The benefits and caveats of using clickstream data to understand student self-regulatory behaviors: opening the black box of learning processes

Rachel Baker[1*], Di Xu[1], Jihyun Park[1], Renzhe Yu[1], Qiujie Li[2], Bianca Cung[1], Christian Fischer[3], Fernando Rodriguez[1], Mark Warschauer[1] and Padhraic Smyth[1]

* Correspondence: rachelbb@uci.edu
[1]University of California, 2060 Education, UC, Irvine, Irvine, CA 92697, USA
Full list of author information is available at the end of the article

Abstract

Student clickstream data—time-stamped records of click events in online courses—can provide fine-grained information about student learning. Such data enable researchers and instructors to collect information at scale about how each student navigates through and interacts with online education resources, potentially enabling objective and rich insight into the learning experience beyond self-reports and intermittent assessments. Yet, analyses of these data often require advanced analytic techniques, as they only provide a partial and noisy record of students' actions. Consequently, these data are not always accessible or useful for course instructors and administrators. In this paper, we provide an overview of the use of clickstream data to define and identify behavioral patterns that are related to student learning outcomes. Through discussions of four studies, we provide examples of the complexities and particular considerations of using these data to examine student self-regulated learning.

Keywords: Online education, Clickstream data, Descriptive analysis, Intervention

Introduction

The ability to monitor progress toward a goal and manage one's own learning behaviors accordingly (e.g., spacing instead of cramming, scheduling work time in advance, working at more ideal times) is critical to learning success (Broadbent and Poon, 2015; Kizilcec, Pérez-Sanagustín, and Maldonado, 2017; Pintrich and De Groot, 1990). While these self-regulated learning (SRL) skills are fundamental to progress in any setting (e.g., traditional classrooms: Paul, Baker, and Cochran, 2012; Trueman and Hartley, 1996; van Den Hurk, 2006; as well as online classes: Elvers, Polzella, and Graetz, 2003; Goda et al. 2015; Michinov, Brunot, Le Bohec, Juhel, and Delaval, 2011), they are particularly critical to success in online, hybrid, and flipped courses, as these classes require a high degree of independence and autonomy (Bawa, 2016; Jaggars, 2011; Jenkins and Rodriguez, 2013; Park et al. 2018; Roll and Winne, 2015). Unlike face-to-face courses, in which students attend lectures at specific days and times, online courses

require that students plan when they will watch course lectures and work on assignments. As a result, understanding students' self-regulatory behaviors and identifying effective ways to scaffold these behaviors is imperative for improving online learning outcomes.

Improved technology and the increasing availability of data from digital learning management systems (LMSs) allow for unique opportunities to capture students' self-regulatory behaviors in these settings in ways that are both more timely and more objective than the methods employed in traditional classrooms. In traditional learning environments, SRL, including students' time management skills, is mainly measured by student retrospective self-report, which can neither capture how SRL unfolds nor provide timely measures to examine how SRL changes with environmental factors. These significant limitations can be partially addressed by student clickstream data that are automatically collected through an LMS. As time-stamped records of students' click events in the course, student clickstream data provide researchers, instructors, and administrators with fine-grained, time-variant information about learning. While these data only provide a partial and noisy record of a student's actions, they offer information at scale about how each student navigates through and interacts with online education resources, therefore promising more objective and richer insight into the learning experience. Clickstream data can also advance how we understand the relationship between self-regulated learning and student achievement by allowing researchers to examine how traditional self-report measures of students' self-regulated learning correspond to students' behavior and engagement with course materials (Li, Baker, and Warschauer, 2020; Rodriguez et al. 2019).

However, analyses of clickstream data often require relatively advanced analytic techniques and a deep and contextualized understanding of the structure of the data, as the data are often sequential, event-based, and bursty. While there is a growing volume of studies that use clickstream data to measure student self-regulatory behaviors, rarely do these studies provide a detailed discussion about the complexities of constructing behavioral measures, the importance of contextual factors required to interpret clickstream data in meaningful ways, and the many caveats associated with these data. As a result, despite the availability of clickstream data and recent advances in analyzing such data, their usage in education research and application to improving instructional design and student learning remains limited.

This paper is designed to help instructors, administrators, and institutional researchers understand the basic concepts of working with clickstream data and the promising ways in which such data can affect the instructional design and student learning. We first provide a brief review of recent literature that has used clickstream data to measure students' SRL in online environments. We then provide a synthesis of four of our own recent research studies that use clickstream data to examine student behaviors and outcomes in online classes. The studies we highlight incorporate student survey data, demographic data, and transcript data in addition to clickstream data. These studies thus allow us to illustrate the promises and potential challenges of working with clickstream data in authentic education settings. Unlike most empirical papers that focus on results and outcomes, the discussions in this paper focus on the process of using clickstream data to understand student learning processes.

Background and existing literature—using clickstream data to trace self-regulated learning

The role and measurement of self-regulated learning

Self-regulation is an overarching construct that captures how students direct and monitors their own learning processes and progress (Pintrich and De Groot, 1990; Pintrich, Smith, Garcia, and McKeachie, 1993). Specifically, SRL is defined as a process where students actively set goals and make plans for their learning, monitor their learning process, and adjust their study plans (Pintrich, 2004). Students with high self-regulatory skills can appropriately apply effective learning strategies to increase effectiveness based on their personal needs and the characteristics of the tasks and the environment (Pintrich, 2004). Due to the flexibility of the course schedule and the limited social interaction, online courses require students to take more responsibility to regulate their own learning. In contrast, in face-to-face classrooms, instructors and peers can monitor and guide student behavior.

Multiple approaches, such as self-report questionnaires, observation, and think-aloud protocols have been used to measure SRL, with self-report questionnaires being the most widely used (Schellings and Van Hout-Wolters, 2011; Winne, 2010). The Motivated Strategies for Learning Questionnaire (MSLQ) developed by Pintrich et al. (1993) is the most commonly adopted instrument for measuring SRL in both face-to-face and online courses (Broadbent and Poon, 2015; Duncan and McKeachie, 2005). MSLQ captures three sets of SRL skills: (1) the use of cognitive strategies, (2) the use of metacognitive strategies, and (3) the management of personal and environmental academic resources including time management, choice and control of the study environment, effort regulation, and help-seeking (Pintrich and De Groot, 1990; Pintrich et al. 1993). MSLQ instructs students to predict/recall the likelihood or frequency of conducting certain SRL behaviors in the future/past. For instance, before a course starts, student time management skills would be measured by several Likert scale statements capturing the extent to which students predict that they can make good use of their study time, spend enough time studying, keep up with the coursework, attend class regularly, and find time to review before an exam in the upcoming course. Extensive research has been conducted to explore the relationship between self-reported SRL skills and online performance. There is consistent evidence that student online performance is associated with self-reported SRL skills overall. The sub-skills of time management, effort regulation, and metacognition have also shown consistent relationships with performance in online classes, but findings are mixed regarding the relationships between other SRL sub-skills, such as the use of cognitive strategies, and performance (Broadbent and Poon, 2015).

While these findings provide suggestive evidence that SRL skills play an important role in the learning process, most previous studies have relied on student self-reported instruments to measure SRL skills and investigate the role of SRL in online learning. As we discuss in the "Using clickstream data to understand SRL" section, self-reported data may not be effective measures of SRL, as many individuals suffer from self-report bias and past memories are often insufficient for students to accurately recall past behavior or predict future events. Therefore, more timely and objective measures of student SRL skills are needed to more accurately capture student SRL skills.

In contrast to the consistent positive correlations between self-reported SRL skills and academic performance, there is less consistent evidence that SRL skills can be meaningfully altered to affect academic performance. Findings from previous interventions that have attempted to improve SRL skills, mainly concentrating on time management, have varied considerably. For instance, previous work that has attempted to support time management in online courses by providing more deadlines, by allowing students to set their own deadlines, or by suggesting that students schedule study time have found mixed results on the effects of these interventions on student performance. Studies examining the effects of externally and self-imposed interim deadlines on course grades have found positive (e.g., Ariely and Wertenbroch, 2002), negative (e.g., Burger, Charness, and Lynham, 2011), and null effects (e.g., Levy and Ramim, 2013). Studies examining the effects of encouraging students to plan when they will do work have also found a mix of positive (Baker, Evans, Li, and Cung, 2019), negative (Baker, Evans, and Dee, 2016), and null (Sitzmann and Johnson, 2012) effects on course and assignment grades.
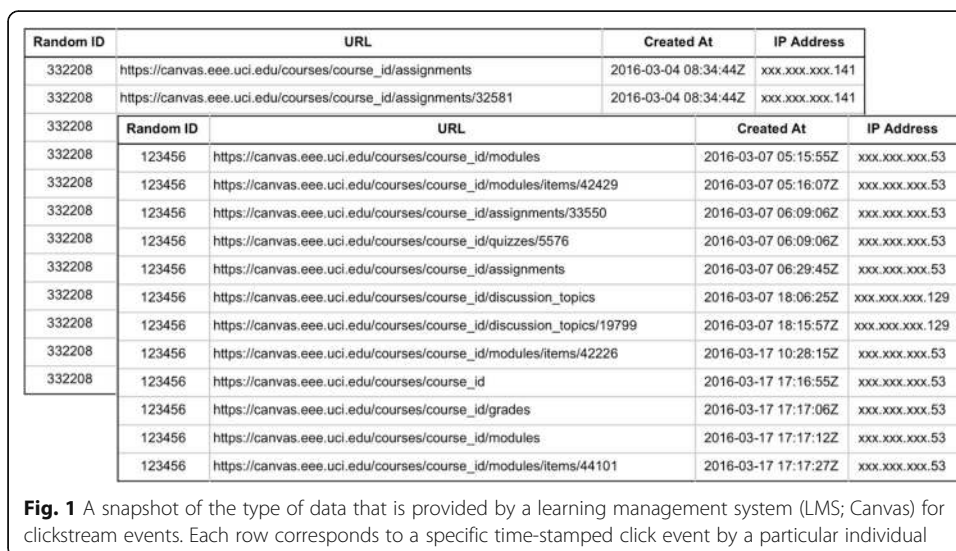
These varied findings underscore the importance of understanding whether SRL time management behaviors (e.g., procrastination, cramming, and time-on-task) are actually affected by these interventions and then whether an improvement in time management behaviors is effective at improving performance. Previous studies have taken on these questions by attempting to examine whether the underlying mechanisms are affected by various time management interventions. However, these studies have used crude measures, such as self-reported time spent per week (Häfner, Oberst, and Stock, 2014), days between completing assignments (Sitzmann and Johnson, 2012), numbers of web-page visits (Bannert, Sonnenberg, Mengelkamp, and Pieger, 2015), self-reports of time management behaviors (Azevedo and Cromley, 2004; van Eerde, 2003), or time of exam submission (Levy and Ramim, 2013). As discussed in the "Using clickstream data to understand SRL" section, nuanced analyses of rich clickstream data can provide more objective and detailed insights into how various interventions are, or are not, affecting student SRL behaviors and can thus allow for better targeted and more efficient interventions.

Clickstream data and its use in higher education research

In the practice and research of higher education, there is an emerging interest in the use of the timely and nuanced clickstream LMS data to better understand and support students' learning. Clickstream data are contained in the detailed logs of time-stamped actions from individuals interacting with LMSs (e.g., Canvas and Blackboard). These actions typically consist of events that a user initiates, such as navigating between web pages, downloading a file, or clicking play on a video. While such data only provide a partial and noisy record of a student's actions, they enable practitioners and researchers to collect information at scale about how students interact with online education resources and thus promise more objective and richer insight into the learning experience than many other methods. In this section, we explain the format of typical clickstream data, introduce major approaches that have been used by researchers in analyzing clickstream data, and provide a brief overview of the current uses of clickstream data in higher education.
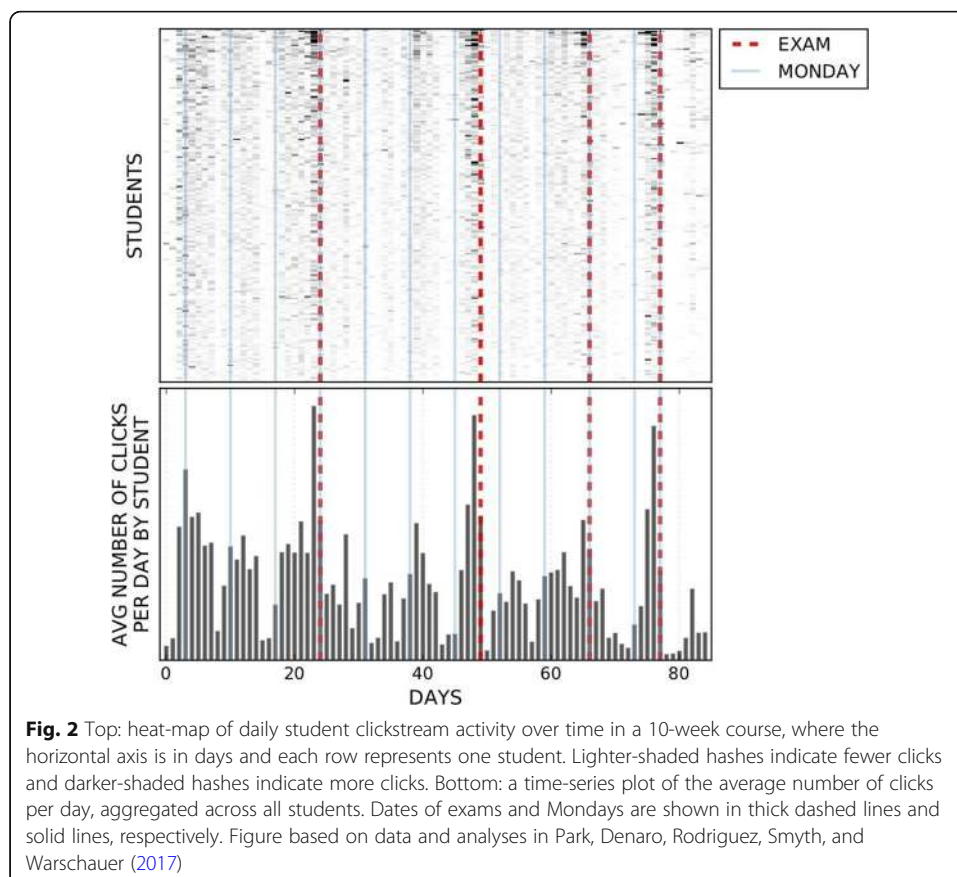
Figure 1 shows an example of the type of data that the LMS Canvas provides, based on students accessing a website associated with a course offering at the University of California, Irvine in 2016. Each row in Fig. 1 corresponds to an event generated by a particular student, identified via his or her (anonymized) Student ID. The URL is the web address of the resource being requested by the student, such as a request to navigate to a particular web page on the site or a request to download a file. One challenge in analyzing this type of data is that the URLs are not semantically meaningful by themselves, although the string names corresponding to the directory paths (e.g., "https://canvas.eee.uci.edu/courses/course_id/grades") often provide useful clues about the content that the student is requesting (in this case, grade information). In practice, most URLs can be readily assigned to categories such as "grades," "file downloads," "assignments," or "quizzes." This type of clickstream data can also be combined with LMS-provided information about additional student activities, such as the text content of search queries, text context in forum discussions, or interactions between students.

There are two somewhat different data analysis strategies that can be used to analyze clickstream data, each with its strengths and weaknesses. The first approach is based on *aggregate non-temporal representations* of the clickstream information per student, in which information is combined over time. An example would be to generate one histogram per student of the counts of actions of different types of activities over the duration of a course (e.g., number of clicks on lecture videos, number of clicks on the gradebook page). This allows for a flattened multivariate representation, with each student represented as a multidimensional vector. The advantage of this representation is that it is amenable to a multitude of statistical analyses, such as multivariate regression for predicting outcomes or clustering of students into groups. The disadvantage, however, is that this type of static aggregate representation does not retain any information about the sequential or temporal aspects of a student's behavior over the duration of a class (Mobasher, 2007; Spiliopoulou, 2000). *Time-dependent or sequence-dependent representations*, on the other hand, can retain more detailed information about a student's behavior over time. A simple example of a time-dependent representation is to count the number of total click events per student recorded per day over the duration of the

Random ID	URL	Created At	IP Address
332208	https://canvas.eee.uci.edu/courses/course_id/assignments	2016-03-04 08:34:44Z	xxx.xxx.xxx.141
332208	https://canvas.eee.uci.edu/courses/course_id/assignments/32581	2016-03-04 08:34:44Z	xxx.xxx.xxx.141

Random ID	URL	Created At	IP Address
123456	https://canvas.eee.uci.edu/courses/course_id/modules	2016-03-07 05:15:55Z	xxx.xxx.xxx.53
123456	https://canvas.eee.uci.edu/courses/course_id/modules/items/42429	2016-03-07 05:16:07Z	xxx.xxx.xxx.53
123456	https://canvas.eee.uci.edu/courses/course_id/assignments/33550	2016-03-07 06:09:06Z	xxx.xxx.xxx.53
123456	https://canvas.eee.uci.edu/courses/course_id/quizzes/5576	2016-03-07 06:09:06Z	xxx.xxx.xxx.53
123456	https://canvas.eee.uci.edu/courses/course_id/assignments	2016-03-07 06:29:45Z	xxx.xxx.xxx.53
123456	https://canvas.eee.uci.edu/courses/course_id/discussion_topics	2016-03-07 18:06:25Z	xxx.xxx.xxx.129
123456	https://canvas.eee.uci.edu/courses/course_id/discussion_topics/19799	2016-03-07 18:15:57Z	xxx.xxx.xxx.129
123456	https://canvas.eee.uci.edu/courses/course_id/modules/items/42226	2016-03-17 10:28:15Z	xxx.xxx.xxx.53
123456	https://canvas.eee.uci.edu/courses/course_id	2016-03-17 17:16:55Z	xxx.xxx.xxx.53
123456	https://canvas.eee.uci.edu/courses/course_id/grades	2016-03-17 17:17:06Z	xxx.xxx.xxx.53
123456	https://canvas.eee.uci.edu/courses/course_id/modules	2016-03-17 17:17:12Z	xxx.xxx.xxx.53
123456	https://canvas.eee.uci.edu/courses/course_id/modules/items/44101	2016-03-17 17:17:27Z	xxx.xxx.xxx.53

Fig. 1 A snapshot of the type of data that is provided by a learning management system (LMS; Canvas) for clickstream events. Each row corresponds to a specific time-stamped click event by a particular individual

class, resulting in a count-valued time-series per student of the number of events per day. Figure 2 presents examples of such representations. A time-dependent representation can reveal more subtle sequential patterns in student behavior than static multivariate representations, such as a change in student activity levels midway through a course (Mobasher, 2007; Spiliopoulou, 2000). But working with time-dependent data is more complicated than working with multivariate representations, and there are typically fewer data analysis tools available for working with such data, particularly with the type of event data that underlies clickstreams.

At the most basic level, using clickstream data in educational contexts allow us to analyze mechanical aspects of student behavior, such as the overall level and frequency of activity on a course website, the temporal patterns of students' online activity (both individually and relative to other students), and choices of which online resources students access. Such descriptions of student behavior, using various visualization and exploratory data mining techniques, were the focus of the earliest research in educational data mining (e.g., Baker and Yacef, 2009; Romero and Ventura, 2007). In recent years, the uses of clickstream data in educational research have expanded far beyond simple descriptions and have introduced both the possibility of empirical examination of educational theories using fine-grained process data and a new wave of data-driven pedagogical interventions (Fischer et al. 2020). The direction of these advances can be categorized into three main groups.

Fig. 2 Top: heat-map of daily student clickstream activity over time in a 10-week course, where the horizontal axis is in days and each row represents one student. Lighter-shaded hashes indicate fewer clicks and darker-shaded hashes indicate more clicks. Bottom: a time-series plot of the average number of clicks per day, aggregated across all students. Dates of exams and Mondays are shown in thick dashed lines and solid lines, respectively. Figure based on data and analyses in Park, Denaro, Rodriguez, Smyth, and Warschauer (2017)

First, clickstream data can help instructors and practitioners understand how students are using the available resources in an effort to improve instructional designs. For instance, instructors can monitor which resources students use most and test different designs that might allow them to better calibrate the course, either to emphasize important resources that are valuable but under-utilized by students or to provide more resources that students favor, affording more targeted guidance and feedback (Bodily and Verbert, 2017; Diana et al. 2017; Shi, Fu, Chen, and Qu, 2015). Second, the real-time accessibility of behavioral clickstream data can be used to develop automatic feedback and intervention modules within the LMS. For example, researchers have built early detection systems for dropout or poor course performance, which can help instructors allocate their attention to the most at-risk students (Baker, Lindrum, Lindrum, and Perkowski, 2015; Bosch et al. 2018; Lykourentzou, Giannoukos, Nikolopoulos, Mpardis, and Loumos, 2009; Whitehill, Williams, Lopez, Coleman, and Reich, 2015). Students can also be provided with adaptive guidance in real-time by, for instance, suggesting collaboration partners (Brusilovsky, 2003; Caprotti, 2017). Third, clickstream data allow for novel analyses that aim to advance understanding of how to identify and cluster student subgroups, as well as to personalize interventions to support learning processes. This includes the identification of student subpopulations with respect to their use of online resources (Gasevic, Jovanovic, Pardo, and Dawson, 2017) or students' engagement patterns in MOOC environments (Guo and Reinecke, 2014; Kizilcec, Piech, and Schneider, 2013). These student clusterings may be used in sequential modeling techniques such as recurrent neural network methods that populate a recommendation system of optimal course progression for different types of learners (e.g., Pardos, Tang, Davis, and Le, 2017).

Using clickstream data to understand SRL

One major line of research on using clickstream data is to measure student SRL behaviors with the goals of better understanding and supporting SRL (Roll and Winne, 2015). Previous studies have explored the use of clickstream data to measure SRL primarily in two types of technology-enhanced learning environments: interactive learning and LMS. The first group of studies has focused on interactive learning environments in which students are offered various tools that are designed to support SRL, including cognitive tools for information processing (e.g., note-taking window), goal-setting tools, reflection tools, and help-seeking tools (Nussbaumer, Steiner, and Albert, 2008; Perry and Winne, 2006; Winne and Jamieson-Noel, 2002). The second large group of studies has focused on student SRL behaviors using clickstream data from LMSs (e.g., blackboard and canvas), which are usually used to deliver learning materials (e.g., text, video, and audio), conduct learning activities (e.g., assignments and discussion), and support different forms of evaluation (e.g., exams and grade book systems; Lewis et al. 2005). The aspects of SRL behaviors that can be inferred using clickstream data collected from the two types of learning environments differ and are largely dependent on the types of interactions students can have within each learning environment.

The interactive learning environments embedded with SRL tools allow students to use one or more SRL tools to explicitly set goals for their learning tasks, monitor their learning process, use different cognition tools to process the information, and reflect

and adjust their learning. SRL behaviors, such as cognitive strategy use, planning, and help-seeking, are measured with data on the frequency of, timing of, characteristic conditions of, and behavioral reactions to the use of these SRL tools (e.g., Nussbaumer et al. 2008; Winne and Jamieson-Noel, 2002). While detailed and diverse SRL behaviors can be inferred from data collected from these interactive learning environments, most of these learning environments are used in laboratory studies (e.g., Perry and Winne, 2006; Winne and Jamieson-Noel, 2002) or for specific domains or topics (e.g., learning the human life cycle; Perry and Winne, 2006) and thus have not been commonly adopted in higher education.

Unlike these interactive learning environments designed for specific domains or topics, LMSs are widely adopted in higher education contexts to support the basic processes that are necessary for learning any subject online (Lewis et al. 2005). Specifically, students usually interact with LMSs by downloading course materials, watching video lectures online, submitting assignments, completing quizzes, posting on the discussion forums, and so on (Lewis et al. 2005). Largely due to the fact that the features of learning management platforms are not set up to explicitly encourage and measure SRL, only a few studies have examined how to use clickstream data from LMSs to measure SRL, and these studies have mainly focused on the sub-concept of time management skill because it is most amenable to measurement (e.g., Baker et al. 2019; Cicchinelli et al. 2018; Crossley, Paquette, Dascalu, McNamara, and Baker, 2016; Lim, 2016; Park et al. 2018; You, 2016). Researchers have used measures such as the frequency with which students view resources pertaining to course dates and deadlines (Cicchinelli et al. 2018; Park et al. 2018), how far in advance students start work on/turn in various assignments (Crossley et al. 2016; Kazerouni, Edwards, Hall, and Shaffer, 2017; Levy and Ramim, 2013), and how close together work sessions are (e.g., Baker et al. 2019; Park et al. 2018) to examine students' time management skills.

In addition, recent work has interrogated the extent to which clickstream measures provide valid inference about various SRL constructs in two ways: (1) by examining whether students' perceptions about their self-regulated learning correspond to their click patterns, and (2) by examining the extent to which clickstream measures complement self-reported measures in predicting student course performance. One recent study found that clickstream data are helpful measures of true time management skills (Li et al. 2020). First, the clickstream measures were strongly correlated with students' self-reported time management skills from a post-course survey (and somewhat correlated with measures from a pre-course survey). Second, the clickstream measures of time management were better predictors of students' performance in the class than were the self-reported measures (Li et al. 2020). These results suggest that clickstream measures of SRL offer insightful and valid information about students' actual learning processes.

The studies above suggest clickstream data can provide objective and timely measures of SRL, such as time management skills, that can be easily scaled up for large student populations. These clickstream measures can be used to examine the relationships between SRL behaviors and performance, which may provide additional and novel information on the role of SRL in online learning beyond existing findings based on self-reported data. Moreover, unlike self-reported measures that are usually collected at only one or limited time points, these measures can be used to investigate how student

SRL behaviors unfold over time and to explore how personal and environmental factors influence SRL behaviors. Finally, the nuanced information on individual learning processes that clickstream data can uncover is useful in understanding how and why an SRL intervention influences student learning outcomes. Indeed, scholars (e.g., Damgaard and Nielsen, 2018) have recently argued that examining the mechanisms that various behavioral interventions affect is "crucial," as interventions can have unintended negative consequences if the likely affected behavioral pathways are not well understood (Damgaard and Nielsen, 2018, p. 313).

In the following sections of this paper, we provide researchers, instructors, and administrators with examples of these promising avenues of research—defining and identifying behavioral patterns that are related to student learning outcomes, suggesting behavioral changes to students for greater success, and providing insights regarding the mechanisms by which education interventions affect student outcomes. In discussing this growing field of research, we specifically highlight the ways in which decontextualized, noisy, and sparse clickstream data can provide only partial answers to many questions by focusing on the specific strategies and cautions necessary for working with these data.

The challenges and benefits of using clickstream data to measure SRL and understand mechanisms: four example studies

In this section, we provide illustrations of three main considerations of using clickstream data to understand student SRL through narrations of four of our own studies that have used clickstream data in applied educational settings. The discussions of these four papers reflect the findings of extant literature by highlighting the ways in which clickstream data can be particularly helpful in this context—by providing time-varying measures, by showing how students use course resources, and by illuminating which specific behaviors interventions affect—but also allow us to demonstrate the unique challenges and considerations that come with working with clickstream data. Specifically, we highlight (1) the importance of measurement and pre-processing clickstream data in studies I and II, (2) the crucial nature of understanding the context from which the clickstream data emerge in study III, and (3) the affordances and limitations of using clickstream data to understand the mechanisms of educational interventions in study IV.

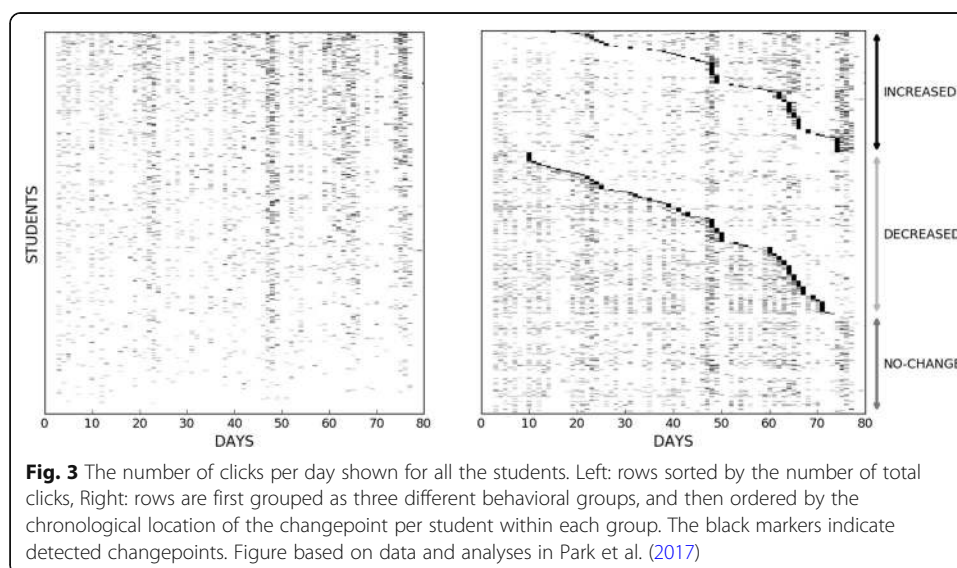### Complications in constructing valid measurement using clickstream data

The first two example studies aggregated raw clickstream data into daily activity counts for each student and applied two different statistical models to summarize the temporal patterns of students' engagement to examine scheduling behavior. The first example study, Park et al. (2017), focused on how student engagement changes over the length of a course. The authors used Poisson regression models to examine how active each student was at a daily level, relative to the class mean. The activity was defined as downloads of course files and each action was classified as either previewing or reviewing course material (described in more depth below). The authors used these daily counts to determine whether and when a student changed this relative activity level during the course. If a change point existed, the authors further determined, based on
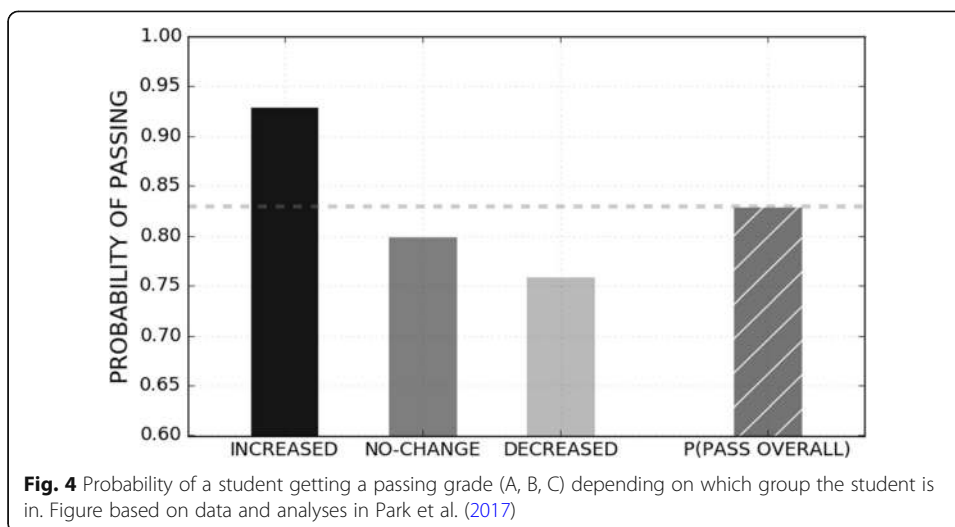
the estimated coefficients, whether the student increased or decreased daily previewing and reviewing activity, and classified students into one of three groups: *increased*, *decreased*, and *no change* (depicted in Fig. 3).

These types of detected behavior changes were highly correlated with student outcomes. Figure 4 shows the probabilities of a student getting a passing grade given that the student was in the increased group, no change group, or decreased group, compared to the unconditional probability of a student passing. Students in the increased group had a higher probability of passing the course, while the opposite was true for the decreased group; students who increased the activity (relative to all of the students in the course) at some point during the course term had a higher chance of passing the class than those who did not. This distinction was true for both the preview and review daily activity but the relationship was stronger for review activity. These findings highlight that such nuanced and dynamic descriptors of behavior (e.g., change in engagement behavior rather than absolute engagement behavior) could enable instructors to better identify students who are at risk of poor learning outcomes. Future work could test these hypotheses in a causal framework.

The second example study examined students' procrastination behaviors in two 5-week online courses (Park et al. 2018). In each week, students were assigned five tasks on Monday, which were designed to be completed on a daily basis but were all due at the end of the week (on Friday). This setting thus gave students flexibility in time scheduling and allowed them to procrastinate each week. Poisson mixture models on latent student profiles—each as measured by how many assignments students completed across days of the week—identified best fit for two latent profiles, procrastinators and non-procrastinators (Fig. 5).
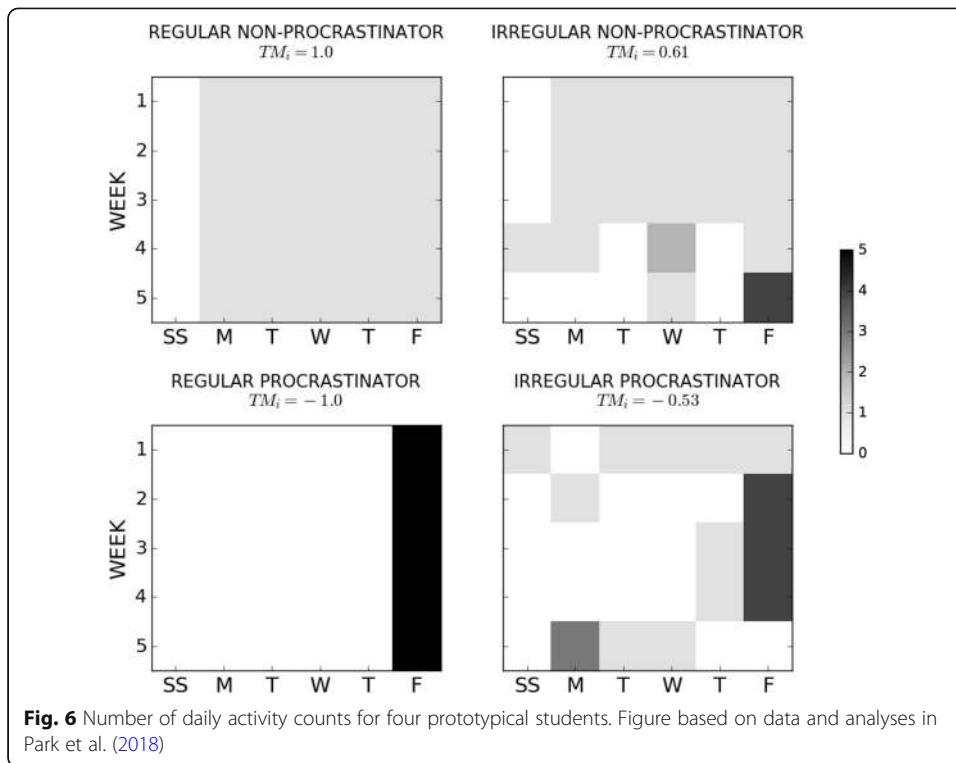
We then extracted two dimensions of time management behavior: overall degree of procrastination (measured by how heavily a student is weighted on the procrastinating profile) and regularity of procrastination (measured by how variant the component weights on the procrastinating profile are across the 5 weeks of the course). We (Park et al. 2018) used a dual-dimension model to assign each student a "Time Management Score." This composite score takes the highest value for students with low overall

Fig. 3 The number of clicks per day shown for all the students. Left: rows sorted by the number of total clicks, Right: rows are first grouped as three different behavioral groups, and then ordered by the chronological location of the changepoint per student within each group. The black markers indicate detected changepoints. Figure based on data and analyses in Park et al. (2017)

Fig. 4 Probability of a student getting a passing grade (A, B, C) depending on which group the student is in. Figure based on data and analyses in Park et al. (2017)

procrastination and high regularity (regular non-procrastinators) and lowest value for students with high overall procrastination and high regularity (regular procrastinators). Students who show less regular behaviors have score values somewhere between the extremes. Examples of four prototypical students are shown in Fig. 6. Interestingly, heatmap plots of the average behavioral patterns of students indicated that procrastination was related to course grades (Fig. 7). Students who received an A are shown in the left panel, with moderate levels of engagement every day throughout the course. In contrast, students who received a C, D, or F are shown in the right most panel, with high levels of engagement, indicated by dark shading, on Fridays and much lower levels

Fig. 5 Poisson mixture component means from modeling aggregated daily task counts for the class in 2016 (upper) and 2017 (lower). The left panels show aggregated daily counts for non-procrastinators and the right panels for procrastinators. Figure based on data and analyses in Park et al. (2018)
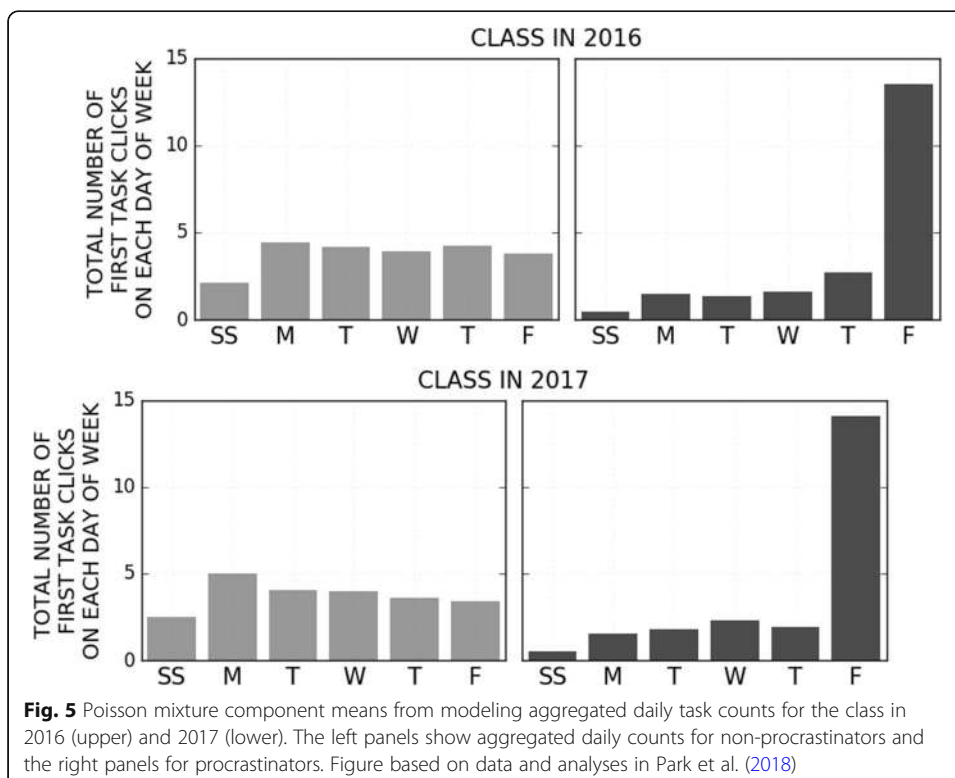
Fig. 6 Number of daily activity counts for four prototypical students. Figure based on data and analyses in Park et al. (2018)
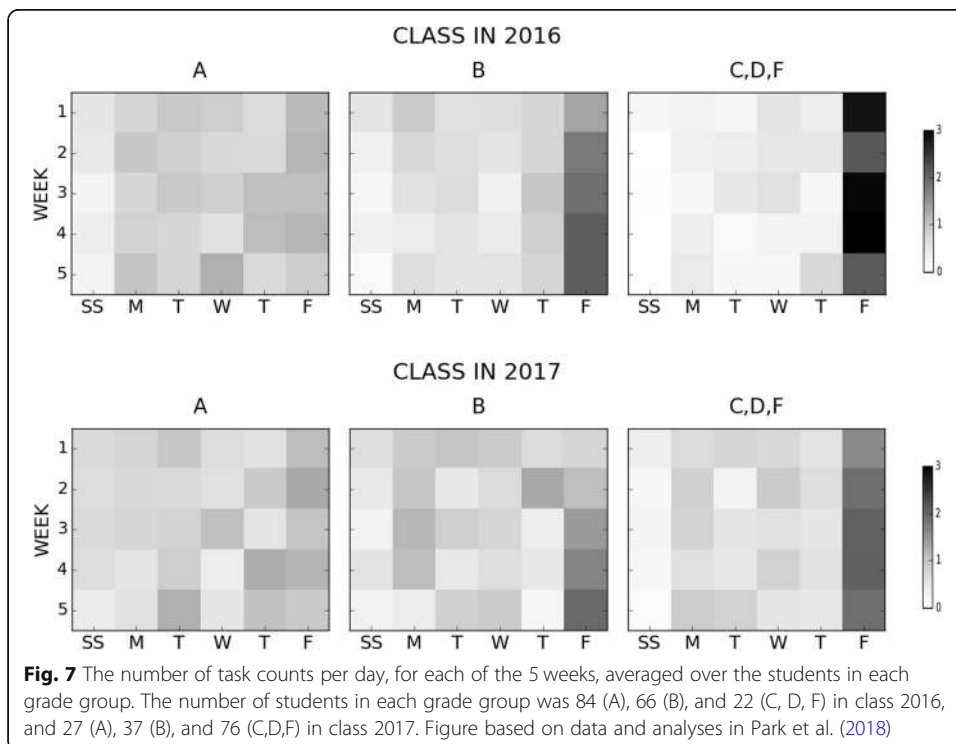
Fig. 7 The number of task counts per day, for each of the 5 weeks, averaged over the students in each grade group. The number of students in each grade group was 84 (A), 66 (B), and 22 (C, D, F) in class 2016, and 27 (A), 37 (B), and 76 (C,D,F) in class 2017. Figure based on data and analyses in Park et al. (2018)

Fig. 8 Distribution of Time Management Score in different grade groups of class in 2016 (left) and in 2017 (right). Lines in bars indicate the median score in each group. H-statistic comes from a Kruskal-Wallis test. Figure based on data and analyses in Park et al. (2018)
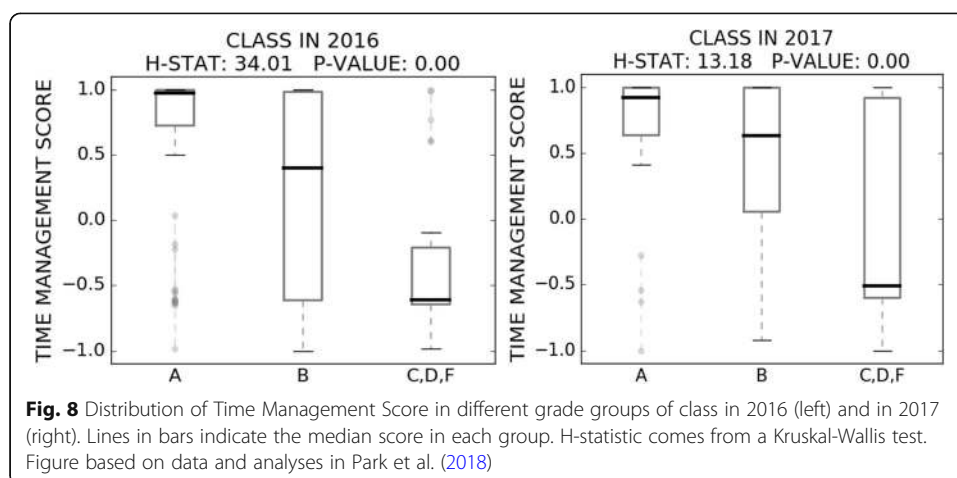
on other days. To more rigorously test, this in the Poisson mixture framework, we ran Kruskal-Wallis tests, which indicated that the overall degree of procrastination, the regularity of procrastination, and the composite Time Management Score were significantly different across grade groups and that the composite Time Management Score (Fig. 8) had more differentiating power than the two individual measures. This suggests that by incorporating two measures, procrastination, and regularity, the Time Management Score amplifies the behavioral information that is predictive of performance, providing a more nuanced view of procrastination. Again, the more nuanced understanding of student behavior that is possible by examining clickstream behavior can provide instructors and researchers with hypotheses to test in more causal frameworks.

In studies like the above two examples, extensive data exploration and pre-processing, as well as careful decisions about measurement, are necessary to obtain valid conclusions. In the first study, in order to produce useful and interpretable results, we needed to carefully filter out relevant data from the Canvas log files. The raw log files contain a record for each student click event in the form of student ID, time-stamp, and URL. In order to make these data useful for analysis, relevant clicks—in this case, preview and review events defined as a student downloading lecture notes or an old exam before (preview) or after (review) the time of the corresponding event (such as a quiz, exam, or lecture) relevant to the file in question—were kept and click events not relevant to the analysis, such as navigation events, were filtered out. Converting the raw logs into streams of preview/review events consisted of (i) manually identifying both the relevant URLs for file downloads and the relevant times (such as that of a lecture or exam) associated with each file and then, given this information, (ii) creating a program to automate the assignment of a "preview" or "review" label to each individual download event by each student. This process highlights the fact that complete automation is not possible in most educational contexts, as instructor autonomy—a hall-mark feature of higher education in the USA—almost always results in unique LMS environments that necessitate some manual steps in interpreting clickstream LMS data.

Viewing and analyzing the clickstream data from different angles can help to determine which part of the data has meaningful information and which data provide little additional information and should be filtered out. This "noise" may exist in different
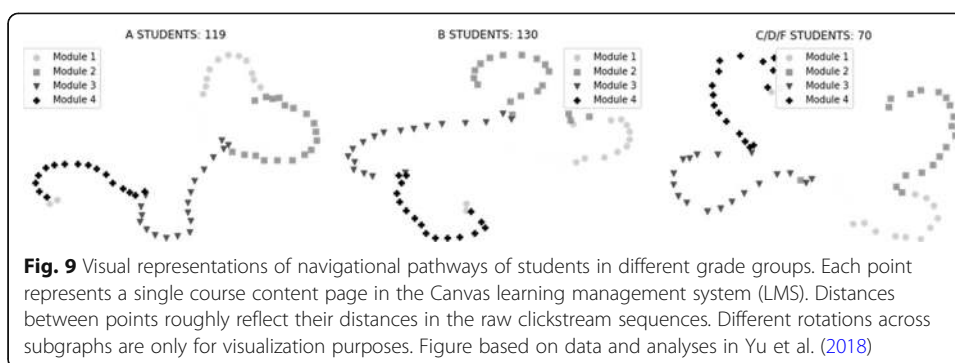
dimensions, such as at the student level or at the click level. For example, students with extremely low levels of activity (0 or 1 total clicks) were excluded from both of the studies. In the second study, these instances of low click activity could be because they did not watch videos or because they accessed the videos in some way other than through their Canvas account (e.g., by watching with peers). Because of the structure of the course and the available LMS data, only clickstream data related to lecture watching was analyzed in this study. Thus, the students with no record of lecture watching were excluded from the analyses as it was impossible to measure and categorize their scheduling and spacing behaviors. Noise may also exist at the click level. For example, researchers may choose to ignore clicks that are not directly related to the behavior of interest. In this case, clicks on irrelevant content (such as navigation activities in the first study) and duplicate actions were identified and removed.

As we discuss in the next section, carefully examining how the course is designed is another key element for analyzing student clickstream data. In the case of our first example study, detecting changes in student behavior may be a more relevant study for a course that is offered for a longer term, and in the case of our second example study, using a weekly vector of daily counts to examine patterns of procrastination patterns is only possible when a course has a weekly repeated structure (that is, roughly the same deadlines, activities, and assignments each week). Understanding the instructor's intention behind the course design, combined with the researcher's interest, should inform decisions on the granularity (e.g., hourly, daily, weekly) and the form (e.g., vector of counts, sequences of activity types) of the processed data, which may also lead to using different modeling techniques.

The crucial nature of understanding the context

The third example study (Yu, Jiang, and Warschauer, 2018) investigated navigational behavior by using neural network methods to examine student pathways through course components in a 10-week fully online introductory STEM course. The course had 4 modules, each containing around 10 videos on the core course content, along with quizzes and other activities. This organization of course content roughly followed the order of knowledge inquiry expected by the instructor. Given the flexibility and freedom with which students could browse content pages within the learning system, students could potentially exhibit a myriad of non-conforming pathways, reflecting individual differences in learning progress and/or study strategies. The raw clickstream data (i.e., time-stamped sequences of students' visits to different course pages) were used to generate descriptive accounts of the order in which students accessed course materials. The authors applied a skip-gram model to extract sequential interdependencies among course pages (i.e., their distances from each other in the raw sequences) and used this information to project each course page into a multidimensional vector. In such a model, described at length in Yu et al. (2018), the more often two-course pages were visited in similar contexts (neighboring pages in the sequence), the closer their vectors would be to each other. If a student strictly followed the designed sequence, the pages would be linearly connected in the designed order in the vector space.

On average, students in this course followed the suggested course structure. However, students who did well in the course (earned an A) followed a more linear pattern

Fig. 9 Visual representations of navigational pathways of students in different grade groups. Each point represents a single course content page in the Canvas learning management system (LMS). Distances between points roughly reflect their distances in the raw clickstream sequences. Different rotations across subgraphs are only for visualization purposes. Figure based on data and analyses in Yu et al. (2018)

than students who earned a C/D/F (Fig. 9). As shown in the first subgraph (left panel), the pages generally followed the intended order from Module 1 to Module 4 for students who earned an A. This trend was somewhat broken among the B students, where the majority of Module 1 pages fell away from the main thread in a separate segment and a few Module 2 pages left their expected positions. Among C/D/F students, the entire page set was further split into two segments, with some "outlier" pages from Module 2 as well as non-linearities within Module 4. This indicates that high-performing students are more likely than low-performing peers to follow the course sequence intended by the instructor (e.g., watching all videos from Module 1 before all videos from Module 2), while low-performing students are more likely to watch videos out of the intended order (e.g., watching videos from Module 4 before videos from Module 1). This finding highlights that understanding the actual pathways that students take could help instructors to identify and redirect struggling students or to redesign courses better aligned with students' navigational behaviors.

Similar to the first two example studies described above, performing such pathway analyses entails a series of data processing strategies that, with potential modifications, can generalize to other contexts. The particular decisions must be driven by the specific context of the course under study. First, the specific content each click points to, or the action it represents, must be determined from the URLs. This often requires the help of the course instructor because the specific coding of the URLs is determined partially by LMS's functionality and partially by the specifics of the course design; understanding what each click actually means requires intimate knowledge of the course structure and content. To give an example, these two URLs (canvas.uci.edu/pages/segment-5 and canvas.uci.edu/files/89283) show the potential complexity of determining content from URLs. The former URL represents a required content video, whereas the latter URL is a visit to a file uploaded by the instructor. Second, the researcher must decide which clicks to keep. In this third example study, the research focus was on the order in which students visited the content videos, as the instructor indicated that the series of lecture videos represented the core content in the class. Thus, only clicks on the video pages were included in the analysis. Third, the remaining clicks from each student were sorted by timestamp and converted to a tokenized sequence, in which each click was assigned the unique ID for the content page it pointed to. These tokenized sequences were then ready for modeling.

However, there are some important particularities to this analytical pipeline. Most prominently, the behavioral contexts of the "interesting" clicks are lost when other

student actions are removed. For example, given two students who watch the course videos in the same order, one student might engage with other course content in an effort to better understand the video, such as by checking the syllabus and/or course requirements more often. Such behavioral differences outside of researchers' focus will be invisible in the cleaned sequences and, like any other omitted variables in traditional education research, might bias researchers' conclusions about the "interesting" clicks.

Also, without a deep understanding of the context, researchers might fail to notice nuances of instructional design that actually play important roles in shaping students' learning behavior. For instance, if a quiz includes a question about a specific point included in an early video, even students who usually follow the designed instructional sequence might choose to re-watch earlier videos. This would result in outliers and non-linearities in the types of graphs generated by the current study. If, however, all quizzes are focused on general comprehension of video content, these types of graphs might be more accurate representations of how different types of students navigate the course space.

In light of these types of caveats, researchers should not simply apply the foregoing approach to other course contexts. Instead, researchers need to closely communicate with the instructors or instructional designers to understand the rationale of the observed course structure and decide on what behaviors should be the research focus. Potentially "noisy" clicks should be carefully examined before being left out, in case they carry meaningful signals of study habits.

The affordances and limitations of using clickstream data to understand mechanisms of education interventions

The fourth example study examined a time management intervention in a for-credit, fully online class at a selective 4-year university (Baker et al. 2019). The authors provided a randomly selected half of the students in a for-credit online physics class with the opportunity to schedule when they would watch the lecture videos in an otherwise asynchronous, unscheduled class. In each of the first 2 weeks of the course, students in the treatment group were sent an email nudging them to think about upcoming coursework and to plan when they would watch each of the five lecture videos for the week. The students were given a link to an online survey tool in which they could explicitly state what day and time they planned to watch the course lecture videos. Students in the control group were also contacted at the beginning of the first and second weeks, but they were sent a survey that contained theoretically inert questions, such as what web browser they used to watch course videos. Students were not told of the two conditions, and they did not have access to the other group's survey. All students received extra credit for completing the weekly surveys, and survey response rates were uniformly high.

The authors found mixed effects on course performance. The intervention produced a significant positive effect on the first weekly quiz and a significant negative effect on the last weekly quiz of the course. The effect of the scheduling intervention on the final exam and course grade was positive but not significant (see Baker et al. 2019, for full results). Like most interventions in education, there are a number of theoretically motivated mechanisms through which this type of intervention could affect student
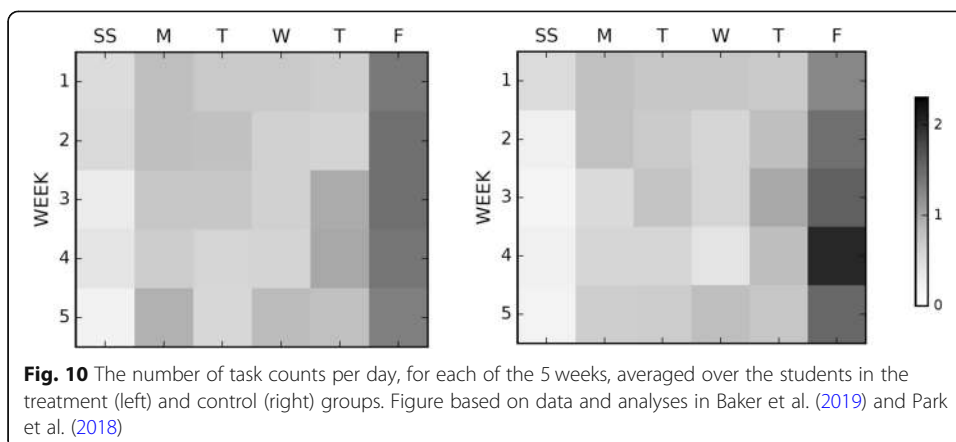
educational outcomes and produce the pattern of positive to negative results found. The positive effects at the beginning of the course could be the result of a number of distinct mechanisms. The suggestion to schedule in advance could reduce the probability that students will work at non-ideal times of day, which has been found to be related to worse academic outcomes (e.g. Carrell, Maghakian, and West, 2011; Goldstein, Hahn, Hasher, Wiprzycka, and Zelazo, 2007; Hahn et al. 2012). It could reduce students' propensity to cram or procrastinate, which has also been found to be negatively related to success in online classes (e.g. Elvers et al. 2003; Michinov et al. 2011). A nudge to schedule time to do coursework could increase the time that students spend on their coursework, which has been found to be positively related to course outcomes (Beattie, Laliberté, Michaud-Leclerc, and Oreopoulos, 2017). This nudge could also serve to reduce academic anxiety and stress, which could positively affect academic outcomes (e.g., Macan, Shahani, Dipboye, and Phillips, 1990; Misra and McKean, 2000). The negative effects at the end of the course could also be the result of a number of mechanisms. Withdrawing the scheduling intervention (and thus removing a support that the students were relying on) might have led to worse time management practices (such as cramming and procrastination) for students in the later weeks of the course. Worse time management practices could also explain the negative effect if the externally imposed time management structure negatively affected students' own intrinsically motivated time management strategies, such as active procrastination (Chu and Choi, 2005; Seo, 2013).

The availability of clickstream data allowed the researchers to examine some, but not all, of these potential mechanisms. Specifically, the authors examined the effect of the treatment on the time of day in which students watched lecture videos, on students' procrastination, and on students' cramming. Clickstream data can provide partial measures of other potential mechanisms (such as student time on task) and do not provide any purchase on understanding other potential mechanisms.

We first present the authors' heatmaps of treatment and control students' average level of interaction over the course of the term in Fig. 10. While these do not provide conclusive evidence about differences (or lack thereof) between the two groups, they do illustrate that there are not immediately obvious differences in behavior that could explain the results. Plots of the time of each interaction for all treatment students and all control students (collapsed down to a 24-h period) allowed the researchers to examine the effect of the intervention on the time of day that students interacted with the course material. Figure 11 shows the treatment and control distributions for the first 2 weeks of the course (left panel) and the last 3 weeks of the course (right panel). There is suggestive visual evidence that the treatment students were, on average, engaged with course materials earlier in the day than control students in the first 2 weeks of the course (when the treatment was active), though this does not stand up to statistical tests. Kolmogorov-Smirnov (K-S) tests show that the two distributions were not statistically significantly different from each other. In weeks three through five of the course, the treatment and control distributions appear quite similar and K-S tests for the joint weeks 3–5 and week 5 alone were also not significant.

To examine the effect of the intervention on students' procrastination and cramming behaviors, the authors relied on measures of spacing and procrastination and a composite time management measure (derived using Poisson mixture models and described

Fig. 10 The number of task counts per day, for each of the 5 weeks, averaged over the students in the treatment (left) and control (right) groups. Figure based on data and analyses in Baker et al. (2019) and Park et al. (2018)

in the "Complications in constructing valid measurement using clickstream data" section above). The authors found that being assigned to treatment had no effect on measured procrastination, spacing, or the composite time management score; students in the treatment group exhibited very similar engagement patterns to students in the control group. Importantly, there were no differences in procrastination and cramming between the treatment and control groups in the 2 weeks in which there was a treatment effect, weeks 1 and 5. There was also no significant difference between the treatment and control groups on the overall Time Management Score (see Baker et al. 2019, for full results).

A plausible secondary mechanism by which the treatment could have affected outcomes is by inducing the treatment students to spend more time on their classwork. This could have resulted if students were induced to start their work earlier (procrastinate less) or work at more ideal times of day (when their work time is less likely to be

Fig. 11 Time of course interactions, control and treatment students, first and second weeks (left panel) and third through fifth weeks (right panel) of 5-week Physics course. Figure based on data and analyses in Baker et al. (2019)
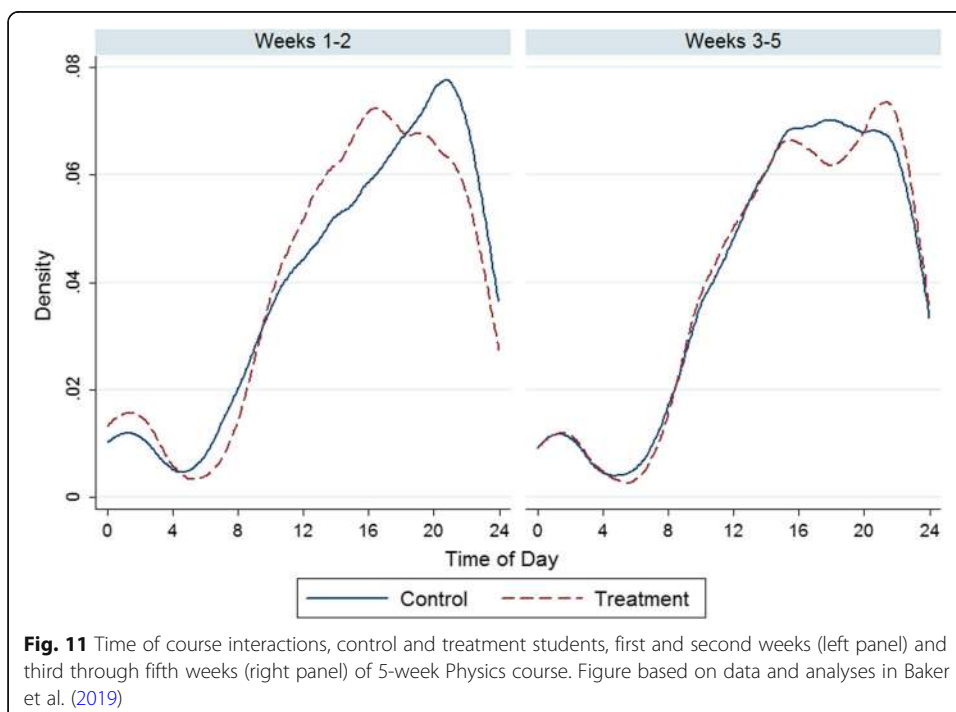
cut short by fatigue or other obligations). Unfortunately, the clickstream data provided by Canvas do not provide direct measures of students' time on task. Therefore, the authors used the total number of clicks per week as a proxy for time on task and found no evidence that the treatment induced students to spend more time engaging with the course platform. This ersatz measure of time-on-task is an example of how clickstream data can provide some, but not sufficient, insight into student behavior online. Clickstream data cannot track what students are doing in between clicks. Time in between clicks ranges from milliseconds to days or even weeks. While it is safe to assume that two consecutive clicks with a half-second between them indicate that the student is still present, and that two consecutive clicks with 3 days in between indicate that the student has logged off, it is difficult to distinguish if a student with 20 min between clicks is reading an article (or watching a lecture video) closely, if the student stepped away briefly, or if the student is engaged in another activity simultaneously. Thus, the number of clicks provides a noisy measure of time on task.

Another difficulty of working with clickstream data is that current measures might not provide a comprehensive enough understanding of student behaviors. For example, the authors were unable to capture the activity of web pages outside of the LMS, so their data on online course-related activity were incomplete. Unfortunately, course content is rarely limited to just the web pages on the LMS—instructors tend to post external links to other websites either because that website already has what the instructor needs or because that website is easier to work with for the instructor's specific needs (for example, many instructors find that uploading lecture videos to Youtube is easier than using the LMS). Even if the researcher is able to obtain click-data on user activity from the website outside of the LMS, matching student information on the external website poses another challenge. Thus, clickstream measures of a number of SRL planning behaviors, such as procrastination, cramming, and time-on-task, are potentially incomplete and noisy due to data availability.

Finally, many interventions that aim to affect academic outcomes by influencing SRL behaviors might actually act through other mechanisms, such as by reducing or increasing anxiety, that are not measurable using clickstream data.

Discussion and conclusion

For many years, educational researchers have tried to crack the "black box" of learning, by better illuminating the learning *processes* that lead to learning *outcomes*. By automatically recording students' interactions with online course materials, clickstream data provide a valuable new source of information on student learning behaviors. These data can be used to define and identify behavioral patterns that are related to student learning outcomes, suggest behavioral changes to students for greater success, and provide insights regarding the mechanisms by which education interventions affect student outcomes. Yet, the raw clickstream data are noisy and thus not useful to educational researchers and course instructors without additional steps of data processing and analysis.

There are important caveats in working with clickstream data. First, there are several limitations in analyzing clickstream data across classes using broad and universal metrics of engagement. For example, individual instructors at the same school may enable or disable specific LMS features, and the general structure of the available materials

may vary across courses. These differences can bias the conclusions researchers may generate about student engagement. As our example study III above shows, it is important to work closely with course instructors to understand the specific course structure, available resources, and the resources which the instructor deems most important.

Course meta-data, such as the information provided on the syllabus, provide important context to raw clickstream data. Unfortunately, for most courses, critical information, such as the dates of exams, are not recorded within the LMS in a standard format and must instead be gleaned via manual inspection of course materials. The manner in which online resources are organized may differ across courses, and these differences can affect researchers' understanding of clickstream data. Unfortunately, examining the structure and meta-data manually on a course-by-course basis necessitates a time-consuming step in the data analysis process.

Second, significant heterogeneity in behavior and variability over time also often complicates analyses. For example, the overall level of click activity may vary greatly across individual students in a course, with some students generating a very large number of clicks on a regular basis and others generating few or even zero clicks. Depending on the questions being asked, it may or may not make sense either to analyze the raw (absolute) daily click counts per student or to analyze standardized (relative) versions of the counts (e.g., by scaling each student's counts by his or her mean daily count).

Third, engagement data from online classes can be missing for a number of reasons. For instance, clickstream data only capture students' interactions with online materials. Researchers, instructors, and administrators may miss important learning processes if students use alternative methods for studying, such as watching external videos, reading books, or reviewing notes. It is sometimes possible to infer some activities that occur outside of the LMS, such as video watching on an external website, but such data gathering usually requires bespoke strategies. Example study IV highlights how such missing data can complicate or invalidate certain important measures that are created using clickstream data.

Fourth, it is important for users of clickstream data to be careful about monitoring data quality. From the time a student generates a click event to the time a researcher obtains a representation of that same click event, that information may have passed through a pipeline of different pieces of software. It is only natural that from time to time there are issues in this data collection and aggregation pipeline. For example, something as simple as a quick power outage at the web server may disrupt time-series clickstream data. Another issue easier to address but important not to overlook is the interpretation of time-stamps, which may be recorded on the web server time in local time or in Greenwich Mean Time depending on the particular log file format being used by the server, with consequences for evaluating student behavior.

Finally, there are complicating, non-analysis factors specific to the educational context. For example, one great promise of clickstream data is that they can be combined with student-level data such as survey data, demographic data, and education/course records from multiple classes. However, such linkages require access to these data, from either instructors or school administrators, as well as access to the necessary linking ids, which do not always exist across courses and datasets. Using clickstream data to their fullest potential requires coordination at the school level and cannot be accomplished by using data from only one class.

These limitations notwithstanding, as teaching and learning continue to gravitate to digital environments, the role of clickstream data will become increasingly valuable in understanding student SRL behaviors. Those involved in promoting educational research, whether professional associations, academic journals, funding agencies, or graduate and undergraduate training programs, will need to consider how they can better prepare the next generation of educational scholars to exploit this valuable data source to illuminate student learning.

Author details
[1]University of California, 2060 Education, UC, Irvine, Irvine, CA 92697, USA. [2]New York University, New York, USA. [3]University of Tübingen, Tübingen, Germany.

References
Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science, 13*(3), 219–224. https://doi.org/10.1111/1467-9280.00441.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*(3), 523–535. https://doi.org/10.1037/0022-0663.96.3.523.

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), 1–18. https://doi.org/10.1177/2332858416674007.

Baker, R., Evans, B., Li, Q., & Cung, B. (2019). Does inducing students to schedule lecture watching in online classes improve their academic performance? An experimental analysis of a time management intervention. *Research in Higher Education, 60*(4), 521–552. https://doi.org/10.1007/s11162-018-9521-3.

Baker, R. S., Lindrum, D., Lindrum, M. J., & Perkowski, D. (2015). Analyzing early at-risk factors in higher education e-learning courses. In *Paper presented at the meeting of the International Educational Data Mining Society, Spain.*

Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: a review and future visions. *JEDM Journal of Educational Data Mining, 1*(1), 3–17. https://doi.org/10.5281/zenodo.3554657.

Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior, 52*, 293–306. https://doi.org/10.1016/j.chb.2015.05.038.

Bawa, P. (2016). Retention in online courses: Exploring issues and solutions—A literature review. *SAGE Open, 6*(1), 1–11. https://doi.org/10.1177/2158244015621777.

Beattie, G., Laliberté, J. W. P., Michaud-Leclerc, C., & Oreopoulos, P. (2017). *What sets college thrivers and divers apart? A contrast in study habits, attitudes, and mental health*, (p. w23588). Cambridge: National Bureau of Economic Research.

Bodily, R., & Verbert, K. (2017). Trends and issues in student-facing learning analytics reporting systems research. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, (pp. 309–318). https://doi.org/10.1145/3027385.3027403.

Bosch, N., Crues, R., Henricks, G. M., Perry, M., Angrave, L., Shaik, N., … Anderson, C. J. (2018). Modeling key differences in underrepresented students' interactions with an online STEM course. *Proceedings of the Technology, Mind, and Society, 6,* 1–6. https://doi.org/10.1145/3183654.3183681.

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education, 27,* 1–13. https://doi.org/10.1016/j.iheduc.2015.04.007.

Brusilovsky, P. (2003). Adaptive navigation support in educational hypermedia: The role of student knowledge level and the case for meta-adaptation. *British Journal of Educational Technology, 34*(4), 487–497. https://doi.org/10.1111/1467-8535.00345.

Burger, N., Charness, G., & Lynham, J. (2011). Field and online experiments on self-control. *Journal of Economic Behavior & Organization, 77*(3), 393–404. https://doi.org/10.1016/j.jebo.2010.11.010.

Caprotti, O. (2017). Shapes of educational data in an online calculus course. *Journal of Learning Analytics, 4*(2), 76–90. https://doi.org/10.18608/jla.2017.42.8.

Carrell, S. E., Maghakian, T., & West, J. E. (2011). A's from Zzzz's? The causal effect of school start time on the academic achievement of adolescents. *American Economic Journal: Economic Policy, 3*(3), 62–81. https://doi.org/10.1257/pol.3.3.62.

Chu, A. H., & Choi, J. N. (2005). Rethinking procrastination: positive effects of "active" procrastination behavior on attitudes and performance. *The Journal of Social Psychology, 145*(3), 245–264. https://doi.org/10.3200/SOCP.145.3.245-264.

Cicchinelli, A., Veas, E., Pardo, A., Pammer-Schindler, V., Fessl, A., Barreiros, C., & Lindstädt, S. (2018). Finding traces of self-regulated learning in activity streams. In *Proceedings of the 8th international conference on learning analytics and knowledge*, (pp. 191–200). https://doi.org/10.1145/3170358.3170381.

Crossley, S., Paquette, L., Dascalu, M., McNamara, D. S., & Baker, R. S. (2016). Combining click-stream data with NLP tools to better understand MOOC completion. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, (pp. 6–14). https://doi.org/10.1145/2883851.2883931.

Damgaard, M. T., & Nielsen, H. S. (2018). Nudging in education. *Economics of Education Review, 64,* 313–342. https://doi.org/10.1016/j.econedurev.2018.03.008.

Diana, N., Eagle, M., Stamper, J., Grover, S., Bienkowski, M., & Basu, S. (2017). An instructor dashboard for real-time analytics in interactive programming assignments. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, (pp. 272–279). https://doi.org/10.1145/3027385.3027441.

Duncan, T. G., & McKeachie, W. J. (2005). The making of the motivated strategies for learning questionnaire. *Educational Psychologist, 40*(2), 117–128. https://doi.org/10.1207/s15326985ep4002_6.

Elvers, G. C., Polzella, D. J., & Graetz, K. (2003). Procrastination in online courses: Performance and attitudinal differences. *Teaching of Psychology, 30*(2), 159–162. https://doi.org/10.1207/S15328023TOP3002_13.

Fischer, C., Pardos, Z., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., … Warschauer, M. (2020). *Mining big data in education: Affordances and challenges. Review of Research in Education*, (p. 44).

Gasevic, D., Jovanovic, J., Pardo, A., & Dawson, S. (2017). Detecting learning strategies with analytics: Links with self-reported measures and academic performance. *Journal of Learning Analytics, 4*(2), 113–128. https://doi.org/10.18608/jla.2017.42.10.

Goda, Y., Yamada, M., Kato, H., Matsuda, T., Saito, Y., & Miyagawa, H. (2015). Procrastination and other learning behavioral types in e-learning and their relationship with learning outcomes. *Learning and Individual Differences, 37,* 72–80. https://doi.org/10.1016/j.lindif.2014.11.001.

Goldstein, D., Hahn, C. S., Hasher, L., Wiprzycka, U. J., & Zelazo, P. D. (2007). Time of day, intellectual performance, and behavioral problems in morning versus evening type adolescents: Is there a synchrony effect? *Personality and Individual Differences, 42*(3), 431–440. https://doi.org/10.1016/j.paid.2006.07.008.

Guo, P. J., & Reinecke, K. (2014). Demographic differences in how students navigate through MOOCs. In *Proceedings of the First ACM Conference on Learning @ Scale Conference*, (pp. 21–30). https://doi.org/10.1145/2556325.2566247.

Häfner, A., Oberst, V., & Stock, A. (2014). Avoiding procrastination through time management: An experimental intervention study. *Educational Studies, 40*(3), 352–360. https://doi.org/10.1080/03055698.2014.899487.

Hahn, C., Cowell, J. M., Wiprzycka, U. J., Goldstein, D., Ralph, M., Hasher, L., & Zelazo, P. D. (2012). Circadian rhythms in executive function during the transition to adolescence: The effect of synchrony between chronotype and time of day. *Developmental Science, 15*(3), 408–416. https://doi.org/10.1111/j.1467-7687.2012.01137.x.

Jaggars, S. S. (2011). *Online learning: Does it help low-income and underprepared students? (CCRC brief no. 52)*. New York: Community College Research Center, Columbia University.

Jenkins, D., & Rodriguez, O. (2013). Access and success with less: Improving productivity in broad-access postsecondary institutions. *The Future of Children, 23*(1), 187–209. https://doi.org/10.1353/foc.2013.0000.

Kazerouni, A. M., Edwards, S. H., Hall, T. S., & Shaffer, C. A. (2017). DevEventTracker: Tracking development events to assess incremental development and procrastination. In *Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education*, (pp. 104–109). https://doi.org/10.1145/3059009.3059050.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. *Computers & Education, 104,* 18–33. https://doi.org/10.1016/j.compedu.2016.10.001.

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

Levy, Y., & Ramim, M. M. (2013). An experimental study of habit and time incentive in online-exam procrastination. In *Proceedings of the Chais Conference on Instructional Technologies Research*, (pp. 53–61).

Lewis, B. A., MacEntee, V. M., DeLaCruz, S., Englander, C., Jeffrey, T., Takach, E., & Woodall, J. (2005). Learning management systems comparison. In *Proceedings of the 2005 Informing Science and IT Education Joint Conference*, (pp. 17–29).

Li, Q., Baker, R., & Warschauer, M. (2020). Using clickstream data to measure, understand, and support self-regulated learning in online courses. *The Internet and Higher Education*, 45. https://doi.org/10.1016/j.iheduc.2020.100727.

Lim, J. M. (2016). Predicting successful completion using student delay indicators in undergraduate self-paced online courses. *Distance Education*, 37(3), 317–332. https://doi.org/10.1080/01587919.2016.1233050.

Lykourentzou, I., Giannoukos, I., Nikolopoulos, V., Mpardis, G., & Loumos, V. (2009). Dropout prediction in e-learning courses through the combination of machine learning techniques. *Computers & Education*, 53(3), 950–965. https://doi.org/10.1016/j.compedu.2009.05.010.

Macan, T. H., Shahani, C., Dipboye, R. L., & Phillips, A. P. (1990). College students' time management: Correlations with academic performance and stress. *Journal of Educational Psychology*, 82(4), 760–768. https://doi.org/10.1037/0022-0663.82.4.760.

Michinov, N., Brunot, S., Le Bohec, O., Juhel, J., & Delaval, M. (2011). Procrastination, participation, and performance in online learning environments. *Computers & Education*, 56(1), 243–252. https://doi.org/10.1016/j.compedu.2010.07.025.

Misra, R., & McKean, M. (2000). College students academic stress and its relation to their anxiety, time management, and leisure satisfaction. *American Journal of Health Studies*, 16(1), 41–51.

Mobasher, B. (2007). Data mining for web personalization. In *The adaptive web*, (pp. 90–135). Berlin, Heidelberg: Springer.

Nussbaumer, A., Steiner, C., & Albert, D. (2008). Visualisation tools for supporting self-regulated learning through exploiting competence structures. In *Proceedings of the International Conference on Knowledge Management (IKNOW 2008)*, (pp. 3–5). Graz: Verlag der Technischen Universität Graz.

Pardos, Z. A., Tang, S., Davis, D., & Le, C. V. (2017). Enabling real-time adaptivity in MOOCs with a personalized next-step recommendation framework. In *Proceedings of the Fourth (2017) ACM Conference on Learning @ Scale*, (pp. 23–32). https://doi.org/10.1145/3051457.3051471.

Park, J., Denaro, K., Rodriguez, F., Smyth, P., & Warschauer, M. (2017). Detecting changes in student behavior from clickstream data. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, (pp. 21–30). Vancouver: ACM Press.

Park, J., Yu, R., Rodriguez, F., Baker, R., Smyth, P., & Warschauer, M. (2018). Understanding student procrastination via mixture models. In *Paper presented at the 11th International Conference on Educational Data Mining (EDM), Raleigh, North Carolina*.

Paul, J. A., Baker, H. M., & Cochran, J. D. (2012). Effect of online social networking on student academic performance. *Computers in Human Behavior*, 28(6), 2117–2127. https://doi.org/10.1016/j.chb.2012.06.016.

Perry, N. E., & Winne, P. H. (2006). Learning from learning kits: gStudy traces of students' self-regulated engagements with computerized content. *Educational Psychology Review*, 18(3), 211–228. https://doi.org/10.1007/s10648-006-9014-3.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, 16(4), 385–407. https://doi.org/10.1007/s10648-004-0006-x.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33–40. https://doi.org/10.1037/0022-0663.82.1.33.

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, 53(3), 801–813. https://doi.org/10.1177/0013164493053003024.

Rodriguez, F., Yu, R., Park, J., Rivas, M. J., Warschauer, M., & Sato, B. K. (2019). Utilizing learning analytics to map students' self-reported study strategies to click behaviors in STEM courses. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, (pp. 456–460). https://doi.org/10.1145/3303772.3303841.

Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*, 2(1), 7–12. https://doi.org/10.18608/jla.2015.21.2.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, 33(1), 135–146. https://doi.org/10.1016/j.eswa.2006.04.005.

Schellings, G., & Van Hout-Wolters, B. (2011). Measuring strategy use with self-report instruments: theoretical and empirical considerations. *Metacognition and Learning*, 6(2), 83–90. https://doi.org/10.1007/s11409-011-9081-9.

Seo, E. H. (2013). A comparison of active and passive procrastination in relation to academic motivation. *Social Behavior and Personality: An International Journal*, 41(5), 777–786. https://doi.org/10.2224/sbp.2013.41.5.777.

Shi, C., Fu, S., Chen, Q., & Qu, H. (2015). VisMOOC: Visualizing video clickstream data from massive open online courses. In *Proceedings of the 8th IEEE Pacific Visualization Symposium*, (pp. 159–166). https://doi.org/10.1109/PACIFICVIS.2015.7156373.

Sitzmann, T., & Johnson, S. K. (2012). The best laid plans: Examining the conditions under which a planning intervention improves learning and reduces attrition. *Journal of Applied Psychology*, 97(5), 967–981. https://doi.org/10.1037/a0027977.

Spiliopoulou, M. (2000). Web usage mining for web site evaluation. *Communications of the ACM*, 43(8), 127–134.

Trueman, M., & Hartley, J. (1996). A comparison between the time-management skills and academic performance of mature and traditional-entry university students. *Higher Education*, 32(2), 199–215. https://doi.org/10.1007/BF00138396.

van Den Hurk, M. (2006). The relation between self-regulated strategies and individual study time, prepared participation and achievement in a problem-based curriculum. *Active Learning in Higher Education*, 7(2), 155–169. https://doi.org/10.1177/1469787406064752.

van Eerde, W. (2003). Procrastination at work and time management training. *The Journal of Psychology*, 137(5), 421–434. https://doi.org/10.1080/00223980309600625.

Whitehill, J., Williams, J. J., Lopez, G., Coleman, C. A., & Reich, J. (2015). Beyond prediction: First steps toward automatic intervention in MOOC student stopout. In *Proceedings of the 8th international conference on educational data mining*. Madrid.

Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational Psychologist*, 45(4), 267–276. https://doi.org/10.1080/00461520.2010.517150.

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology, 27*(4), 551–572. https://doi.org/10.1016/S0361-476X(02)00006-1.

You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning. *The Internet and Higher Education, 29*, 23–30. https://doi.org/10.1016/j.iheduc.2015.11.003.

Yu, R., Jiang, D., & Warschauer, M. (2018). Representing and predicting student navigational pathways in online college courses. In *Proceedings of the 5th ACM Conference on Learning at Scale (L@S)*.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.