Patterns of engagement in a flipped undergraduate class: Antecedents and outcomes

[1]Rosenberg, J., [1]Lee, Y., [1]Robinson, K., [2]Ranellucci, J., [1]Roseth, C., & [1]Linnenbrink-Garcia, L.

Michigan State University, Hunter College

Author note. Paper to be presented at the Annual Meeting of the American Educational Research Association, New York, NY. Correspondence regarding this manuscript may be directed to Joshua Rosenberg (jrosen@msu.edu).

Abstract

The flipped classroom is a promising and increasingly popular instructional approach used in STEM courses to promote active engagement and deeper learning (Gross, Pietri, Anderson, Moyano-Camihort, & Graham, 2017)). However, it is assumed that successful flipping depends on students' engagement with online lecture materials, including the frequency and duration of their video views. To better understand how students engaged with online lecture videos in flipped classrooms, patterns of engagement with online video lectures in an undergraduate anatomy course were examined (*N* = 272). Focusing on the time period between first and second exams, a mixed effects modeling approach was used to estimate trajectories of students' engagement, which showed a slight decline between the first and second exams with a sharp increase just prior to the second exam. Person-specific estimates were used to examine the relations of antecedents (perceived competence, cost) and outcomes (exam 2 performance, final exam performance, course grade) to the trajectories. Neither perceived competence nor effort cost significantly predicted students' engagement; however, the findings for effort cost showed a trend suggesting that high effort cost was associated with declines in engagement and then a sharp increase just prior to the second exam. Students' initial levels of engagement were positively associated with higher achievement; a sharp increase in viewing just prior to exam was associated with lower achievement.

Lecture classes are common, particularly at the post-secondary level (Mazur, 2009; Smith & Valentine, 2012). Faculty and instructors at the post-secondary level, however, have targeted the lecture as a context with the potential to be rethought (Freeman et al., 2014; Mazur, 2009). Recent pedagogical trends (with a focus on student-centered learning) and technological trends (with a focus on learning and communicating with digital tools, such as through learning management systems and social media online communities) have made it possible--and often desirable--for instructors to video-record and assign or make available to students recordings of their lectures or their content (Baepler, Walker, & Driessen, 2014; Gross, et al., 2017). Students viewing recordings of lectures, in turn, provides for time in-class to be used for other purposes such as discussion, inquiry-based activities, and feedback from instructors. This shift in instruction, in which students view video-recordings of lectures outside of class and engage in other activities in class, is commonly described as a *flipped class*.

The flipped class has become an increasingly common design, but research on impacts of flipping is quite limited. One key question for research to answer about flipped classes concerns students' participation in outside of class activities such as viewing online lectures. This is a key consideration because it is a central feature of the flipped class: students now must learn core content outside of class that was previously delivered via in-class lectures. Moreover, understanding students' engagement outside of class is particularly important given prior research that highlights the importance of student engagement as a key determinant of learning and academic achievement (e.g., Christenson, Reschly, & Wylie, 2012; Fredricks, Blumenfeld, & Paris, 2004; Sinatra, Heddy, & Lombardi, 2015).

Thus, the present study investigates students' viewing of online video lectures, which we consider as a form of behavioral engagement (Fredricks et al., 2004). Specifically, we: (1)

examine patterns of engagement using longitudinal growth models and examine potential antecedents (perceived competence, effort cost) and outcomes (course-specific achievement) of these longitudinal patterns of engagement. Below, we begin by providing a brief overview of our theoretical framework for studying outside of class engagement, namely expectancy-value theory. Next, we briefly review the relatively small extant literature on "flipped classrooms" and their impact on students' engagement. Finally, we provide the context for our use of online lecture views as an indicator of behavioral engagement, reviewing the use of intensive data (e.g., datasets with multiple repeated measurements of individuals' behaviors) to measure students' engagement.

Theoretical Background

We utilized Eccles and Wigfield's (Eccles, 1983; Wigfield & Eccles, 2000) modern expectancy-value theory as the theoretical framework for studying students' engagement in flipped classrooms. Expectancy-value theory posits that the most proximal predictors of students' achievement-related choices are their expectations of success (*Can I do this?*) and subjective task values (*Why do I want to do this?*). That is, student success depends on how much they believe they will succeed and the level of value students' associate with this success. Modern expectancy-value theory includes four different components of task value: (a) intrinsic or interest value, the enjoyment one derives from performing a task; (b) attainment value, the importance of succeeding in a task for one's identity; (c) utility value, the usefulness of performing a task for one's future goals; and (d) perceived costs, the anticipated consequences associated with a task.

Generally, expectancies are stronger predictors of academic achievement and task values are stronger predictors of academic choices and engagement or persistence (Wigfield &

Cambria, 2010). Furthermore, there is a growing body of evidence that perceived costs play an important role in both academic achievement and choice/persistence behaviors (Barron & Hulleman, 2015; Battle & Wigfield, 2003; Flake et al., 2015; Perez et al., 2014; Trautwein et al., 2012). Given the relative dearth of research on cost relative to the other task values and the potentially strong role that costs may play in students' decisions to engage outside of class (e.g., by viewing the online lectures), we focused our analyses on expectancies and perceived effort costs. Effort cost consists of the perceived amount of time, effort, or the workload necessary to engage in the activity (Flake et al., 2015). We expected that students' perceptions of high effort costs associated with the flipped classroom would predict their subsequent learning behavior, as a flipped classroom could be perceived to require extra work due to video lectures watched outside of class and more student-led activities during class time.

Prior Research on Flipped Instruction

Online video lectures are a popular instructional tool that can be incorporated into traditional face-to-face, online, or flipped courses (Lo, Lie, & Hew, 2018; O'Flaherty & Phillips, 2015). While the research on flipped classrooms is relatively scarce, there are a few rigorous research studies, including those that make use of an experimental design, to shed light on the potential benefits of flipped instruction. For instance, a recent field experiment conducted by He, Holton, Farkas, and Warschauer (2016) compared student engagement and achievement between introductory chemistry courses taught using a flipped version versus a traditional lecture format. The authors found that students studied more before class in the flipped classroom and had higher achievement, with the largest achievement effects observed for the midterm. A subsequent study found similar benefits of a flipped versus traditional chemistry course in terms of student achievement in subsequent chemistry courses (Casasola, Nguyen, Warschauer, & Schenke, 2017).

Although research suggests that video lectures promote learning and achievement outcomes (e.g., Chen & Wu, 2015), for video lectures to be an effective instructional tool, it is necessary for students to actually watch and pay attention to the video content. What still remains unclear is the way the flipped instruction positively relates to students' academic achievement; thus, the current study focuses on students' behavior of viewing video lectures over time to better understand how specific forms of engagement may explain why and how flipped classroom are beneficial for achievement.

Using Intensive Data to Study Engagement

One of the challenges of studying engagement in the flipped context is that it is difficult to accurately and consistently assess engagement that occurs outside of the classroom. To address this challenge, we turned to a promising new approach to measure engagement that involves analyzing behavioral trace data from students (e.g., log-data tracked in online learning environments, intelligent-tutor systems, or in flipped classrooms; Azevedo, 2015). This approach can be considered as a subset of intensive data approaches (Bolger & Laurencau, 2013). While such approaches commonly use diary or Experience Sampling Method (ESM) methodologies, a number of scholars have also sought to use new, digital sources of data to obtain repeated measures of individuals' behaviors based on intensive data (D'Mello, Dieterle, & Duckworth, 2017). Technology-rich learning environments, such as online microworlds (Gerard, Ryoo, McElhaney, Liu, Rafferty, & Linn, 2015) and intelligent tutors (Gobert, Baker, & Wixon, 2015), provide a valuable context for exploring the use of these new sources of data because such data are often generated as traces left behind through individuals' interactions (Lazer et al., 2009; Snee et al., 2016).

Behavioral trace measures are especially useful when assessing indicators of behavioral engagement, such as physical actions, utterances, body language, or interactions with peers or teachers (see Azevedo, 2015). Importantly, the present study uses trace data of students' online lecture viewing as a behavioral indicator of behavioral engagement. This type of basic engagement with the content is essential for learning gains in any educational context and is especially relevant when students are watching online video lectures. By using online video lecture viewing as an observable proxy of student engagement in a flipped classroom, we advance a unique understanding of student engagement across the whole semester in an increasingly popular learning context. Recent research suggests that there is potentially a relationship between log-trace, and observational, and self-report measures of engagement, though how measures are constructed and how the data they generate are important, and open, issues in the field (Henrie, Bodily, Larsen, & Graham, 2015; Henrie, Bodily, Manwaring, Graham, advance online publication). Using trace data as an indicator of student engagement makes it possible to measure students' actual behavior with minimal biases from respondents' recall limitations, the temporal granularity of experiences, and subjective beliefs (e.g., Black, Dawson, & Priem, 2008; Golder & Macy, 2011; Koskey, Karabenick, Woolley, Bonney, & Dever, 2010; Robinson & Clore, 2002). Furthermore, using students' viewing data to measure behavioral engagement is non-intrusive and enables multiple assessments throughout the whole semester.

The Present Study

The purpose of the current study is to explore patterns of behavioral engagement, as assessed with log-trace data, and the antecedents and outcomes of these patterns in an undergraduate flipped classroom. The specific goals are to (1) describe patterns of online video-lecture views throughout the duration of a flipped anatomy course, (2) investigate cost, value and perceived competence as predictors of pattern of video viewing, and (3) examine the relations between these viewing patterns and three metrics of academic achievement (Exam 2, the cumulative Final Exam, and students' final Course Grade).

Based on expectancy-value theory (Wigfield & Eccles, 2000) and considering the context of a flipped classroom, we hypothesize that students who are more confident in their ability (high perceived competence) to learn the course material and view the costs of engaging in the course (low perceived costs) will be more likely to watch the video lectures online and will subsequently have higher course-specific achievement.

Recent scholarship has also shown that behavioral engagement is strongly associated with achievement-related outcomes at school and in other learning environments (Fredricks, 2016; Sinatra et al., 2015). Although the significance of behavioral engagement in relation to important learning outcomes is well established, to our knowledge, no research has examined longitudinal fluctuations in behavioral engagement and its relation to academic achievement over a full semester and across the academic week. Investigating the trajectory of behavioral engagement can provide a better understanding of changes in students' learning behavior and its relation to academic outcomes that would not have been captured by a correlational design at one time point. Furthermore, persistence or continuous engagement in a flipped classroom context is particularly important in that this innovative learning environment requires that students watch online lectures regularly for class preparation. This line of research is critical given the increased popularity of flipped classrooms paired with minimal empirical understanding of the underlying mechanisms and effectiveness of this new instructional approach.

Method

Participants and Procedures

Participants were 272 undergraduates (75% female, 85% Caucasian) taking a large introductory anatomy course at a Midwestern public university in Spring 2015. The course employed a "flipped" design, such that students watched video lectures at home and participated in active learning activities during class. Students were assigned 58 videos total over the 15-week semester; the average length of each video was 12.62 ($SD = 5.47$) minutes. Students were also assigned readings from an anatomy textbook, however students were not graded for work outside of class, nor were they penalized if they did not watch the assigned videos. In-class activities largely consisted of small-group work with short (1-5 minutes) review lectures to reinforce material learned outside of class. Small-group activities included problem solving, identification/recall exercises, and case study problems.

Students completed questionnaires assessing their motivation (perceived competence and perceived effort cost) as part of a larger study. The survey was administered several weeks into the semester and five days prior to the first exam. Students received course credit (up to 3% of the total points available in the course) for completing the survey and provided informed consent for participating in the study. The course instructor did not see students' survey responses, and students received course credit for survey completion regardless of whether they agreed to participate in the study. As described in greater detail below, log files of online lecture views were collected throughout the semester. Exam scores and final grade were collected directly from the course instructor at the end of the semester.

Measures

We created measures of students' behavioral engagement using log files of online lectures. Student motivation was assessed with self-report items from the questionnaire Academic achievement was measured using exam and course grades. Below, we provide additional data about each of these measures.

Behavioral engagement. Log files of students' lecture-video views were collected directly from the Learning Management System (LMS). Person-specific data included time-stamped records indicating which videos each student watched, when videos were watched, and how many minutes each student watched the online lecture videos. These data were processed into one data frame, which consisted of 40,668 unique log entries. Then, these log entries were aggregated by day, so that the sum of the number of minutes students viewed on each day of the semester was calculated for every student. A value of 0 minutes was recorded for any days in which students viewed no videos. This resulted in a data frame with a row for each student by day combination, for a total of 30,303 rows.

Motivation. We collected self-report measures of the two motivational predictors. Specifically, *perceived competence* (5 items; Midgley et al., 2000; alpha = .87) assessed students' beliefs that they would be able to learn the course material. A sample item for perceived competence was "I'm certain I can master the skills taught in anatomy." *Effort cost* (5 items; Conley, 2012; alpha = .61) assessed students' beliefs about the excessive time and effort needed to be successful in the course. A sample item was "Studying anatomy requires more effort than I'm willing to put in." All items were assessed using Likert-type scales from 1 to 5, with 1 = *strongly disagree* and 5 = s*trongly agree*. Items for perceived competence and effort cost were averaged to create composite scores for each variable. High ratings for perceived competence indicate high levels of competence. High ratings for effort cost reflect students' beliefs that the course required more effort than they we were willing to exert.

Academic achievement. Three measures of academic achievement were used: one proximal (Exam 2) and two distal (final exam score and final course grade). Exam 2 reflected the material learned in the first few weeks of class, and the final exam grade reflected the material learned throughout the semester. The final course grade included all semester exams and class assignments. The three outcomes were standardized ($M = 0$, $SD = 1$) to facilitate comparisons of results with respect to the three outcomes, however unstandardized scores were used to compute descriptive statistics and correlations.

Data Analysis

To understand students' patterns of video views over time, we specified a series of mixed effects growth models using the Linear and Nonlinear Mixed Effects Models (nlme) package (Pinheiro, Bates, DebRoy, Sarkar, and R Core Team, 2018) using R version 3.4.3 (R Core Team, 2018). Mixed effects models are commonly viewed as multilevel or hierarchical linear models (Raudenbush & Bryk, 2002), and these models can account for repeated measures nested within students. The models, variations of which are specified first as part of a "top-down" strategy methodologists describe as useful for when a specific functional form can be identified from descriptive plots of the data (or can be tested given theory or past research; West et al., 2015), is as follows, for person *j* and engagement wave *i*:

Models 1A-1C:

$$Y_{\text{minutes-viewed-}i} = \beta_{00} + \text{linear-term}_i * \beta_1 + \text{linear-slope}_i * \beta_2 + \text{quadratic-term}_i * \beta_3 + \varepsilon_j$$

$$\beta_{00} = \beta_{00} + \text{person-specific-intercept-effect}_j * \mu_1$$

$\beta_1 = \beta_{10}$ + person-specific-linear-slope-effect$_j$*μ_2

$\beta_2 = \beta_{20}$ + person-specific-quadratic-term-effect$_j$*μ_3

We saved person-specific estimates of video view intercepts, and linear slopes, and quadratic growth terms from the mixed effects model. These estimates are the sum of the fixed effects estimate (β_1, β_2, and β_3) in the model above plus the person-specific predictions (μ_1, μ_2, and μ_3); these are commonly termed *Best Linear Unbiased Predictors* (BLUPs; West et al., 2015), and are similar to the estimates that would be obtained from estimating a regression model for each student, but they are less biased because the outcome is considered to be nested within each student. These person-specific estimates for the intercept, linear slopes, and quadratic growth terms then used these growth parameter estimates as *outcome variables* in three separate models to examine motivation as predictors of person-specific trajectories. For these models, person-specific intercepts, linear slopes, and quadratic terms were predicted by students' effort cost and perceptions of competence:

Model 2A: $Y_{\text{person-specific-intercept-j}} = \beta_0 + pc_j{*}\,\beta_1 + cost_j{*}\,\beta_2 + \varepsilon_j$

Model 2B: $Y_{\text{person-specific-linear-term-j}} = \beta_0 + pc_j{*}\,\beta_1 + cost_j{*}\,\beta_2 + \varepsilon_j$

Model 2C: $Y_{\text{person-specific-quadratic-term-j}} = \beta_0 + pc_j{*}\,\beta_1 + cost_j{*}\,\beta_2 + \varepsilon_j$

We also used the person-specific estimates *as three predictor variables* in three separate models predicting academic achievement. For these models, person-specific intercepts, linear slopes, and quadratic terms were used to predict students' Exam 2 and final exam scores and their final grade:

Model 3A:

$Y_{exam\text{-}2\text{-}j} = \beta_0$ + person-specific-intercept$_j$* β_1 + person-specific-linear-term$_j$* β_2 + person-specific-quadratic-term$_j$* $\beta_3 + \varepsilon_j$

Model 3B:

$Y_{final\text{-}exam\text{-}j} = \beta_0$ + person-specific-intercept$_j$* β_1 + person-specific-linear-term$_j$* β_2 + person-specific-quadratic-term$_j$* $\beta_3 + \varepsilon_j$

Model 3C:

$Y_{final\text{-}grade\text{-}j} = \beta_0$ + person-specific-intercept$_j$* β_1 + person-specific-linear-term$_j$* β_2 + person-specific-quadratic-term$_j$* $\beta_3 + \varepsilon_j$

For this reason, the part of data analysis for estimating growth models can be considered a *within-person* analysis (Murayama, Goetz, Malmberg, Pekrun, Tenaka, & Martin, 2017) in that repeated measures over time are used to estimate person-specific trajectories which are then related to antecedents (i.e., motivation) and outcomes (i.e., academic achievement) in a more conventional between-persons analytic procedure.

Results

Preliminary Results

Prior to our primary analyses, we created descriptive plots of the minutes of *all* videos viewed per day for the semester in order to gain a better sense of students' behavioral engagement (online lecture views) throughout the semester and the nature of this rather unique dataset (see Figure 1). We noted trends in engagement associated with the timing of the three course exams (Exams 1-3) and the Final Exam with online lecture views rising and peaking just before each exam (particularly for Exam 2, Exam 3, and the Final Exam). Moreover, there was a great deal of variability in patterns across the entire semester. In order to better analyze these complex patterns, we focused our analyses on the period between the first and second exam, which had a duration of 26 days. Focusing on this period, as opposed to the period before Exam 1, allowed us to examine students' video-viewing behavior pattern after they had become used to this new instructional approach, reducing the potential influence of its novelty. This approach also allowed us to examine perceived competence and cost (measured six days before the first video log after the first exam) as predictors of both initial video views (intercept) and the change in viewing patterns over time (slope). In future research, we plan to test whether our approach for this initial engagement period can be applied to time sequences later in the semester and to examine whether the findings reveal a similar pattern.

Using this period of engagement between Exams 1 and 2 as the focus for the present study, we then sought to summarize the 26 days of data into a fewer number of time points (waves) by plotting and examining two to ten waves to reflect various configurations of the time points (see Appendix A). Our goal was to select the minimum number of waves (to simplify the analysis) that reflected the underlying patterns we observed and were substantively meaningful given our knowledge of the structure of the class and its assessments. That is, because of the importance of the exam to students' grades, we expected an increasing growth rate prior to the exam as students prepared for it. Based on our theoretical expectation and the identification of what appeared to be a quadratic growth trajectory in the descriptive plots (Figure 1) and the plots with each of the numbers of waves (Appendix A), we selected the five-wave configuration to group the time points as the minimum configuration that sufficiently captured the descriptive variation in the data. The five-wave configuration, named *Engagement Wave 1-5*, was the most parsimonious configuration that appeared to effectively capture the trend in increasing views in advance of the exam (at the conclusion of the last wave) observed in the plots.

Using this five-wave approach, we next examined the correlations and descriptive statistics for all of the study variables (see Table 1). Perceived competence (M = 4.25, SD = 0.54), measured on a 5-point Likert scale, was high indicating that students felt competent in their ability to learn the course material. Perceived effort cost (M = 2.36, SD = 0.71) was moderate-low suggesting that students perceived there to be a relatively low cost of the anatomy class for them. In terms of the engagement waves, the values across the five waves indicate increases in the number minutes that students watched the online videos as the second exam approached along with increases in variability.

The correlations indicate small to moderate relations between the effort cost (rs = -.18 to -.20) and perceived competence measures and the measures of achievement (rs = .23 to .28), and small relations with the measures of behavioral engagement across all five waves (i.e., amount of videos watched (in minutes); effort cost: rs = -.13 to .05; perceived competence: rs = -.05 to .04). The three achievement measures were strongly correlated with one another (rs = .65 to .85), and the measures of behavioral engagement were moderately (and highly variable) correlated with the measures of achievement (rs = .05 to .36).

Mixed Effects Growth Models

Our first goal in conducting this research was to provide a descriptive analysis of the trends of outside of class behavioral engagement using mixed effects models. Towards this end, we tested a series of models (see Table 2). Following West, Welch, and Galecki's (2015) recommendations, we started with fitting a full model, a quadratic growth model. We chose this model on the basis of the descriptive plot and the assumption that the growth rate of students' video views would be increasing towards an exam. First, we fit a model with a “loaded mean structure,” or with fixed and random effects for each of the coefficients of the model (Model 1A). Specifically, this model included estimates of both a fixed (i.e., population-level) and random (i.e., varying by student) intercepts, linear coefficients, and quadratic coefficients. Next, to determine whether person-specific variability at the quadratic level was present, we fit a model without this random effect (Model 1B), and then compared the log-likelihood values and information criteria (AIC and BIC) between Models 1A and 1B. The results suggested that the model with the person-specific variability in the quadratic coefficients (Model 1A) fit better than the model without it (Model 1B). Finally, we varied specifications for the residuals, namely, first-order autoregressive covariance structure that structures the residuals to reflect the correlation between time points nearer together in time (Brady, West, & Galecki, 2015).

The model with the first-order autoregressive residual structure (Model 1Ai) exhibited superior fit compared to the model without this structure (Model 1A). In addition to comparing the models based on the information criteria and likelihood-ratio tests, we also plotted the fitted model (see Figure 2). On the basis of the fit indices and bolstered by the interpretability of the fitted values (indicated from the plot), we selected model 1Ai, specifying quadratic change and autoregressive paths between video view waves as the best-fitting model (see Table 3). Overall, the model-implied trajectory indicated that, on average, students’ patterns of viewing were somewhat flat with a slight decline toward the beginning of the time period and gradually increased in advance of Exam 2 (see Figure 2), with large variability in patterns of viewing between students. Specifically, the intercept of video views was 34.55 minutes (SE = 4.08, p < .001), with a large random effect standard deviation of 34.83 minutes. These results indicate that students watched about 35 minutes of video during (approximately) the first five days with a large degree of person-specific variation in the intercept. The fixed linear slope (-16.09, *SE* = 3.58, $p < .001$) indicates that across all students, there is a decreasing pattern of engaging (accounting for the value of the intercept and quadratic term) over time. There was also a significant fixed quadratic effect (3.85, *SE* =0.66, $p < .001$), indicating that patterns of viewing videos were not linear but rather showed an increased rate of growth prior to Exam 2 (see Figure 2;). Noteworthy, the partial R^2 value for the quadratic term (partial R^2 = .031) was about twice as large as the value for the linear term (partial R^2 = .015), suggesting that the quadratic trend may explain students' trajectories to a greater extent than the linear trend, helping to explain why the negative overall slope is not as observed in the fitted values (Figure 2) as the *change* in the rate (i.e., what is indicated by the quadratic growth term) of viewing. Overall, this pattern suggests that students remain relatively stable in their viewing of online videos, with a sharp increase in online video viewing just prior to the second exam. Moreover, there was substantial variability in students' trajectories in terms of the intercept (*SD* = 34.83), linear slope (*SD* = 39.80), and quadratic term (*SD* = 8.34). These values indicate the mean difference of the person-specific estimates from the fixed effects estimates. For the intercept, for example, the person-specific predictions have a mean difference of 51.79 from the fixed effect estimate of 34.55. These values, relative to the value of the residual (*SD* = 28.92) and of the fixed effects highlight the need to consider individual variability in these trajectories.

Motivation as Antecedents of Video View Trajectories

Next, we examined perceived competence and effort cost as predictors of person-specific trajectory estimates from the quadratic growth model; namely, the person-specific intercept, linear slope, and quadratic slope estimates of video views between the first and second exams. These models (see Models 2A, 2B, 2C in Table 4) can be interpreted in terms of the effect of one of the predictors (perceived competence or effort cost) accounting for the independent effect of the other predictor upon person-specific initial levels (intercept) and linear and quadratic growth terms.

As shown in Table 4, there were not statistically significant relations between the antecedents (perceived competence and effort cost) and the person-specific trajectory estimates. Somewhat surprisingly, perceived competence was not significantly related to the intercept, linear growth, or quadratic growth terms for behavioral engagement (i.e., video viewing) suggesting that students' perceptions of their ability to learn anatomy were not associated with their engagement online video lecture viewing patterns. Effort cost was also not significantly related to the intercept, linear slope (B = -7.03, SE = 3.77, p = .063), or the quadratic growth term (B = 1.37, SE = 0.81, p = .088); however, it was marginally related to both growth terms. The trends suggested that students who perceived that the anatomy course required more work than they were willing to give were somewhat less likely to watch the videos over time, but then showed an increase in viewing just prior to the second exam.

Taken together, this pattern of results suggests that both perceived competence and effort cost were not significantly related to the engagement pattern in a flipped classroom context, though effort cost showed a weak tendency of the relation to the engagement pattern but did not reach conventional levels of statistical significance ($p < .10$).

Academic Achievement as Outcomes of Video View Trajectories

Our third research question investigated how patterns of behavioral engagement between the first and second exams predicted student performance on the second exam as well as their overall performance in the course, as indicated by the final exam and final course grade. In these models, we use the person-specific trajectory estimates to predict Exam 2, Final Exam, and Final Grade scores.

As shown in Table 5, initial levels of behavioral engagement (e.g., intercept) positively predicted Exam 2 performance ($B = 0.13$, $SE = 0.038$, $p < .001$). Specifically, for each one-minute more of video watching during the initial wave, there was an associated 0.13 standard deviation increase in Exam 2 scores suggesting that students who watched more videos at the very start of the period of time after the first exam did better on the second exam. Similar patterns were observed for the effect of initial levels of video watching with respect to the Final Exam ($B = 0.09$, SE = 0.038, $p = .019$) and Final Grade ($B = 0.10$, SE = 0.03), $p = .004$). The person-specific linear term was not found to be associated with any of the outcomes, indicating that students' linear rate of change in video watching was not significantly related to any the measures of academic achievement. The person-specific quadratic growth term was negatively and significantly related to Exam 2 achievement ($B = -1.03$, $SE = 0.46$, $p = .027$), indicating that a sharp increase in the rate of video watching near Exam 2 were associated with *lower* achievement. This suggests that a dramatic increase in time spent watching video lectures right before Exam 2 was not beneficial in terms of Exam 2 performance. The quadratic growth term was not significantly related to Final Exam or Final Grade.

Discussion

As more instructors move to a flipped classroom format, it is important to understand the degree to which students engage with a key element of this instructional design, namely, the actual watching of the assigned online lecture videos. Instructors may assume that students are completing these assignments, but there is limited empirical evidence to provide evidence that students are actually watching online lectures. Wang (2017), for example, examined engagement on the basis of student behavior in the flipped class, but did not use data on their viewing of video-recordings, instead measuring specific actions such as reading or writing content in the course. Boeve, Meier, Bosker, Vugteveen, Hoekstra, and Albers (2017) *asked* students to report their study habits using a diary method, finding individual differences in students' study habits as well as weak relations between students' self-reported engagement and their acheivement. Thus, the current study provides critical information regarding the nature of online video lecture viewing in a flipped classroom setting. Focusing on the period of time between the first and second exams, we found that a quadratic model--one with an intercept, linear slope (representing change in behavioral engagement as the exam approaches), and quadratic growth term (representing change in the rate of behavioral engagement), fit the data well. Specifically, students' lecture views started at a moderately high level and exhibited a statistically significant pattern of *change* in the rate of viewing over time, suggesting that students view a greater number of minutes of video as the exam approaches. This finding aligns with past research in the literature on time management and procrastination (Chu & Choi, 2005; Crede & Kuncel, 2008). While past research has shown this to be the case with self-report measures, this study is distinct in using a behavioral measure constructed from log-trace data. In this way, these findings align with research in learning analytics (Gerard et al., 2015; Gobert et al., 2015). This study also has the potential to contribute to debates on the alignment of log-trace (and other "objective" or external measures) to "subjective," self-report measures (Henrie et al. 2015; Henrie et al., advance online publication). Moreover, we found substantial individual variability in these trajectories, highlighting the importance of considering person-specific trajectories in our subsequent analyses.

Having identified a suitable model for students' pattern of engagement, we sought to find out whether two motivational constructs, students' effort cost and their perceptions of competence, might predict their behavioral engagement. The relations between effort cost with students' rate of viewing and the change in their rate of viewing (i.e., their linear slopes and quadratic growth terms) did not meet the criterion for statistical significance. However, higher effort cost was negatively associated with students' rate of viewing and positively associated with the change in their rate of viewing, suggesting that students who perceive a class to require a cost in terms of their effort show slight declines in lecture viewing immediately after an exam and then rapidly increase in their lecture views just prior to the exam. Particularly for students who have had past, negative experiences with the flipped classes, required viewing of video recording outside of class may believe that such a class requires an undesirable exertion of effort. Though this relation was not statistically significant, it suggests that future work, perhaps with a different sample or at different times in the semester, may consider the role of effort cost in terms of students' behavioral engagement, identified as an important predictor of achievement in past research from similar groups of undergraduate students (Perez, Cromley, & Kaplan, 2014).

In addition to effort cost, we examined the relations of students' perceptions of competence to their patterns of engagement. We did not find any statistically significant relations between perceptions of competence and students' trajectories of behavioral engagement. This suggests that students view videos for reasons other than how good at the subject or content they are learning about they are. Although no significant relation between perceived competence and patterns of video viewing is unexpected, this finding is not entirely inconsistent with prior work given the expectations from expectancy-value theory. Perceived competence is more strongly associated with academic achievement (Wigfield & Cambria, 2010), whereas perceived costs are associated with academic choice and engagement behaviors as well as academic achievement. Our finding provides implication that in learning environments that involve more work outside classroom, such as a flipped classroom, it is particularly important to help students perceive lower cost associated with their academic work.

In terms of relations of the growth terms to outcomes, we used the estimates of the person-specific trajectories as predictors. We found consistent, positive patterns between person-specific intercepts and the measures of achievement. Having higher initial levels of engagement--accounting for each students' person-specific linear slope and quadratic growth term--was associated with higher achievement. That students who began the period after an exam viewing more videos (and engaging to a greater extent) suggests that such a pattern is adaptive in terms of students' achievement. We did not find relations between the person-specific linear terms and students' achievement, suggesting that, accounting for the effects of the other terms that describe students' trajectories of engagement, neither increasing (nor decreasing) one's rate of viewing is associated with differences in achievement. Students' change in their rate of viewing--indicated by the relation between the quadratic growth term and the measures of achievement--was only statistically significant for Exam 2. This relationship was strong and negative, indicating that the achievement of students whose rate of viewing itself increases is lower. This pattern is characteristic of "cramming" for the exam, and it appears to not be adaptive in terms of students' achievement. The corollary of this finding is that the achievement of students who do not increase their rate of viewing is higher, suggesting that a more consistent pattern of viewing has benefits to students' achievement.

Our results also highlight the importance of studying engagement using growth curve modeling. Specifically, we observed that it was the pattern of watching video lectures over time that predicted students' learning outcomes. In this way, these findings also highlight the benefit of a growth modeling approach to understanding the antecedents and outcomes of students' achievement. While scholars have used mixed effects models with intensive repeated measures data to study students' engagement over time (i.e., Patall, Steingut, Vasquez, Trimble, Pituch, and Freeman's [2018] analysis of data from their daily diary study and Strati, Schmidt, and Shumow's (2017) analysis of student engagement, both in high school science classes), no studies that have taken a growth curve modeling approach to the analysis of such data. Along the same lines, no studies have used a growth curve modeling approach to the study of log-trace data, including that for engagement; though studies have used sophisticated methodological approaches (Gobert et al., 2015), they have focused on the development of *predictive* models, often using machine learning methods, whereas growth curve modeling is not as strong as some machine learning for predicting student outcomes but is highly interpretable. As such, these findings also suggest the potential of this approach for other, closely-related log-trace sources of data, such as that from intelligent tutoring systems, online microworlds, and other technology-enhanced learning environments (Azevedo, 2013; Gerard et al., 2015).

Limitations and Recommendations for Future Research

The use of mixed effects models allows us to account for the nested structure of the data and to investigate person-specific trajectories of engagement. There was substantial person-specific variability relative to the overall, population-level (i.e., associated with the fixed effects estimates) model. While we used a growth model that allowed for person-specific deviations from this population model, other modeling approaches could even better be used to consider the heterogeneity in students' engagement trajectories. One candidate is a growth mixture modeling, which could be used to identify multiple groups of person-specific trajectories; such a model could be used to better explain the data than one overall matter. Acknowledging this potential limitation, it is also possible that--so long as person-specific estimates are made so that individual differences are considered--the pattern of viewing before the exam is best characterized by one, rather than multiple patterns. Another candidate modeling approach is one that treats the trajectory more flexibly, namely models that interpolate the data in a (more) non-linear fashion, such as Generalized Additive Models (GAMs) or spline models. A potential drawback of this approach is that it is often not easy to interpret the form of the model that is estimated, such that we could not as easily interpret students' initial level, rate of viewing, and change in the rate of viewing as for the present models.

While the mixed effects modeling approach was strength of this study, this approach also has some limitations relative to other models. Latent variable models, in particular, address some of the limitations of the mixed effects modeling approach. A benefit of a latent variable modeling approach can be considered in light of the use of person-specific predictions for students' engagement trajectories in the present study. These predictions, or BLUPs, have advantages over fitting a model to each students' data and to fitting one overall model for all of the data. However, they are made with a standard error. This standard error is not accounted for when the person-specific predictions are used in *other models* (i.e., the models for relations between the antecedents and the person-specific intercepts, linear slopes, and quadratic growth terms and the models for relations between the person-specific intercepts, linear slopes, and quadratic growth terms and the outcomes). Accordingly, these relations are likely over-confident (Houslay & Wilson, 2017), suggesting that findings with p-values lower than the criterion for statistical significance by a small amount may be above it were the standard errors accounted for in the modeling. A latent variable model could better account for the uncertainty in the person-specific predictions. Another benefit of such an approach is the presence of global fit indices (i.e., the Root Mean Squared Error of Approximation [RMSEA]), which are included as part of the output from the models we fit.

How engagement was measured is a strength in the current study, as it can be used to measure engagement in the context that students engage and in a fine-grained way, but it is possible that log-trace data do not directly reflect the quality of students' viewing. For instance, students could be concentrating when viewing videos or may be playing them in the background while multitasking (or even may have left the computer to do another task). Some scholars have addressed these concerns by validating log-trace measures of engagement with self-report measures (Henrie et al., 2015; Henrie et al., advance online publication). Others have sought to address some of the limitations of log-trace measures of engagement through coupling their use with machine learning analyses (Baker, 2016; Berland et al., 2014). Reducing the potential distracting factors for measuring behavioral engagement based on log-trace data would be a continuous challenge that should be addressed in future research.

A final limitation concerns how cost was conceptualized and used in this study. Scholars have recently pointed out the multi-dimensional nature of cost beliefs. While we examined effort cost, or students' perceptions of how much one needs to put forth to engage in a task, students' outside *effort*, *loss of valued alternatives*, and *emotional cost* may play a role in terms of explaining students' engagement in the flipped class (Flake et al., 2015). Moreover, while we situated the present study in expectancy-value theory and considered students' perceptions of their competence and of effort cost (as one dimension of the broader value construct), there are other types of values that may also impact students' engagement and achievement in the flipped class. Most generally, students' task value may be considered as a predictor of students' engagement in future research.

Future research should explore how consistent this pattern of viewing is in other classes. This can lead to better understanding of how specific features of the flipped class can contribute to different patterns of viewing. In the context of the present study, we could also consider whether this pattern of findings replicates during other time periods in the semester. Other directions for future research concern how individual differences in students impact their viewing patterns. Potential relations may exist between students' initial interest in the flipped or other technology-supported learning environments and students' experiences in such contexts. Relatedly, how students' self-regulation may impact their viewing is an issue that has been raised by other authors (Wandler, & Imbriale, 2017). Finally, how students' engagement compares to their self-reported engagement is a possible question that can be further understood.

Implications for Practice

One key implication of this study for practice concerns viewing patterns. Our findings provide initial evidence regarding the benefits of a more consistent viewing of videos over time. Not surprisingly, students who consistently watched the videos throughout the weeks leading up to the exam had better performance than those who "crammed" by watching multiple videos in the days leading up to the exam. Instructors, too, may consider encouraging more regular viewing and less cramming, at least in the context of the flipped class. For instance, instructors might include additional course structures that encourage viewing throughout the semester (e.g., providing course points for weekly viewing).

Another implication for practice concerns students' initial reasons for taking classes, particularly those that are flipped. Students who perceive the class to require less exertion of effort that incurs a cost to view more videos. Similarly, students who have a higher degree of perceived competence may expect to exhibit higher academic achievement. For these students, support for cost beliefs and, potentially, self-regulated learning strategies may help students to improve their academic achievement.

Conclusion

Recent pedagogical and technological trends have aligned to make it possible for instructors to video-record recordings of their lectures and to use time in class for other purposes. This approach, of "flipping" classes that are traditionally lecture-focused, has become increasingly common at the post-secondary level, but there is limited research on the impacts of flipping (Baepler et al., 2014; Gross et al., 2017). This study examined a key characteristic of the flipped class--the video lecture and student viewing of them, conceptualized as student behavioral engagement, in this out-of-class activity. Using a mixed effects modeling approach and focusing on the period of time before one exam, we found that a model for students' initial levels of engagement, their rate of viewing over time, and the change in their rate of viewing could be used to explain students' trajectories of engagement. We examined motivational antecedents, effort cost and perceived competence, and found that neither significantly predicted any of the person-specific terms for students' pattern of engaging, though there was a trend suggesting that students' effort cost was negatively associated with their rate of viewing and positively associated with their quadratic growth term. We found that students' initial level of engagement was associated with higher achievement. We also found an increase in students' rate of viewing, suggestive of a pattern of viewing more videos just prior to the exam was associated with lower performance on that specific exam but was not related to overall course achievement.

This study has implications for the design of the flipped class and the instruction and support that is provided as part of such a design; namely, that consistently engaging with out-of-class video-recordings of lectures may be more beneficial to students' learning than doing so immediately before important assessments. Moreover, this study has implications for how behavioral engagement is studied in flipped classroom and other technology-supported contexts, as we were able to use and interpret results from the use of such a measure to better understand students' involvement in the flipped class.

References

Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist, 50*(1), 84-94.

Baepler, P., Walker, J. D., & Driessen, M. (2014). It's not about seat time: Blending, flipping, and efficiency in active learning classrooms. *Computers and Education, 78,* 227-236.

Barron, K. E. & Hulleman, C. (2015). Expectancy-value-cost model of motivation. In J. S. Eccles & K. Salmelo-Aro (Eds.), *International encyclopedia of social and behavioral sciences* (2nd ed., pp. 503–509). New York, NY: Elsvier.

Battle, A., & Wigfield, A. (2003). College women’s value orientations toward family, career, and graduate school. *Journal of Vocational Behavior, 62*, 56–75.

Chen, C.-M., & Wu, C.-H. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education, 80*, 108-121.

Black, E. W., Dawson, K., & Priem, J. (2008). Data for free: Using LMS activity logs to measure community in online courses. *The Internet and Higher Education, 11*(2), 65-70.

Boevé, A. J., Meijer, R. R., Bosker, R. J., Vugteveen, J., Hoekstra, R., & Albers, C. J. (2017). Implementing the flipped classroom: an exploration of study behaviour and student performance. *Higher Education, 74*(6), 1015-1032.

Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods*. New York, NY: Guilford.

Casasola, T., Nguyen, T., Warschauer, M., & Schenke, K. (2017). Can flipping the classroom work? Evidence from undergraduate chemistry. *International Journal of Teaching and Learning in Higher Education, 29*(3), 421-435.

Chen, C.-M., & Wu, C.-H. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers &Education, 80*, 108-121.

Christenson, S. L., Reschly, A. L., & Wylie, C. (2012). *The handbook of research on student engagement*. New York: Springer Science.

Chun 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.

Conley, A. M. (2012). Patterns of motivation beliefs: Combining achievement goal and expectancy-value perspectives. *Journal of Educational Psychology, 104*, 32-47.

Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science, 3*(6), 425-453.

D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educational Psychologist, 52*(2), 104-123.

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

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal analysis* (1st ed.). Hoboken, NJ: Wiley.

Flake, J. K., Barron, K. E., Hulleman, C., McCoach, B. D., & Welsh, M. E. (2015). Measuring cost: The forgotten component of expectancy-value theory. *Contemporary Educational Psychology, 41*, 232–244. doi:10.1016/j.cedpsych.2015.03.002

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

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). End of lecture: active learning increases student performance across the STEM disciplines. *Proceedings of the National Academy of Sciences of the United States of America, 111,* 8410-8415.

Gerard, L. F., Ryoo, K., McElhaney, K. W., Liu, O. L., Rafferty, A. N., & Linn, M. C. (2016). Automated guidance for student inquiry. *Journal of Educational Psychology, 108*(1), 60.

Gobert, J. D., Baker, R. S., & Wixon, M. B. (2015). Operationalizing and detecting disengagement within online science microworlds. *Educational Psychologist, 50*(1), 43-57.

Golder, S. A., & Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science, 333*(6051), 1878-1881.

Gross, D., Pietri, E. S., Anderson, G., Moyano-Camihort, K., & Graham, M. J. (2017). Increased preclass preparation underlies student outcome improvement in the flipped classroom. *CBE Life Sciences, 14,* 1-8.

He, W., Holton, A., Farkas, G., & Warschauer, M. (2016). The effects of flipped instruction on out-of-class study time, exam performance, and student perceptions. *Learning and Instruction, 45,* 61-71.

Henrie, C. R., Bodily, R., Larsen, R., & Graham, C. R. (advance online publication) Exploring the potential of LMS log data as a proxy measure of student engagement. *Journal of Computing in Higher Education*, 1-19.

Henrie, C. R., Bodily, R., Manwaring, K. C., & Graham, C. R. (2015). Exploring intensive longitudinal measures of student engagement in blended learning. *The International Review of Research in Open and Distributed Learning, 16(*3).

Koskey, K. L., Karabenick, S. A., Woolley, M. E., Bonney, C. R., & Dever, B. V. (2010). Cognitive validity of students' self-reports of classroom mastery goal structure: What students are thinking and why it matters. *Contemporary Educational Psychology, 35*(4), 254-263.

Lo, C. K., Lie, C. W., & Hew, K. F. (2018). Applying "First Principles of Instruction" as a design theory of the flipped classroom: Findings from a collective study of four secondary school subjects. *Computers & Education, 118,* 150-165.

Mazur, E. (2009). Farewell, lecture. *Science, 323,* 50-51.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., et al. (2000). *Manual for the Patterns of Adaptive Learning Scales (PALS)*. Ann Arbor, MI: University of Michigan.

Murayama, K., Goetz, T., Malmberg, L. E., Pekrun, R., Tanaka, A., & Martin, A. J. *Within-person analysis in educational psychology: Importance and illustrations.* In British Journal of Educational Psychology Monograph Series II: Psychological Aspects of Education—Current Trends: The Role of Competence Beliefs in Teaching and Learning (pp – 71-87). Oxford: Wiley.

O'Flaherty, J., & Phillips, C. (2015). The use of flipped classrooms in higher education: A scoping review. *Internet and Higher Education, 25,* 85-95.

Patall, E. A., Steingut, R. R., Vasquez, A. C., Trimble, S. S., Pituch, K. A., & Freeman, J. L. (2018). Daily autonomy supporting or thwarting and students' motivation and engagement in the high school science classroom. *Journal of Educational Psychology, 110*(2), 269.

Perez, T., Cromley, J. G., & Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *Journal of Educational Psychology, 106*, 315–329.Raudenbush,

Perez, T., Cromley, J. G., & Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. Journal of educational psychology, 106(1), 315.

Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. *Journal of Educational Psychology, 105*, 579-595.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128*(6), 934.

S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage Publications.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist, 50*(1), 1–13. https://doi.org/10.1080/00461520.2014.1002924

Smith, D. J., & Valentine, T. (2012). The use and perceived effectiveness of instructional practices in two-year technical colleges. *Journal on Excellence in College Teaching, 23*, 133-161.

Strati, A. D., Schmidt, J. A., & Maier, K. S. (2017). Perceived challenge, teacher support, and teacher obstruction as predictors of student engagement. *Journal of Educational Psychology, 109*(1), 131.

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of Educational Psychology, 104*, 763–777. doi:10.1037/a0027470

Wandler, J. B., & Imbriale, W. J. (2017). Promoting undergraduate student Self-Regulation in *Online Learning 21*(2), 1-16.

Wang, F. H. (2017). An exploration of online behaviour engagement and achievement in flipped classroom supported by learning management system. *Computers & Education, 114*, 79-91.

West, B. T.,Welch, K. B., & Galecki, A. T. (2015). *Linear mixed models: A practical guide using statistical software* (2nd ed.). Boca Raton, FL: Chapman and Hall/CRC Press.

Wigfield, A., & Cambria, J. (2010). Expectancy-value theory: Retrospective and prospective. In *The decade ahead: Theoretical perspectives on motivation and achievement* (pp. 35–70). Emerald Group Publishing Limited.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68-81.

Tables

Table 1. Correlations and descriptive statistics for all variables.

	Effort Cost	Perceived Competence	Exam 2	Final Exam	Final Grade	Engagement Wave 1	Engagement Wave 2	Engagement Wave 3	Engagement Wave 4	Engagement Wave 5
Effort Cost	1									
Perceived Competence	-0.39	1								
Exam 2	-0.2	0.23	1							
Final Exam	-0.18	0.28	0.72	1						
Final Grade	-0.2	0.25	0.76	0.85	1					
Engagement Wave 1	-0.1	0.04	0.32	0.25	0.3	1				
Engagement Wave 2	-0.07	0.03	0.14	0.22	0.23	0	1			
Engagement Wave 3	-0.13	-0.02	0.17	0.16	0.23	0.18	0.08	1		
Engagement Wave 4	-0.03	-0.05	0.27	0.33	0.36	0.4	0.25	0.26	1	
Engagement Wave 5	0.05	-0.01	0.14	0.05	0.1	-0.01	-0.01	0.15	-0.11	1
M (*SD*)	2.36 (0.614)	4.25 (0.545)	77.736 (13.339)	142.227 (22.666)	999.386 (128.455)	16.36 (25.542)	25.68 (32.14)	20.60 (24.43)	24.414 (30.884)	56.1 (67.421)
n	252	252	256	256	256	272	272	272	272	272

Table 2. Fit statistics for mixed effects growth models.

Model number	Fixed Terms	Random Terms	Residual structure	LL	AIC	BIC	LRT
1A	Intercept, linear slope, quadratic term	Intercept, linear slope, quadratic term	Full	6849.322	13718.64	13770.77	
1B	Intercept, linear slope, quadratic term	Intercept, linear slope	Full	6919.282	13852.57	13889.06	139.92 ($p < .001$)
1Ai	**Intercept, linear slope, quadratic term**	**Intercept, linear slope, quadratic term**	**Auto-regressive**	**6820.379**	**13662.76**	**13720.10**	**57.88 ($p < .001$)**

Note. LL = log-likelihood; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; LRT = Likelihood Ratio Test (for each model, this is for the model in the row before it). The model selected for use in subsequent analyses is in bold.

Table 3. Parameters for model with fixed and random effects for the intercept, slope, and quadratic terms and autoregressive residual structure.

Term	Estimate
Fixed effects	*B (SE)*
Intercept	34.55 (4.08, $p < .001$)
Linear term	-16.09 (3.59, $p < .001$)
Quadratic term	3.85 (0.66, $p < .001$)
Random effects	*SD*
Intercept	34.83
Linear term	39.80
Quadratic term	8.34
Residual	28.92
Correlation structure	*Phi*
First-order auto-regressive (AR1)	-0.33

Table 4. Relations of antecedents to person-specific intercepts, linear slopes, and quadratic slopes.

Estimate (*SE*)	Model 2A: Intercept	Model 2B: Linear slope	Model 2C: Quadratic growth term
Model intercept	21.92 (20.38)	6.934 (23.14)	-0.413 (4.94, p =
Perceived competence	0.78 (3.32, p = .833)	-1.77 (4.25, p = .676)	0.301 (0.909, p = .740)
Effort cost	4.49 (3.32, p = .177)	-7.03 (3.77, p = .063)	1.37 (0.806, p = .088)
R2	.001	.014	.019
F	.971 (p = .381)	1.787 (p = .169)	2.44 (p = .089)

Note. Model intercept refers to the intercept for the model (Models 2A-2C) with the person-specific intercept, linear slope, and quadratic growth terms as outcomes.

Table 5. Relations of growth terms with outcomes

Estimate (*SE*)	Model 3A: Exam 2	Model 3B: Final Exam	Model 3C: Final Grade
Model intercept	-2.31 (0.60)	-1.63 (0.60)	-1.95 (0.58)
Intercept growth term	0.13 (.038, p < .001)	0.09 (.038, p = .019)	0.10 (0.03, p = .004)
Linear growth term	-0.11 (.068, p = .106)	-0.03 (0.06, p = .628)	-0.03 (0.06, p = .570)
Quadratic growth term	-1.03 (.46, p = .027)	-0.52 (0.46, p = .264)	-0.60 (0.45, p = .180)
R2	.148	.140	.198
F	14.60 (p < .001)	13.76 (p < .001)	20.80 (p < .001)

Note. Model intercept refers to the intercept for the model (Models 3A-3C) with the achievement measures (Exam 2, Final Exam, and Final Grade) as outcomes.

Figures

Figure 1. Descriptive plot of minutes of videos viewed per day of the semester by student per day for random sample (50%) of students.

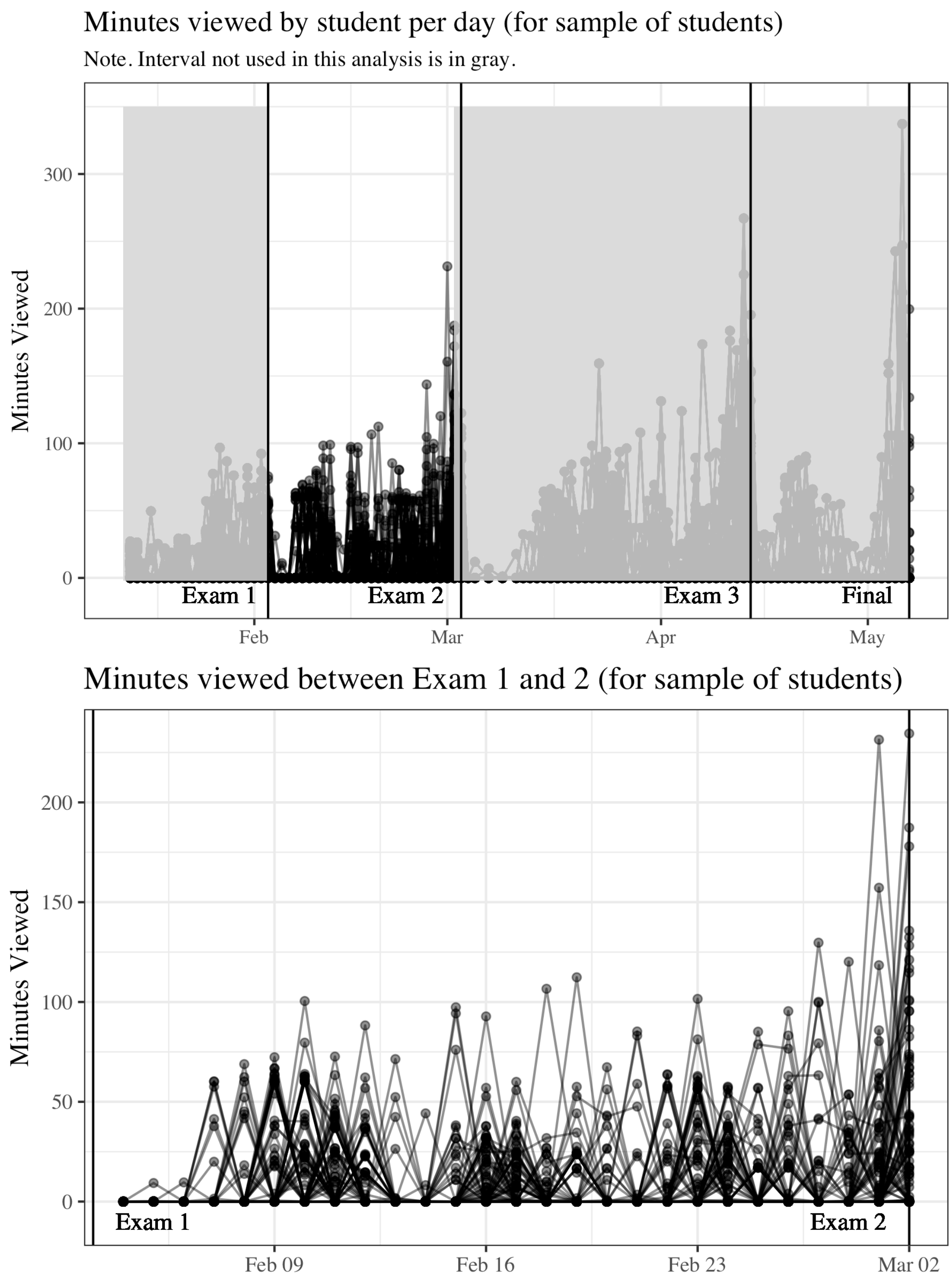
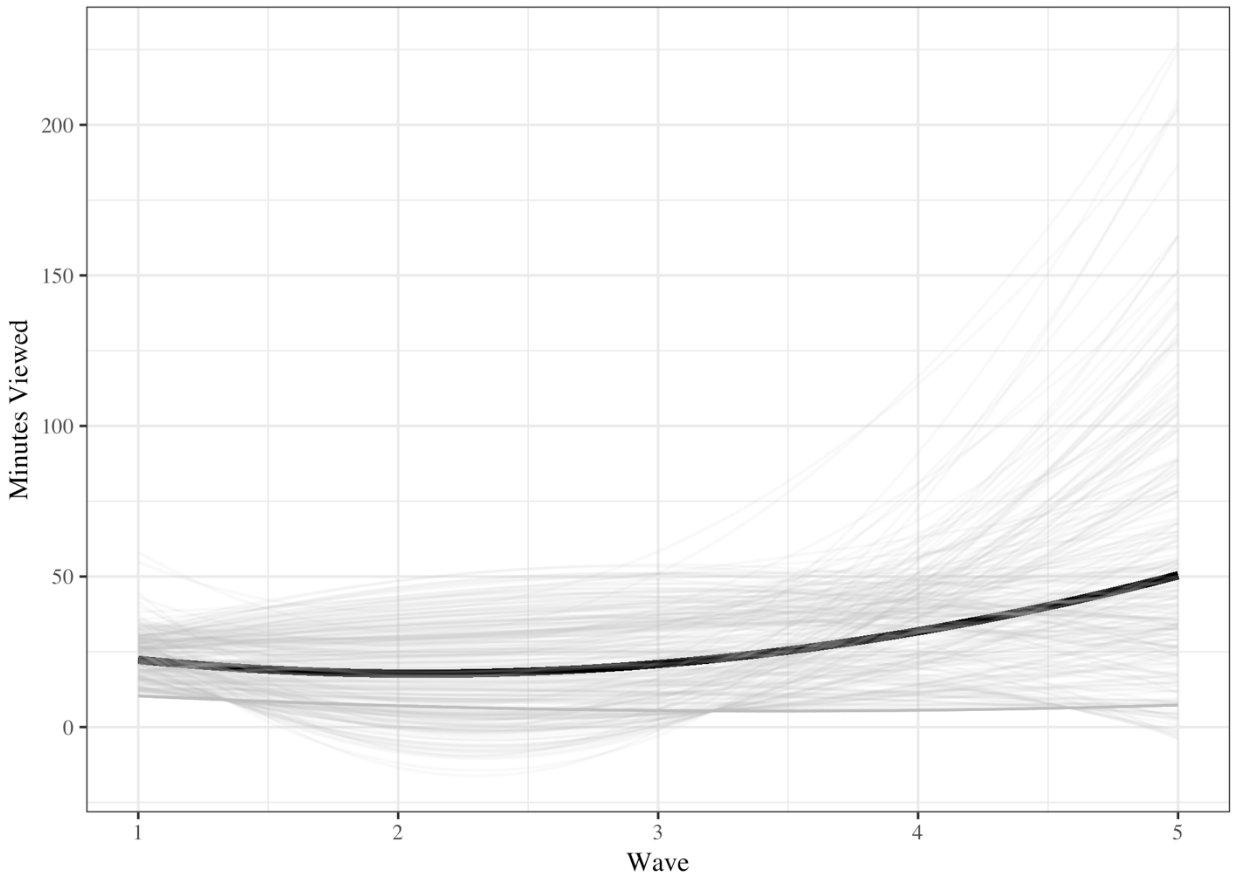
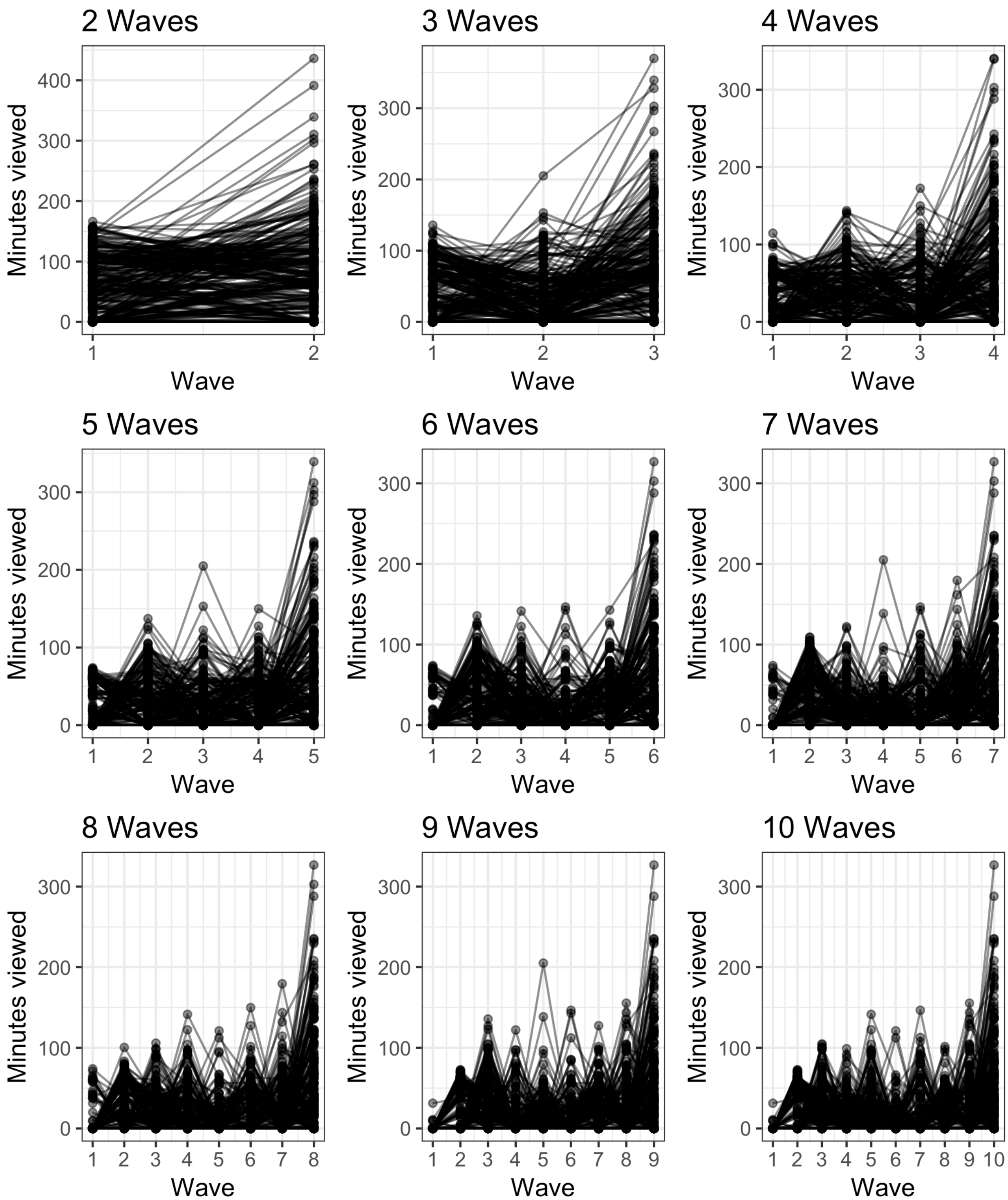


Figure 2. Fitted quadratic growth model (black) and student trajectories (gray).



Note. Fitted values from linear and quadratic models and individual student trajectories. Each wave is associated with approximately five days. Model is estimated with random effects for the intercept and slope (for the linear model) and the intercept, linear, and quadratic terms (for the quadratic model). Residuals are specified with a first-order autoregressive structure.

Appendix A. Minutes viewed for different numbers of waves between Exams 1 and 2.



Appendix B: Link to code to reproduce analysis

A link to the code to reproduce this analysis is available here:
https://github.com/jrosen48/flipped-engagement/blob/master/script.r