2	0.009	−0.064*	−0.060	−0.122***
2 < Average pause ≤ 3	−0.109	−0.088*	−0.222	−0.028
3 < Average pause	−0.211*	−0.025	−0.333*	−0.049
0 < Average backward seeking	−0.145***	−0.024	−0.139***	−0.052
0 < Average slow watching	0.071***	−0.071**	0.089***	−0.172***

Halfway-Course Engagement	Using different cutoffs for lecture watching and quiz submission					
	Course average + 0.5SD		Course average + 1SD		10 lectures and 6 quizzes	
	All-rounder	Quiz-taker	All-rounder	Quiz-taker	All-rounder	Quiz-taker
Quiz coverage	0.595***	1.451***	0.777***	1.862***	0.368***	1.173***
Lecture coverage	0.565***	−0.176***	0.587***	−0.156**	0.547***	−0.179***
0 < Average pause ≤ 1	0.171	−0.002	0.076	−0.026	0.115	−0.004
1 < Average pause ≤ 2	0.056	−0.052	−0.067	−0.096**	0.016	−0.075**
2 < Average pause ≤ 3	−0.072	−0.033	−0.218	−0.075	−0.100	−0.090*
3 < Average pause	−0.202	0.059	−0.349**	0.054	−0.203*	−0.026
0 < Average backward seeking	−0.122***	0.008	−0.113**	0.005	−0.144***	−0.041
0 < Average slow watching	0.114***	−0.071**	0.131***	−0.057	0.070***	−0.086***

Note. We generated the subgroups with classification rules in addition to the ones used in the main analysis. Using those subgroups in the regression model, the direction of the associations between engagement variables and course grades remains consistent for both all-rounders and quiz-takers across classification rules, although we found changes in the statistical significance of some coefficients. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table B3
Robustness Check for Overall Lecture Coverage Using Different Cutoffs for Group Classification

Halfway-Course Engagement	Main analysis			Using different cutoffs for quiz-lecture ratio		
	All-rounder	Auditor	Disengager	All-rounder	Auditor	Disengager
Quiz coverage	0.017*	−0.009	−0.074***	−0.046**	−0.257***	−0.082***
Lecture coverage	1.009***	0.907***	0.605***	1.065***	0.913***	0.595***
0 < Average pause ≤ 1	0.067	−0.022	−0.066***	0.184**	−0.026	−0.066***
1 < Average pause ≤ 2	0.068	0.017	−0.058***	0.181*	0.026	−0.058***
2 < Average pause ≤ 3	0.074	−0.047	−0.047***	0.181*	−0.068	−0.047***
3 < Average pause	0.062	0.017	−0.054***	0.186**	0.008	−0.054***
0 < Average backward seeking	−0.010	0.007	0.015*	−0.024	0.018	0.016**
0 < Average slow watching	0.020*	−0.021	−0.001	0.021*	−0.041*	0.001

Halfway-Course Engagement	Using different cutoffs for lecture watching and quiz submission								
	Course average + 0.5SD			Course average + 1SD			10 lectures and 6 quizzes		
	All-rounder	Auditor	Disengager	All-rounder	Auditor	Disengager	All-rounder	Auditor	Disengager
Quiz coverage	0.061***	−0.009	−0.041***	0.097***	−0.01	−0.033***	0.013	−0.01	−0.082***
Lecture coverage	1.012***	0.918***	0.754***	1.019***	0.922***	0.805***	1.008***	0.908***	0.583***
0 < Average pause ≤ 1	0.152*	−0.084	−0.079***	0.140*	−0.078	−0.085***	0.066	−0.02	−0.063***
1 < Average pause ≤ 2	0.154*	−0.042	−0.071***	0.139	−0.04	−0.076**	0.069	0.019	−0.055***
2 < Average pause ≤ 3	0.156*	−0.086	−0.064***	0.132	−0.074	−0.068***	0.075	−0.043	−0.042***
3 < Average pause	0.146*	−0.011	−0.075***	0.137	0.029	−0.098***	0.062	0.02	−0.055***
0 < Average backward seeking	0.016	0.028	0.008	0.016	0.038	0.004	−0.011	0.005	0.017**
0 < Average slow watching	0.023**	−0.035	0.0003	0.030**	−0.025	−0.004	0.019*	−0.02	0.0003

Note. We generated the subgroups with classification rules in addition the ones used in the main analysis. Using those subgroups in the regression model, the direction of most associations between engagement and overall lecture coverage remains consistent across classification rules, although we found changes in significance of some coefficients. One large difference is that the small coefficient on halfway-course quiz coverage for all-rounders switched from significantly positive to significantly negative in one robustness check model. *p < .05, **p < .01, ***p < .001.

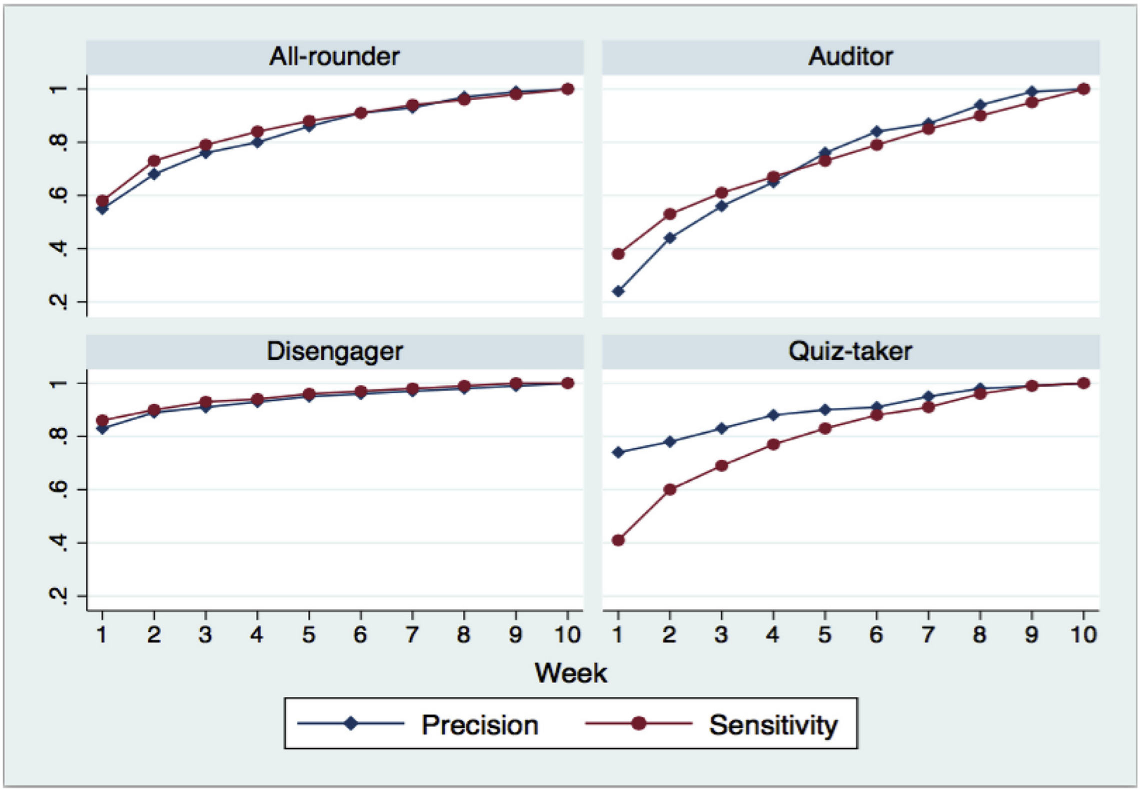

Fig. B1. Precision and sensitivity using cumulative data from weeks 1–10 to define subgroups. Notes: Precision is defined as the percentage of participants who were classified in a subgroup who were indeed in that subgroup, where correct group membership is defined as the classification based on the full ten weeks of data. Sensitivity refers to the percentage of participants in a given group who were correctly classified as such. Results show that later weeks provide more accurate classifications in terms of precision and sensitivity. Using data from the first five weeks allowed us to correctly classify more than 70% of the participants in each group.

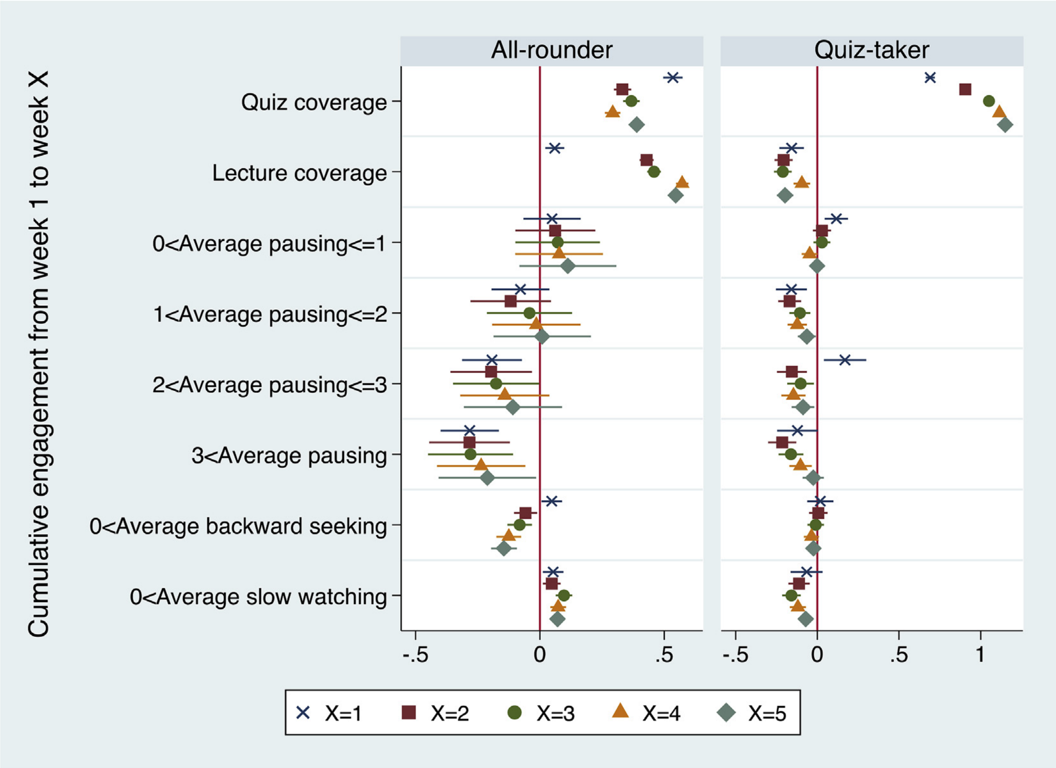

Fig. B2. The estimated relationships between cumulative engagement and course grade for all-rounders and quiz-takers using cumulative data from weeks 1–5. Notes: Regression analysis were conducted using cumulative data from weeks 1–5 to measure participant engagement, generate participant subgroups, and examine course grade. Results show that the estimated relationships were similar to those of week 5 (in magnitude and direction) if we were instead to use cumulative data from weeks 2–4.

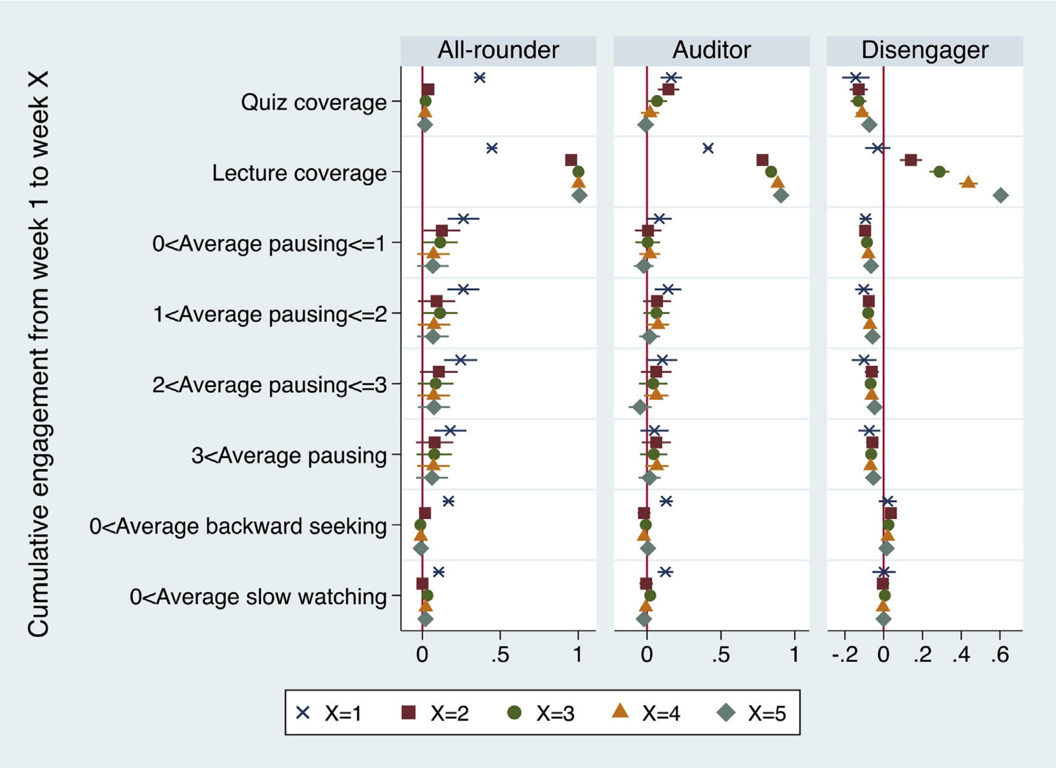

Fig. B3. The estimated relationships between cumulative engagement and overall lecture coverage for all-rounders, auditors, and disengagers using cumulative data from weeks 1–5. Notes: Regression analysis were conducted using cumulative data from weeks 1–5 to measure participant engagement, generate participant subgroups, and examine overall lecture coverage. Results show that the estimated relationships were similar to those of week 5 (in magnitude and direction) if we were instead to use cumulative data from weeks 2–4.

Appendix C. Additional Results

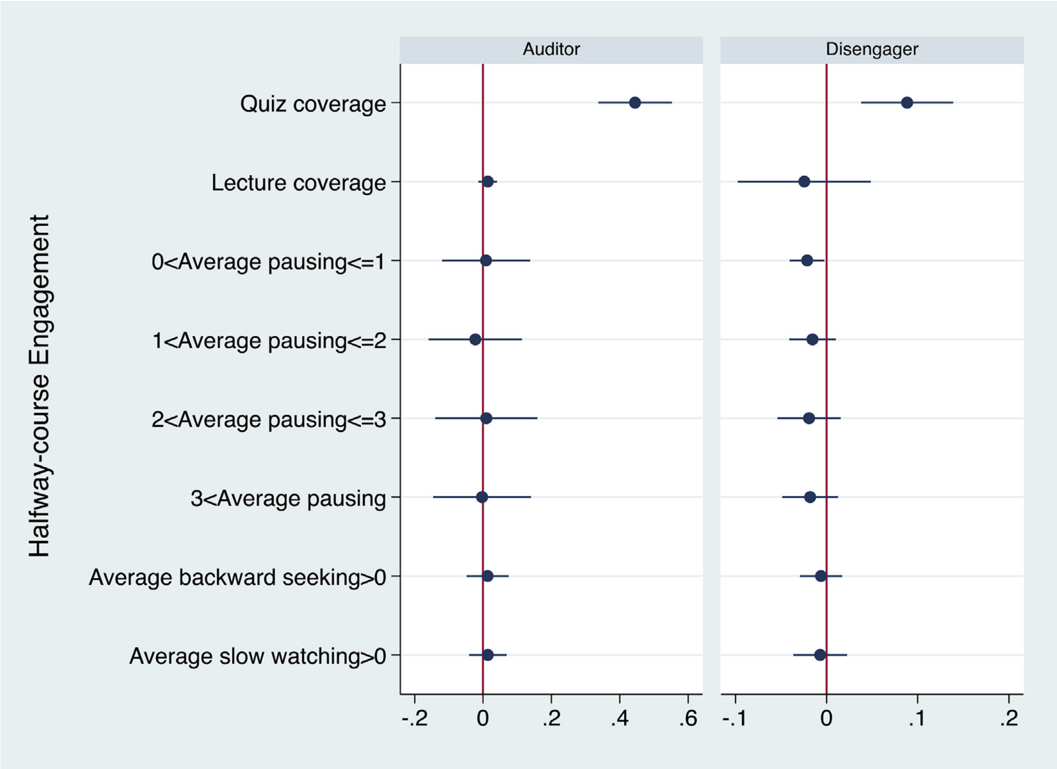

Fig. C1. The estimated relationships between halfway-course engagement and course grade for auditors and disengagers.

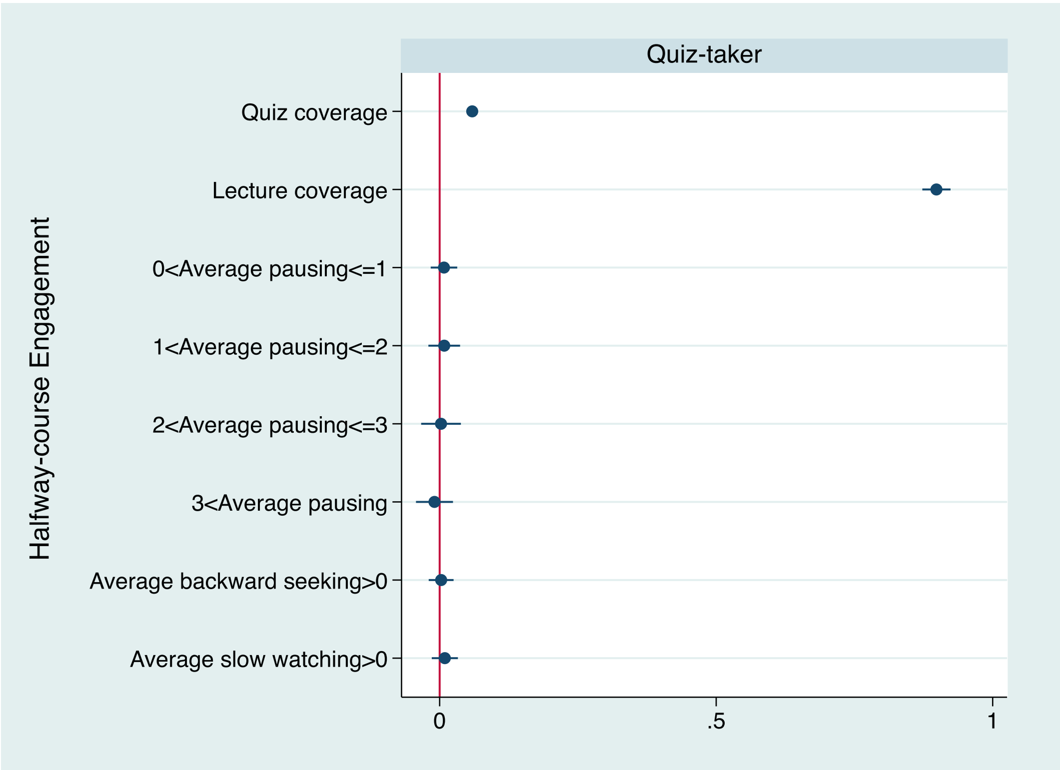

Fig. C2. The estimated relationships between halfway-course engagement and overall lecture coverage for quiz-takers.

References

Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2014). April). Engaging with massive online courses. *Proceedings of the 23rd international conference on World wide web* (pp. 687–698). Association for Computing Machinery. https://doi.org/10.1145/2566486.2568042.

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology, 44*(5), 427–445. https://doi.org/10.1016/j.jsp.2006.04.002.

Archambault, I., Janosz, M., Morizot, J., & Pagani, L. (2009). Adolescent behavioral, affective, and cognitive engagement in school: Relationship to dropout. *Journal of School Health, 79*(9), 408–415. https://doi.org/10.1111/j.1746-1561.2009.00428.x.

Arora, S., Goel, M., Sabitha, A. S., & Mehrotra, D. (2017). Learner groups in massive open online courses. *American Journal of Distance Education, 31*(2), 80–97. https://doi.org/10.1080/08923647.2017.1300461.

Attewell, P., Monaghan, D., & Kwong, D. (2015). *Data mining for the social sciences: An introduction.* University of California Press. Retrieved from http://www.jstor.org/stable/10.1525/j.ctt13x1gcg.

Baker, R., Evans, B., & Dee, T. (2016). A randomized experiment testing the efficacy of a scheduling nudge in a massive open online course (MOOC). *AERA Open, 2*(4), https://doi.org/10.1177/2332858416674007.

Balakrishnan, G., & Coetzee, D. (2013). *Predicting student retention in massive open online courses using hidden markov models.* Electrical Engineering and Computer Sciences University of California at Berkeley. Retrieved from https://www2.eecs.berkeley.edu/Pubs/TechRpts/2013/EECS-2013-109.pdf.

Barba, P. G., Kennedy, G. E., & Ainley, M. D. (2016). The role of students' motivation and participation in predicting performance in a MOOC Motivation and participation in MOOCs. *Journal of Computer Assisted Learning.* https://doi.org/10.1111/jcal.12130.

Baturay, M. H. (2015). An overview of the world of MOOCs. *Procedia-social and Behavioral Sciences, 174*, 427–433. https://doi.org/10.1016/j.sbspro.2015.01.685.

Biggs, J. (1987). *Student approaches to learning and studying.* Melbourne: Australian Council for Educational Research.

Biggs, J. B., & Rihn, B. A. (1984). The effects of intervention on deep and surface approaches to learning. In J. Kirby (Ed.). *Cognitive strategies and educational performance* (pp. 279–294). Orlando, FL: Academic Press.

Boyer, N. R., & Usinger, P. (2015). Tracking pathways to success: Triangulating learning success factors. *International Journal of Self-directed Learning®, 22*.

Brinton, C. G., Buccapatnam, S., Chiang, M., & Poor, H. V. (2015). *Mining MOOC clickstreams: On the relationship between learner video-watching behavior and performance.* Retrieved from http://arxiv.org/pdf/1503.06489.pdf.

Brunken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 53–61. https://doi.org/10.1207/S15326985EP3801_7.

Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in Higher Education, 47*(1), 1–32.

Chen, B., Haklev, S., Harrison, L., Najafi, H., & Rolheiser, C. (2015). April). How do MOOC learners' intentions relate to their behaviors and overall outcomes. *In Proceedings of the AERA annual meeting*.

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823.

Cocea, M., & Weibelzahl, S. (2011). Disengagement detection in online learning: Validation studies and perspectives. *IEEE Transactions on Learning Technologies, 4*(2), 114–124. https://doi.org/10.1109/TLT.2010.14.

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

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671–684.

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268.

Crossley, S., Paquette, L., Dascalu, M., McNamara, D. S., & Baker, R. S. (2016, April). Combining clickstream data with NLP tools to better understand MOOC completion. *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 6–14). Association for Computing Machinery. https://doi.org/10.1145/2883851.2883931.

Cutumisu, M., Blair, K. P., Chin, D. B., & Schwartz, D. L. (2015). Posterlet: A game-based assessment of children's choices to seek feedback and to revise. *Journal of Learning Analytics, 2*(1), 49–71. https://doi.org/10.18608/jla.2015.21.4.

Davis, D., Jivet, I., Kizilcec, R. F., Chen, G., Hauff, C., & Houben, G. J. (2017). March). Follow the successful crowd: Raising MOOC completion rates through social comparison at scale. *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 454–463). ACM.

DeBoer, J., Ho, A. D., Stump, G. S., & Breslow, L. (2014). *Changing "course" reconceptualizing educational variables for massive open online courses.* Educational Researcher doi: 0013189X14523038.

DeBoer, J., Stump, G. S., Seaton, D., & Breslow, L. (2013a). Diversity in MOOC students' backgrounds and behaviors in relationship to performance in 6.002 x. *Proceedings of the sixth learning international networks consortium conference: Vol. 4*, (pp. 16–19).

DeBoer, J., Stump, G. S., Seaton, D., Ho, A., Pritchard, D. E., & Breslow, L. (2013b). Bringing student backgrounds online: MOOC user demographics, site usage, and online learning. *Educational data mining 2013*.

Dillon, J., Bosch, N., Chetlur, M., Wanigasekara, N., Ambrose, G. A., Sengupta, B., et al. (2016). Student emotion, Co-occurrence, and dropout in a MOOC context. *Proceedings of the 9th international conference on educational data mining*.

Dissanayake, D., Perera, T., Elladeniya, C., Dissanayake, K., Herath, S., & Perera, I. (2018). Identifying the learning style of students in MOOCs using video interactions. *International Journal of Information and Education Technology, 8*(3).

Do, C. B., Chen, Z., Brandman, R., & Koller, D. (2013, September). Self-driven mastery in massive open online courses. *MOOCs forum*https://doi.org/10.1089/mooc.2013.0003.

Dominguez, M., Bernacki, M. L., & Uesbeck, P. M. (2016). Predicting STEM achievement with learning management system data: Prediction modeling and a test of an early warning system. *EDM* (pp. 589–590). .

Dupeyrat, C., & Mariné, C. (2005). Implicit theories of intelligence, goal orientation, cognitive engagement, and achievement: A test of Dweck's model with returning to school adults. *Contemporary Educational Psychology, 30*(1), 43–59. https://doi.org/10.1016/j.cedpsych.2004.01.007.

Eccles, J. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.). *Achievement and achievement motives* (pp. 75–146). San Francisco: Freeman.

Evans, B. J., Baker, R. B., & Dee, T. S. (2016). Persistence patterns in massive open online courses (MOOCs). *The Journal of Higher Education, 87*(2), 206–242. https://doi.org/10.1353/jhe.2016.0006.

Ferguson, R., & Clow, D. (2015, March). Examining engagement: Analysing learner subpopulations in massive open online courses (MOOCs). *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 51–58). ACM.

Finn, J. D. (1993). *School engagement and students at risk.* Washington, DC: National Center for Education Statistics.

Finn, J. D., Folger, J., & Cox, D. (1991). Measuring participation among elementary grade students. *Educational and Psychological Measurement, 51*(2), 393–402.

Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology, 82*(2), 221.

Fredricks, J. A., Blumenfeld, P. B., Friedel, J., & Paris, A. (2002, April). Increasing engagement in urban settings: An analysis of the influence of the social and academic context on student engagement. *Annual meeting of the American educational research association, new Orleans*.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059.

Gamino, J. F., Chapman, S. B., Hull, E. L., & Lyon, R. (2010). Effects of higher-order cognitive strategy training on gist-reasoning and fact-learning in adolescents. *Frontiers in Psychology, 1*, 188.

Giannakos, M. N., Chorianopoulos, K., & Chrisochoides, N. (2015). Making sense of video analytics: Lessons learned from clickstream interactions, attitudes, and learning outcome in a video-assisted course. *International Review of Research in Open and Distance Learning, 16*(1).

Grainger, B. (2013). *Massive open online course (MOOC) report 2013.* University of London.

Greene, B. A., Miller, R. B., Crowson, H. M., Duke, B. L., & Akey, K. L. (2004). Predicting high school students' cognitive engagement and achievement: Contributions of classroom perceptions and motivation. *Contemporary Educational Psychology, 29*(4), 462–482. https://doi.org/10.1016/j.cedpsych.2004.01.006.

Halawa, S., Greene, D., & Mitchell, J. (2014). Dropout prediction in MOOCs using learner activity features. *Proceedings of the European MOOC stakeholder summit (EMOOCS 2014), Lausanne, Switzerland*. Retrieved from http://www.stanford.edu/~halawa/cgi-bin/files/emoocs2014.pdf.

Hansen, J. D., & Reich, J. (2015). Democratizing education? Examining access and usage patterns in massive open online courses. *Science, 350*(6265), 1245–1248. https://doi.org/10.1126/science.aab3782.

He, J., Bailey, J., Rubinstein, B. I., & Zhang, R. (2015, February). Identifying at-risk students in massive open online courses. *Proceedings of the twenty-ninth AAAI conference on artificial intelligence* (pp. 1749–1755). . Retrieved from http://www.ruizhang.info/publications/AAAI2015-MOOC.pdf.

Helme, S., & Clarke, D. (2001). Identifying cognitive engagement in the mathematics classroom. *Mathematics Education Research Journal, 13*(2), 133–153.

Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education, 90*, 36–53. https://doi.org/10.1016/j.compedu.2015.09.005.

Hew, K. F. (2015). Student perceptions of peer versus instructor facilitation of asynchronous online discussions: Further findings from three cases. *Instructional Science, 43*(1), 19–38.

Ho, A. D., Chuang, I., Reich, J., Coleman, C. A., Whitehill, J., Northcutt, C. G., et al. (2015). *Harvardx and mitx: Two years of open online courses fall 2012-summer 2014. (HarvardX Working Paper No. 10)*. https://doi.org/10.2139/ssrn.2586847.

Jiang, S., Williams, A., Schenke, K., Warschauer, M., & O'dowd, D. (2014, July). Predicting MOOC performance with week 1 behavior. *Educational data mining 2014*.

Jordan, K. (2014). *MOOC completion rates: The data*. Retrieved from http://www.katyjordan.com/MOOCproject.html..

Kahan, T., Soffer, T., & Nachmias, R. (2017). Types of participant behavior in a massive open online course. *The International Review of Research in Open and Distributed Learning, 18*(6).

Karpicke, J. D., Butler, A. C., & Roediger, H. L., III (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory, 17*(4), 471–479.

Käser, T., Hallinen, N. R., & Schwartz, D. L. (2017). March). Modeling exploration strategies to predict student performance within a learning environment and beyond. *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 31–40). ACM.

Khalil, M., & Ebner, M. (2017). Clustering patterns of engagement in massive open online courses (MOOCs): The use of learning analytics to reveal student categories. *Journal of Computing in Higher Education, 29*(1), 114–132.

Khezrlou, S. (2011). Cognitive strategy training: Improving reading comprehension in the language classroom. *Journal of Teaching Language Skills, 30*(4), 77–98.

Kizilcec, R. F., & Halawa, S. (2015). March). Attrition and achievement gaps in online learning. *Proceedings of the second (2015) ACM conference on learning@ scale* (pp. 57–66). ACM. https://doi.org/10.1145/2724660.2724680.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013, April). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. *Proceedings of the third international conference on learning analytics and knowledge* (pp. 170–179). . https://doi.org/10.1145/2460296.2460330.

Kizilcec, R. F., & Schneider, E. (2015). Motivation as a lens to understand online learners: Toward data-driven design with the OLEI scale. *Transactions on Computer-human Interaction, 22*(2), 6. https://doi.org/10.1145/2699735.

Kovacs, G. (2016). April). Effects of in-video quizzes on MOOC lecture viewing. *Proceedings of the third conference on learning@ scale* (pp. 31–40). ACM.

Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., & Gonyea, R. M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *The Journal of Higher Education, 79*(5), 540–563.

Lee, O., & Anderson, C. W. (1993). Task engagement and conceptual change in middle school science classrooms. *American Educational Research Journal, 30*(3), 585–610.

Lee, O., & Brophy, J. (1996). Motivational patterns observed in six-grade science classrooms. *Journal of Research in Science Teaching. The Official Journal of the National Association for Research in Science Teaching, 33*(3), 303–318.

Li, Q., & Baker, R. (2016). Understanding engagement in MOOCs. In T. Barnes, M. Chi, & M. Feng (Eds.). *Proceedings of the 9th international conference on educational data mining (pp. 605-606). Raleigh, NC, USA*.

Li, N., Kidzinski, L., Jermann, P., & Dillenbourg, P. (2015). How do in-video interactions reflect perceived video difficulty? *Proceedings of the European MOOCs stakeholder summit 2015* (pp. 112–121). PAU Education. Retrieved from https://infoscience.epfl.ch/record/207968/files/emooc2015_howdodiff.pdf.

Linnenbrink, E. A., & Pintrich, P. R. (2003). The role of self-efficacy beliefs instudent engagement and learning in the classroom. *Reading & Writing Quarterly, 19*(2), 119–137. https://doi.org/10.1080/10573560308223.

Littlejohn, A., Hood, N., Milligan, C., & Mustain, P. (2016). Learning in MOOCs: Motivations and self-regulated learning in MOOCs. *The Internet and Higher Education, 29*, 40–48.

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education, 54*(2), 588–599. https://doi.org/10.1016/j.compedu.2009.09.008.

Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of massive open online courses (MOOCs). *Computers & Education, 80*, 77–83. https://doi.org/10.1016/j.compedu.2014.08.005.

Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I—outcome and process. *British Journal of Educational Psychology, 46*(1), 4–11.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*(2), 390.

Meece, J. L., Blumenfeld, P. C., & Hoyle, R. H. (1988). Students' goal orientations and cognitive engagement in classroom activities. *Journal of Educational Psychology, 80*(4), 514.

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science, 18*(3), 159–163.

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review, 15*(1), 174–179.

Montague, M., Krawec, J., Enders, C., & Dietz, S. (2014). The effects of cognitive strategy instruction on math problem solving of middle-school students of varying ability. *Journal of Educational Psychology, 106*(2), 469.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*(3), 309–326.

Morris, L. V., Finnegan, C., & Wu, S. S. (2005). Tracking student behavior, persistence, and achievement in online courses. *The Internet and Higher Education, 8*(3), 221–231. https://doi.org/10.1016/j.iheduc.2005.06.009.

Nawrot, I., & Doucet, A. (2014). April). Building engagement for MOOC students: Introducing support for time management on online learning platforms. *Proceedings of the 23rd international conference on world wide web* (pp. 1077–1082). ACM. https://doi.org/10.1145/2567948.2580054.

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. *Handbook of research on student engagement* (pp. 259–282). Springer US.

Pellas, N. (2014). The influence of computer self-efficacy, metacognitive self-regulation and self-esteem on student engagement in online learning programs: Evidence from the virtual world of Second Life. *Computers in Human Behavior, 35*, 157–170. https://doi.org/10.1016/j.chb.2014.02.048.

Peters, L., Shmerling, S., & Karren, R. (2011). Constructivist pedagogy in asynchronous online education: Examining proactive behavior and the impact on student engagement levels. *International Journal on E-Learning, 10*(3), 311–330.

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory Into Practice, 41*(4), 219–225.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33. https://doi.org/10.1037/0022-0663.82.1.33.

Pintrich, P. R., & Garcia, T. (1991). Student goal orientation and self-regulation in the college classroom. *Advances in Motivation and Achievement: Goals and Self-regulatory Processes, 7*, 371–402.

Pursel, B. K., Zhang, L., Jablokow, K. W., Choi, G. W., & Velegol, D. (2016). Understanding MOOC students: Motivations and behaviours indicative of MOOC completion. *Journal of Computer Assisted Learning, 32*(3), 202–217. https://doi.org/10.1111/jcal.12131.

Puzziferro, M. (2008). Online technologies self-efficacy and self-regulated learning as predictors of final grade and satisfaction in college-level online courses. *American Journal of Distance Education, 22*(2), 72–89.

Ramesh, A., Goldwasser, D., Huang, B., Daume, H., III, & Getoor, L. (2014). June). Learning latent engagement patterns of students in online courses. *Twenty-eighth AAAI conference on artificial intelligence*.

Ramsden, P., & Entwistle, N. J. (1981). Effects of academic departments on students' approaches to studying. *British Journal of Educational Psychology, 51*(3), 368–383.
Reilly, C. (2013). October). MOOCs deconstructed: Variables that affect MOOC success rates. *e-learn: World conference on e-learning in corporate, government, healthcare, and higher education* (pp. 1308–1338). Association for the Advancement of Computing in Education (AACE).
Rodrigues, R. L., Ramos, J. L., Silva, J. C. S., Gomes, A. S., de Souza, F. D. F., & Maciel, A. M. A. (2016, July). Discovering level of participation in MOOCs through clusters analysis. *Advanced learning technologies (ICALT), 2016 IEEE 16th international conference on* (pp. 232–233). IEEE.
Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Educational Researcher, 39*(5), 406–412.
Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020.
Scevak, J. J., & Moore, P. J. (1998). Levels of processing effects on learning from texts with maps. *Educational Psychology, 18*(2), 133–155.
Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and Instruction, 14*(3), 293–305. https://doi.org/10.1016/j.learninstruc.2004.06.005.
Shah, D. (2017). *By the numbers: MOOCS in 2017 - class central.* Retrieved from https://www.class-central.com/report/mooc-stats-2017/.
Sharma, K., Jermann, P., & Dillenbourg, P. (2015). *Identifying styles and paths toward success in MOOCs.* International Educational Data Mining Society.
Siemens, G. (2013). *Massive open online courses: Innovation in education. Open educational resources: Innovation, research and practice*, *Vol. 5*. Retrieved from http://s3.amazonaws.com/academia.edu.documents/31833499/pub_PS_OER-IRP_web.pdf?AWSAccessKeyId=AKIAJ56TQJRTWSMTNPEA&Expires=1472125327&Signature=QhjAkoHiVXiTTw%2FNg4S%2F7Qj%2FMz0%3D&response-content-disposition=inline%3B%20filename%3DOpen_Educational_Resources_Innovation_Re.pdf#page=31.
Sinha, T., Jermann, P., Li, N., & Dillenbourg, P. (2014). Your click decides your fate: Inferring information processing and attrition behavior from MOOC video clickstream interactions. *2014 empirical methods in natural Language processing workshop on modeling large scale social interaction in massively open online courses*.
Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology, 85*(4), 571. https://doi.org/10.1037/0022-0663.85.4.571.
Stiglitz, J. E. (2014, June). *Inequality is not inevitable.* The New York Times. Retrieved from http://opinionator.blogs.nytimes.com/author/joseph-e-stiglitz/?version=meter.
Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5.
Taylor, C., Veeramachaneni, K., & O'Reilly, U. M. (2014). *Likely to stop? predicting stopout in massive open online courses.* Retrieved from http://arxiv.org/abs/1408.3382.
Thompson, J. R., Klass, P. H., & Fulk, B. M. (2012). Comparing online and face-to-face presentation of course content in an introductory special education course. *Teacher Education and Special Education, 35*(3), 228–242.
Tinto, V. (1997). Classrooms as communities: Exploring the educational character of student persistence. *The Journal of Higher Education, 68*(6), 599–623. https://doi.org/10.1080/00221546.1997.11779003.
Van Merrienboer, J. J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17*(2), 147–177. https://doi.org/10.1007/s10648-005-3951-0.
Wang, X., Yang, D., Wen, M., Koedinger, K., & Rosé, C. P. (2015). *Investigating how student's cognitive behavior in MOOC discussion forums affect learning gains.* International Educational Data Mining Society.
Weinstein, C. E., Mayer, R. E., & Wittrock, M. C. (1986). *Handbook of research on teaching. Handbook of research on teaching*.
Wellborn, J. G., & Connell, J. P. (1987). *Manual for the Rochester assessment package for schools.* Rochester, NY: University of Rochester.
Wen, M., Yang, D., & Rosé, C. P. (2014). Sentiment analysis in MOOC discussion forums: What does it tell us? *Proceedings of the 7th international conference on educational data mining*.
Whitehill, J., Williams, J. J., Lopez, G., Coleman, C. A., & Reich, J. (2015). *Beyond prediction: First steps toward automatic intervention in MOOC student stopout.* Available at: SSRN 2611750.
Whitmer, J. C. (2013). *Logging on to improve achievement: Evaluating the relationship between use of the learning management system, student characteristics, and academic achievement in a hybrid large enrollment undergraduate course*Ph.D. thesis Davis: University of California. Retrieved June 13, 2018 from https://www.learntechlib.org/p/121047/.
Wieling, M. B., & Hofman, W. H. A. (2010). The impact of online video lecture recordings and automated feedback on student performance. *Computers & Education, 54*(4), 992–998. https://doi.org/10.1016/j.compedu.2009.10.002.
Williams, A., Birch, E., & Hancock, P. (2012). The impact of online lecture recordings on student performance. *Australasian Journal of Educational Technology, 28*(2), https://doi.org/10.14742/ajet.869.
Wuellner, M. (2015). Student success factors in two online introductory-level natural resource courses. *Natural Sciences Education, 44*(1), 51–59.
Yang, D., Sinha, T., Adamson, D., & Rose, C. P. (2013, December). Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses. *Proceedings of the 2013 NIPS Data-driven education workshop: Vol. 11*, (pp. 14–). .
Zahn, C., Barquero, B., & Schwan, S. (2004). Learning with hyperlinked videos—design criteria and efficient strategies for using audiovisual hypermedia. *Learning and Instruction, 14*(3), 275–291. https://doi.org/10.1016/j.learninstruc.2004.06.004.
Zhang, D., Allon, G., & Van Mieghem, J. (2015). *Does social interaction improve learning? Field evidence from massive open online education (Northwestern University Working Paper)*. Retrieved from http://www.kellogg.northwestern.edu/faculty/vanmieghem/htm/pubs/2015_MOOC_Zhang-Allon-VanMieghem_16Aug2015.pdf.
Zhang, H., Lin, L., Zhan, Y., & Ren, Y. (2016). The impact of teaching presence on online engagement behaviors. *Journal of Educational Computing Research, 54*(7), 887–900.
Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker, J. F. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & Management, 43*(1), 15–27. https://doi.org/10.1016/j.im.2005.01.004.
Zheng, S., Rosson, M. B., Shih, P. C., & Carroll, J. M. (2015). February). Understanding student motivation, behaviors and perceptions in MOOCs. *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing* (pp. 1882–1895). ACM.
Zimmerman, B. J., & Martínez-Pons, M. (1992). *Perceptions of efficacy and strategy use in the self-regulation of learning. Student perceptions in the classroom*.