Logs and Surveys of Reviewing Behaviors in an Introductory Computer Science Course: Their Motivational Antecedents and Relation to Performance

Hye Rin Lee
University of Delaware
hyerin@udel.edu
Teomara Rutherford
University of Delaware
teomara@udel.edu
Austin C. Bart
University of Delaware
acbart@udel.edu

ABSTRACT
This study draws on theories of Self-Regulated Learning and Situated Expectancy-Value Theory (SEVT) to understand how logs and self-reports of reviewing behaviors relate to each other and to motivation and performance within a university introductory computer science course. Only logs of reviewing from within coding problems were associated with comparable surveys when total course engagement clicks were controlled. Moreover, only course cost, an aspect of SEVT, predicted logs and self-reports of reviewing behaviors. When motivation and both types of reviewing behaviors were considered together as predictors of final course grade, only course cost was a statistically significant predictor, with a negative relationship. Results can inform assumptions of validity between logs and self-reports of reviewing behaviors within a system of Self-Regulated Learning and demonstrate the important role perceptions of cost may play in both self-regulation and performance.

Keywords
Computer science education, self-regulated learning, motivation, log data, survey data

1. INTRODUCTION
Existing research points to the promise of using log data from a Learning Management System (LMS), such as Canvas, to understand students' learning [e.g., 29, 33]. Log data from LMS refers to time-stamped files of student interactions, in this case, records of click events in the course. Scholars in the field posit that log data can provide more objective measures of dispositional constructs (e.g., motivation, self-regulated learning) than do surveys, because with logs, participants are not responding to socially desirable questions, nor do they have to accurately remember what they did [45, 46]. However, studies that compare logs of student actions with surveys that directly query the same constructs present conflicting evidence as to whether logs really can more accurately measure the constructs than self-report surveys [cf 6, 22, 37, 48]. In addition, the landscape of this type of research is still incomplete as to both domain coverage (e.g., Computer Science) and construct coverage—for example, specific aspects of motivation or self-regulation. There is still considerable room to extend comparisons of logs and surveys by context and construct.

In this study, we investigate logs of self-regulated learning [SRL; 51] behaviors from within an introductory university computer science (CS) course, comparing these logs to surveys of the same behaviors. We examine how theoretically related constructs (i.e., motivation) predict both logs and surveys and how each predicts course grades. Our analysis enables us to bracket the usefulness of log behaviors within a system of SRL, where students' motivations inform their studying behavior and this behavior can impact learning and performance [see 44, 52].

1.1 Theoretical Framework, Self-Regulated Learning
We frame our work with models of SRL, such as the Efklides [11] Metacognitive and Affective Model of SRL (MASRL). Within MASRL, SRL is a dynamic process that involves the interplay between broader motivational and regulation dispositions and competencies at the person level (e.g., value for CS) and the specific motivational and regulation dispositions and competencies at the task X person level (e.g., value for the course or a specific topic in CS). Students who are self-regulated successfully engage in several processes to support their learning: generally, theorists of SRL note that within SRL, students plan and set goals, monitor their progress toward those goals, adjust and control behavior in line with monitoring, and evaluate the result of their efforts [11, 30, 49].

1.1.1 SRL and Motivation
Many SRL theorists note the importance of motivational beliefs and goals to the different phases of SRL [31, 39, 50]. For example, in the forethought and planning phase, students are more likely to plan for a task if they are positively motivated, such as feeling competent in the task or seeing the task as interesting, useful, or important. In the performance control phase, students might devalue or value the outcome based on how well they are doing or how vital they perceive the outcome to be. In the self-reflection phase, students might adjust their goal from their recent learning situation. In the current study, we focus on motivation specifically framed within Situated Expectancy–Value Theory [SEVT; 10], where student performance and action are predicated on expectancies and task values. Expectancies for success refer to how successfully a student believes they will accomplish a task; task values refer to how much a student values a task. There are at least four different task value facets: intrinsic value, attainment value, utility value, and cost. Intrinsic value is the enjoyment a student expects to gain from doing a task. Attainment value is how central a task is to a student's personal and collective identities. Utility value is the extent to which a task is useful to a student's short- and

long-term present and future goals. Cost is what a student perceives they must lose or give up by engaging in a task, including the psychological and social costs of pursuing a task. Eccles and Wigfield [10] point out that cost is conceptualized under task values because both benefits and costs of a task contribute to the net value of students' motivation. Each of the task values are on a continuum to very negative to highly positive. Ultimately, both expectancies (i.e., "Can I do it?") and task values (i.e., "Do I want to do it?") play influential roles to SRL behaviors [43].

Previous studies have examined the association between expectancies, values, and SRL strategies [1, 20, 21]. For example, Lawanto [20] focused on investigating the association between values and overall SRL behaviors, as well as values and subcomponents of SRL behaviors, in an introductory mechanical engineering course. They found a statistically significant positive association between values, overall SRL behaviors, and four sub-components of SRL behaviors (i.e., goal setting, task strategies, help seeking, and self-evaluation), but did not find a statistically significant association between values and the other two sub-components of SRL behaviors (i.e., environment structuring and time management). Within a different context of two massive open online courses (MOOCs) in mathematical probability, Lee [21] found that the more MOOC learners had high self-efficacy regarding their course success and believed that the course materials were useful and important, the more they were likely to use SRL strategies. Most of the prior literature that investigated the association between SRL and motivation either did not focus on a specific domain (e.g., examined several different online courses) or focused on non-CS courses (e.g., engineering courses, mathematical probability courses). To our knowledge, no studies have systematically examined the association between SRL and motivation in an in-person CS course context. In-person CS courses allow us to gain important insights on whether findings replicate across contexts and delivery modes. As computer scientists are needed more than ever to support our nation's economy and global competitiveness [17], studying motivation and SRL within a CS context can provide valuable information necessary to enhance student performance and persistence in CS. Additionally, students who take in-person courses typically come from different backgrounds than students who take online courses. For example, older students [25], women [28], and single parents [12] are more likely to enroll in online courses compared to younger students and men because they perceive that the online format allows them to complete their studies alongside other competing responsibilities, such as taking care of their children [3, 24]. Studies such as ours expand work that is typically conducted in MOOC and other online settings to different populations. Further, our study examines a negatively-valenced value, cost, which is an important element of motivation [10], but has seldom been included in research on links between motivation and SRL strategies. Empirical evidence suggests that positive versus negative emotions predict different SRL behaviors [2]. Therefore, positively- versus negatively-valenced motivation may relate to different SRL behaviors as well.

1.1.2 SRL and Performance

SRL has been long-been associated with achievement [49]. Students as young as elementary school who are more self-regulated perform better on academic tasks [8, 18]. SRL behaviors have been linked with course grades in undergraduate CS courses [5, 23, 36]. However, there have been mixed findings on this association depending on the type of SRL behavior and context. For instance, undergraduate students in an introductory CS course who had higher performance used more SRL strategies, such as metacognitive strategies (i.e., planning, monitoring, regulating strategies, and critical thinking) and resource management (i.e., time and study environment strategies and effort regulation strategies) than those with lower performance [5]. In contrast, Sands and colleagues [36] found the SRL behavior, help seeking, as the only statistically significant predictor of performance in a high school CS MOOC. Prior research studies have also found that self-efficacy (motivation) can serve as a mediator of the relationship between SRL metacognitive strategies and performance for CS undergraduate students [23].

1.1.3 SRL and Log Data

Within models of SRL [11, 30], students leverage their motivational and regulation dispositions and competencies through behavioral action to succeed in learning and academic performance. In this way, the behaviors that students exhibit are representative of both their SRL and motivation. Existing literature has investigated the accuracy of convergence between observed behaviors of SRL (as measured by logs) and self-reported behaviors of SRL [as measured by surveys; 6, 35, 48]. For example, Salehian Kia and colleagues [35] found substantial agreement in the SRL behavior responses between log and self-reported data, particularly in relation to the task definition and planning phases. In contrast, Zhou & Winne [48] found that the self-reported goal orientations of undergraduate students from various departments did not correlate with goals measured by digital traces. Various studies also conclude that SRL log data and self-reported data predict differently to achievement, where logs are generally more predictive of achievement than self-reports [6, 32]. Taken together, these studies show mixed results on the association between log and self-report survey data for SRL from either non-domain-specific or non-CS contexts. Moreover, few studies have systematically validated specific SRL behaviors from logs with self-reports of those specific behaviors, as most studies have investigated the validation between conceptually broader SRL self-reports to specific logs (e.g., planning self-report measures to spacing vs. cramming log measures compared to going back to a prior lesson self-report and log measure). The current study contributes to methodologically validating the association between specific SRL behavior log data and SRL self-reported survey data.

1.2 Context

The context for this study is an introductory CS course at a large, public university in the Mid-Atlantic region of the United States. One instructor taught one of the sections and one instructor—the third author of this study—taught two of the sections. Each instructor followed the same curriculum. Each instructor also embedded four surveys into their course. Surveys were administered at the start of the course, after week six (the first midterm), after week 10 (the second midterm), and at week 14 (the very end of the course). Each survey took approximately 10 minutes. Students received 1% overall course credit for completing each survey, regardless of their answers. Although these surveys were part of the students' coursework, IRB approval was secured for the research analysis of the data.

Course content focused on teaching introductory CS concepts and practices through the use of Python. The course was structured around both programming fundamentals (functions, loops, conditionals) and software engineering basics (testing, functional decomposition, ethics). Outside of class, students were responsible for completing readings, quizzes, and programming problems that taught the core material. These materials were delivered via the course LMS and an instructor-created interface [4]. Students also met in person two to three times a week for lecture and once a week for lab, where they received high-level overviews of the material and participated in activities that reinforced what they were learning. Along the way, they had three exams (two midterms and a final) and several projects of escalating size and complexity.

1.3 Current Study

Within the current study we expand upon prior research defining log behaviors within a system of SRL to specifically investigate reviewing behaviors: times students clicked from their current lesson to a prior lesson (in general, while they were taking a quiz, while they were completing a coding activity). We theorize that this behavior represents an aspect of SRL control triggered by a perception through monitoring during the current activity that the students' knowledge was not adequate. We compare these logs to self-reports of the same behaviors and investigate motivational antecedents to both logs and self-reports and how both predict final course grade.

Our specific research questions are:

1) To what extent do logs of reviewing behaviors align with student reports of reviewing behaviors?
2) To what extent does motivation for the course predict reviewing behaviors (both logs and self-reports)?
3) To what extent do reviewing behaviors (both logs and self-reports) predict final grades in the course?

2. METHOD

2.1 Participants

There were 154 students across the three courses. One hundred and four of these students (68%) provided consent to the surveys. This was largely consistent across the three sections (range 65% to 69%). The number of participants varied across surveys: survey one: 104, two: 95, three: 76, four: 79. Students reported on their gender and on their race. Gender questions asked students to indicate whether they were men (82), women (19), trans (1), not listed (1), or preferred not to answer (1). An open-ended follow-up question indicated the two students who answered not listed or prefer not to answer both identified as non-binary (NB). Students who were not men were combined into one group of 22 students for analysis. Students reported on their race/ethnicity by selecting one or more from American Indian or Alaskan Native (1); Asian (30); Black or African American (11); Hispanic, Latino, or Spanish origin (10); Middle Eastern or North African (2); White (50); or some other race, ethnicity, or origin (3). All students chose between one and three racial/ethnic categories. Because we were interested in understanding the motivation of those students from races or ethnicities underrepresented in CS, we created a binary variable labeling those who reported as identifying as any group other than White or Asian as underrepresented in CS [7, 27]. The percentages of women and URM students among the participants was similar to that in the introductory CS courses at this university and did not differ across course sections. The current study is limited to 94 CS students who completed the second survey and had a final grade for the course.

2.2 Measures

2.2.1 Logs of Reviewing Behaviors.

Three different measures of reviewing behaviors were created: general, quizzing, and coding. As the course is chronologically structured in the LMS and instructor-created interface, going back to prior course content was operationalized as a reviewing behavior. Logs of general reviewing behaviors were operationalized as the number of times the log recorded an event with a lesson number less than the immediately prior event's lesson number. The same method was applied to quizzes and coding reviewing behaviors, except that these counts referred to the types a student returned to a prior lesson while they were in the process of completing a quizzing or coding problem. Counts of general reviewing behaviors include reviewing from quizzes and coding problems, but also reviewing when navigating other aspects of the course, such as reading/watching videos.

2.2.2 Logs of Total Event Counts

Logs of total event counts were operationalized as the raw number of events in the logs associated with a given participant.

2.2.3 Student Reports of Reviewing Behaviors

In waves two through four students were specifically asked about their reviewing behaviors (e.g., "Please consider how often you have done each of the items below since last month… Look back at prior lessons while completing a quiz"). Students were presented with statements about reviewing in general, during quizzes, and during coding problems; participants then indicated how often they did that action on a scale from 0 to 5 (I haven't done this, once or twice, once a module, a few times each module, every lesson, or multiple times a lesson).

2.2.4 Motivation

Students answered questions about their expectancy and value for the course on a six-point Likert-type scale. Questions were structured as self-efficacy (e.g., "How good would you be at learning something new in this course?") to represent expectancy, as the two constructs are theoretically similar and often empirically indistinguishable [34, 35]. Questions were asked at all four surveys about both course expectations (two questions, alpha .75; e.g., "How well do you expect to do in this course?") and attainment value for the course (two questions, alpha .82; e.g., "How important is it to you to do well in this course?"). Value questions around course interest (three questions, alpha .83; e.g., "To you, how much fun is this course?"), utility (three questions, alpha .73; e.g., "How useful is what you are learning in this course for your future career?"), and cost (six questions, alpha .90; e.g., "How much time does this course demand?") were only asked during waves two through four, as students needed some experience with the course to answer these questions. Expectancy and positive values questions were based on constructs from SEVT [10] and borrowed/adapted language from Gaspard [14] to focus on the specific context of this course. Cost questions were those from Flake [13] but rephrased from statements to questions. Within the wave, motivation questions correlated as expected based on prior research [9].

2.3 Analysis

To answer the first research question, rank-order correlations were examined between logs of reviewing behaviors, total event counts, and students' reports of reviewing behaviors. Because we used a total count of reviewing behaviors it could be that student overall engagement with the course drove results. Therefore, we conducted additional regression analyses to control for total engagement operationalized as event (click) counts. In addition, regression models included gender, professor, and year in college as controls.

To answer the second research question, rank-order correlations were examined between motivation and reviewing behaviors. Due to the high correlations between the motivation constructs, a scale variable was created for positively-valenced value by averaging interest, attainment, and utility value (alpha .82). This variable was then included with expectancy and cost in one series of models predicting the three types of logs of reviewing behaviors and in another series of models predicting the three types of student reports of reviewing behaviors. Again, all models included total event count, gender, professor, and year in college as controls.

To answer the third research question, for parsimony, we created two scales combining logs and student reports of reviewing behaviors together (alphas .82, .79 respectively) because of the strong correlations between the three types of reviewing behaviors within modality. These variables were entered together as predictors of final course grade, along with motivation, total events, and demographics predicting final grade in the course.

3. RESULTS

3.1 Research Question 1

Rank-order correlations showed statistically significant positive associations between logs and surveys of general reviewing behaviors (r = .29, p = .005) as well as logs and surveys of coding reviewing behaviors (r = .36, p < .001; see Table 1). There was no statistically significant association between logs and surveys regarding reviewing behavior from quizzes. In addition, total event count—as a measure of general student engagement—was positively associated with all logs of reviewing behavior (general: r = .70, p < .001; quizzing: r = .30, p = .004; coding: r = .81, p < .001) and all student reports of reviewing behavior (general: r = .42, p < .001; quizzing: r = .35, p < .001; coding: r = .29, p = .004). When considering regressions of log behaviors on student reports of reviewing behavior controlling for total event counts, only coding reviewing behaviors were associated with comparable surveys (β = .17, p = .01; see Table 2). In all models, total event count remained a statistically significant predictor of logs of reviewing behaviors (general: β = .67, p < .001; quizzing: β = .28, p = .008; coding: β = .75, p < .001). Results remained largely unchanged when demographic controls were included.

Table 1. Correlation between all variables

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Expectancy for Success	--												
2. Intrinsic Value	0.39^{c}	--											
3. Utility Value	0.34^{c}	0.56^{c}	--										
4. Attainment Value	0.43^{c}	0.65^{c}	0.55^{c}	--									
5. Cost	-0.37^{c}	-0.09	-0.13	-0.02	--								
6. Went to prior lesson log	-0.02	0.08	0.13	0.18	0.26^{b}	--							
7. Quizzing went to prior lesson log	-0.15	-0.13	0.05	0.03	0.33^{b}	0.56^{c}	--						
8. Coding went to prior lesson log	-0.05	0.06	0.07	0.22^{b}	0.31^{b}	0.84^{c}	0.44^{c}	--					
9. Event count log	0.03	0.19	0.11	0.30^{b}	0.34^{c}	0.70^{c}	0.30^{b}	0.81^{c}	--				
10. Went to prior lesson survey	-0.08	0.15	0.08	0.18	0.40^{c}	0.29^{b}	0.10	0.34^{c}	0.42^{c}	--			
11. Quizzing went to prior lesson survey	-0.17	0.01	0.01	0.10	0.40^{c}	0.29^{b}	0.16	0.38^{c}	0.35^{c}	0.54^{c}	--		
12. Coding went to prior lesson survey	-0.25^{b}	-0.04	0.01	0.08	0.43^{c}	0.24^{b}	0.14	0.36^{c}	0.29^{b}	0.60^{c}	0.61^{c}	--	
13. Grade	0.41^{c}	0.22^{b}	0.17	0.33^{b}	-0.46^{c}	0.10	-0.24^{b}	0.18	0.18	-0.18	-0.16	-0.18	--

$^{a}p < 0.05$ $^{b}p < 0.01$ $^{c}p < 0.001$.

Table 2. Regression of logs of students' reviewing behaviors on surveys

	Went to Prior Lesson Logs		Quizzing Went to Prior Lesson Logs		Coding Went to Prior Lesson Logs	
Reviewing behaviors survey	0.03 [-0.13, 0.19]	0.03 [-0.14, 0.19]	0.08 [-0.12, 0.29]	0.10 [-0.11, 0.31]	0.17* [0.04, 0.29]	0.17* [0.04, 0.29]
Event count	0.67*** [0.51, 0.83]	0.67*** [0.50, 0.83]	0.28** [0.08, 0.49]	0.30** [0.09, 0.51]	0.75*** [0.63, 0.88]	0.76*** [0.63, 0.88]
Women		0.06 [-0.10, 0.21]		-0.10 [-0.30, 0.10]		0.03 [-0.10, 0.15]
Professor		-0.02 [-0.18, 0.14]		-0.19 [-0.39, 0.01]		-0.15* [-0.27, -0.02]
Year 2		0.10 [-0.44, 0.63]		0.17 [-0.51, 0.85]		-0.01 [-0.43, 0.42]
Year 3 & 4		0.02 [-0.42, 0.46]		-0.26 [-0.82, 0.30]		-0.08 [-0.43, 0.27]
R^2	0.46	0.47	0.10	0.15	0.65	0.67

[a] The reported numbers are as follows from top to bottom: standardized beta and confidence interval. The reference group for women was men. Women included women, trans, and non-binary. Professor was dichotomized as 0 = *the professor who taught two sections of the course* and 1 = *the professor who taught one section of the course*. Event count log referred to total engagement in the course. $^{*}p < 0.05$ $^{**}p < 0.01$ $^{***}p < 0.001$.

3.2 Research Question 2

Rank-order correlations showed there were statistically significant positive associations between self-reported course cost and all logs of reviewing behaviors: general (r = .23, p = .02), quizzing (r = .34, p = .001), and coding (r = .26, p = .01; see Table 1). There were also statistically significant positive associations between self-reported course cost and student reports of reviewing behaviors: general (r = .41, p < .001), quizzing went to prior lessons (r = .35, p = .001), and coding (r = .46, p < .001). None of the other motivational constructs showed statistically significant association with reviewing behaviors from logs or surveys. Regressions showed only a statistically significant positive association between self-reported course cost and logs of reviewing from quizzes (β = .28, p = .02; see Table 3). Also, regressions showed a significant positive association between cost and all student reports of reviewing behavior (general: β = .35, p = .003; quizzing: β = .27, p = .02; coding: β = .39, p < .001; see Table 4). However, there were no statistically significant associations between expectancy for success, positively-valenced values, and student reports of reviewing behaviors.

Table 3. Regression of motivation predicting logs of reviewing behaviors

	Went to Prior Lesson Logs		Quizzing Went to Prior Lesson Logs		Coding Went to Prior Lesson Logs	
Expectancy	0.08 [-0.12, 0.28]	0.09 [-0.12, 0.30]	0.04 [-0.21, 0.29]	0.07 [-0.18, 0.33]	-0.07 [-0.23, 0.10]	-0.04 [-0.21, 0.13]
Pos. Value	-0.07 [-0.26, 0.11]	-0.08 [-0.27, 0.11]	-0.08 [-0.31, 0.15]	-0.09 [-0.33, 0.15]	-0.04 [-0.20, 0.11]	-0.06 [-0.22, 0.09]
Cost	0.02 [-0.16, 0.20]	0.03 [-0.16, 0.22]	0.27* [0.05, 0.50]	0.28* [0.05, 0.51]	-0.05 [-0.20, 0.10]	-0.03 [-0.18, 0.12]
Event count	0.68*** [0.52, 0.85]	0.68*** [0.51, 0.85]	0.22* [0.01, 0.44]	0.24* [0.03, 0.46]	0.81*** [0.67, 1.00]	0.82*** [0.68, 1.00]
Woman		0.06 [-0.10, 0.22]		-0.08 [-0.27, 0.12]		0.04 [-0.09, 0.17]
Professor		-0.03 [-0.20, 0.13]		-0.19 [-0.39, 0.01]		-0.13* [-0.27, < 0.01]
Year 2		0.06 [-0.49, 0.61]		0.05 [-0.63, 0.73]		-0.04 [-0.49, 0.40]
Year 3 & 4		0.02 [-0.42, 0.47]		-0.22 [-0.77, 0.33]		-0.12 [-0.49, 0.24]
R^2	0.47	0.47	0.16	0.20	0.63	0.65

[a] Pos. value refers to intrinsic + attainment + utility value. The reported numbers are as follows from top to bottom: standardized beta and confidence interval. The reference group for women was men. Women included women, trans, and non-binary. Professor was dichotomized as 0 = *the professor who taught two sections of the course* and 1 = *the professor who taught one section of the course*. Event count log referred to total engagement in the course.
$^{*}p < 0.05$ $^{**}p < 0.01$ $^{***}p < 0.001$.

Table 4. Regression of motivation predicting survey reports of reviewing behaviors

	Went to Prior Lesson Surveys		Quizzing Went to Prior Lesson Surveys		Coding Went to Prior Lesson Surveys	
Expectancy	-0.04 [-0.28, 0.20]	-0.03 [-0.28, 0.22]	-0.55 [-0.30, 0.20]	-0.06 [-0.32, 0.20]	-0.14 [-0.38, 0.10]	-0.15 [-0.40, 0.10]
Pos. Value	0.13 [-0.09, 0.36]	0.12 [-0.12, 0.35]	0.19 [-0.14, 0.32]	0.10 [-0.14, 0.34]	0.13 [-0.09, 0.35]	0.12 [-0.11, 0.35]
Cost	0.33** [0.11, 0.55]	0.35** [0.12, 0.57]	0.28* [0.05, 0.50]	0.27* [0.04, 0.51]	0.39*** [0.17, 0.61]	0.39*** [0.17, 0.62]
Event count	0.20 [-0.01, 0.40]	0.18 [-0.02, 0.39]	0.18 [-0.03, 0.39]	0.17 [-0.05, 0.39]	0.06 [-0.14, 0.27]	0.06 [-0.14, 0.27]
Woman		0.07 [-0.12, 0.27]		0.03 [-0.17, 0.23]		0.07 [-0.12, 0.26]
Professor		0.00 [-0.20, 0.20]		0.06 [-0.15, 0.26]		0.05 [-0.15, 0.24]
Year 2		-0.13 [-0.79, 0.54]		0.13 [-0.56, 0.82]		-0.13 [-0.79, 0.54]
Year 3 & 4		0.13 [-0.41, 0.66]		0.29 [-0.27, 0.85]		-0.18 [-0.72, 0.35]
R^2	0.22	0.23	0.16	0.18	0.23	0.24

[a] Pos. value refers to intrinsic + attainment + utility value. The reported numbers are as follows from top to bottom: standardized beta and confidence interval. The reference group for women was men. Women included women, trans, and non-binary. Professor was dichotomized as 0 = *the professor who taught two sections of the course* and 1 = *the professor who taught one section of the course*. Event count log referred to total engagement in the course.
$^{*}p < 0.05$ $^{**}p < 0.01$ $^{***}p < 0.001$.

3.3 Research Question 3

Regressions showed that there was a significant negative association between self-reported course cost and grade in the course after controlling for demographics (β = -.48, p < .001; see Table 5). There were no other statistically significant associations between motivation and grade in the course. Moreover, reviewing behaviors did not statistically significantly predict grades in the course.

Table 5: Reviewing behavior predicting grade

	Standardized Model	
	Beta	
Reviewing behavior	0.02 [-0.20, 0.25]	0.02 [-0.21, 0.25]
Expectancy	0.19 [-0.03, 0.40]	0.20 [-0.02, 0.42]
Pos. Value	0.07 [-0.13, 0.27]	0.05 [-0.16, 0.26]
Cost	-0.50*** [-0.69, -0.30]	-0.48*** [-0.68, -0.28]
Event count	0.36** [0.13, 0.60]	0.36** [0.11, 0.60]
Woman		-0.01 [-0.18, 0.16]
Professor		-0.02 [-0.20, 0.15]
Year 2		-0.17 [-0.76, 0.42]
Year 3 & 4		0.19 [-0.28, 0.66]
R^2	0.40	0.41

[a] Pos. value refers to intrinsic + attainment + utility value. The reported numbers are as follows from top to bottom: standardized beta and confidence interval. The reference group for women was men. Women included women, trans, and non-binary. Professor was dichotomized as 0 = *the professor who taught two sections of the course* and 1 = *the professor who taught one section of the course*. Event count log referred to total engagement in the course.
$^{*}p < 0.05$ $^{**}p < 0.01$ $^{***}p < 0.001$.

4. DISCUSSION

In the current study, we triangulated logs of SRL reviewing behaviors to student reports of the same behaviors in three sections of an introductory CS course. We also examined the extent to which motivation (expectancies and values) predicted logs and student reports of SRL reviewing behaviors and how each predicted the final course grade. Prior research has looked at a limited number of SRL behaviors [36] and work in CS has focused on introductory CS students within a large online course rather than an in-person course [5]. Our research focused specifically on reviewing behaviors, an aspect of control within a system of SRL [26], by investigating when students returned to prior content in general, when taking quizzes, and when completing coding problems during online engagement within an otherwise in-person CS course.

Prior research is mixed with respect to matches between logs and student reports [6, 22, 37, 48]. Looking at the bivariate relations between logs and student reports of SRL reviewing behaviors, there were statistically significant associations for reviewing behaviors overall and those specific to coding. There was no statistically significant association between logs of the quizzing reviewing behaviors and student reports of quizzing reviewing behaviors. In regression analyses that controlled for total events/clicks, only coding reviewing behaviors were associated with comparable surveys. Our work suggests that students are better at providing accurate information on some tasks compared to others. Regarding the specific tasks, students may recognize the relevance of prior content when completing problem-based activities, such as the coding problems in the introductory CS course. This finding aligns with previous research on problem solving in that students who can successfully solve problems reflect back on prior content [40]. The lack of alignment between our other log and self-report data also suggest that these measures might best be used as complementary information about SRL rather than interchangeable measures. One implication from this finding is that instructors can ask students to keep track of their time when returning to prior content in general and when taking quizzes to help students get a more accurate picture of SRL skills and learning. Another implication from this finding is that instructors cannot merely rely on log data as representative of students' SRL. Future research should investigate how scholars and practitioners can leverage survey and log data in a way that complements each other rather than using these measures in disparate ways.

In regression analyses predicting SRL reviewing behaviors from motivation, self-reported course cost emerged as the only statistically significant motivation predictor of reviewing behaviors or reports. In our study, cost measures included questions regarding perceived course time and work demands. That these questions were related both to student reviewing behavior and to student total event clicks suggests students who felt they invested a lot did engage more with the course. This may also suggest that students who feel more frustrated and stressed are more likely to review materials from the course in order to alleviate the emotional cost. Perhaps, some level of perceived cost is optimal for learning [19, 38]. Future studies should investigate whether there is a curvilinear association between cost and SRL reviewing behaviors (e.g., do CS students who perceive too much cost decide to disengage later on in the course?). Contrary to prior research [1, 20, 21], we did not find statistically significant associations between the positively-valenced values (interest, utility, importance) and SRL reviewing behaviors. One potential explanation is that only certain SRL behaviors associate with the positively-valenced values, whereas others do not—prior studies have not explicitly examined reviewing behaviors [2].

Finally, we found that there was a statistically significant association between self-reported cost and grade in the course. Aligned with SEVT [10], cost was a statistically significant predictor of achievement. However, contrary to previous studies, there were no statistically significant associations between expectancies, positively-valenced values, and course grades [1, 20, 21]. Results also showed that SRL reviewing behaviors did not statistically significantly predict grades in the course. This suggests that students' own perceptions mattered more than actual behavior. Future research should examine whether reviewing behaviors mediate the relationship between motivation and course performance. One aspect of investigating these associations that might be tricky is that cost was positively associated with logs and reports of reviewing behaviors even though cost was negatively related to grades. Perhaps, there are two facets of perceived cost. In other words, cost promotes SRL behaviors, but also hinders performance. Aligned with Eccles and Wigfield's [10] SEVT, cost is on a spectrum from negative to positive. Our results provide an example of this negative and positive cost phenomenon. Instructors should consider optimizing feelings of cost to the point where students value the course enough to invest effort, but also not to the point where they face negative consequences from engaging in the course. This balance may be related to that seen in stress research [e.g., 34, 47].

One limitation in our approach is that we examined aggregate measures of reviewing behaviors instead of examining temporally distinct reviewing events (e.g., after a midterm exam). Investigating the temporality of reviewing can provide insight into how students change their SRL behaviors after key events in the course. Future studies can leverage the multiple time points of log and survey data throughout the academic term. In addition, reviewing, as an SRL control process, operates within a system of SRL [26]. Although we examined motivational aspects of this system, we did not examine other SRL skills and processes, such as metacognitive monitoring [30, 49]. Future studies could examine these aspects of SRL to note how they align to self-reports within and between constructs. Finally, in the current study, logs of students' reviewing behavior were measured in the context of the course LMS and an instructor-created interface, which might have not captured all reviewing behaviors related to this course. For instance, students might review through Google or other online sources. However, we believe that this construct validity concern is low, because preliminary results from interviewing students about their learning experiences found that most students utilized the abundant course resources given.

These results provide insight into students' SRL, motivation, and performance within in-person CS courses. Our mixed alignment between logs and self-reports of reviewing behaviors calls into question the interchangeability of these measures. Their relations with motivation (i.e., perceived course cost) provide insight into how motivation and control processes operate within a system of SRL. Understanding what SRL logs represent and predict can inform future work to distill these student behaviors into actionable insights for instructors. Our work also demonstrates the importance of perceived course cost; an importance that should be considered for future motivational interventions that to this point have largely focused on positive aspects of task values [15, 16].

5. ACKNOWLEDGMENTS

We thank the participating instructors and students. We thank our undergraduate research assistants, Yiqun Zhang, Meile Li, Juliana Baer, and Sotheara Veng for formatting the references and paper.

6. REFERENCES

[1] Anthony R. Artino and Jason M. Stephens. 2006. Learning Online: Motivated to Self-Regulate?. *Academic Exchange Quarterly* 10, 4 (2006), 176–182.

[2] Anthony R. Artino and Kenneth D. Jones. 2012. Exploring the complex relations between achievement emotions and self-regulated learning behaviors in online learning. *The Internet and Higher Education* 15, 3 (2012), 170–175. DOI: 10.1016/j.iheduc.2012.01.006

[3] Matthew Bailey, Dirk Ifenthaler, Maree Gosper, Mandy Kretzschmar, and Cheryl Ware. 2015. The Changing Importance of Factors Influencing Students' Choice of Study Mode. *Tech Know Learn* 20, 2 (July 2015), 169–184. DOI: 10.1007/s10758-015-9253-9

[4] Austin C. Bart, Teomara Rutherford, and James Skripchuk. 2020. Evaluating an Instrumented Python CS1 Course." In *CSEDM@ EDM*.

[5] Susan Bergin, Ronan Reilly, and Desmond Traynor. 2005. Examining the role of self-regulated learning on introductory programming performance. In *Proceedings of the 2005 International Workshop on Computing Education Research - ICER '05*, ACM Press, Seattle, WA, USA, 81–86. DOI: 10.1145/1089786.1089794

[6] Moon-Heum Cho and Jin Soung Yoo. 2017. Exploring online students' self-regulated learning with self-reported surveys and log files: a data mining approach. *Interactive Learning Environments* 25, 8 (November 2017), 970–982. DOI: 10.1080/10494820.2016.1232278

[7] Leslie Cintron, Yunjeong Chang, James Cohoon, Luther Tychonievich, Brittany Halsey, Devon Yi, and Genevieve Schmitt. 2019. Exploring Underrepresented Student Motivation and Perceptions of Collaborative Learning-Enhanced CS Undergraduate Introductory Courses. In *2019 IEEE Frontiers in Education Conference (FIE)*, 1–6. DOI: 10.1109/FIE43999.2019.9028463

[8] Charlotte Dignath and Gerhard Büttner. 2008. Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition Learning* 3, 3 (December 2008), 231–264. DOI: 10.1007/s11409-008-9029-x

[9] Jacquelynne S. Eccles and Allan Wigfield. 2002. Motivational Beliefs, Values, and Goals. *Annu. Rev. Psychol.* 53, 1 (2002), 109–132. DOI: 10.1146/annurev.psych.53.100901.135153

[10] Jacquelynne S. Eccles and Allan Wigfield. 2020. From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology* 61, (April 2020), 101859. DOI: 10.1016/j.cedpsych.2020.101859

[11] Anastasia Efklides. 2011. Interactions of Metacognition with Motivation and Affect in Self-Regulated Learning: The MASRL Model. *Educ. Psychol.* 46, 1 (January 2011), 6–25. DOI: 10.1080/00461520.2011.538645

[12] Maya Escueta, Vincent Quan, Andre Joshua Nickow, and Philip Oreopoulos. 2017. *Education Technology: An Evidence-Based Review*. National Bureau of Economic Research, Cambridge, MA. DOI: 10.3386/w23744

[13] Flake, Jessica Kay, et al. "Measuring Cost: The Forgotten Component of Expectancy-Value Theory." *Contemporary Educational Psychology*, vol. 41, Apr. 2015, pp. 232–44. *DOI:* 10.1016/j.cedpsych.2015.03.002.

[14] Hanna Gaspard, Isabelle Häfner, Cora Parrisius, Ulrich Trautwein, and Benjamin Nagengast. 2017. Assessing task values in five subjects during secondary school: Measurement structure and mean level differences across grade level, gender, and academic subject. *Contemporary Educational Psychology* 48, (January 2017), 67–84. DOI: 10.1016/j.cedpsych.2016.09.003

[15] Judith M. Harackiewicz, Elizabeth A. Canning, Yoi Tibbetts, Stacy J. Priniski, and Janet S. Hyde. 2016. Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *Journal of Personality and Social Psychology* 111, 5 (November 2016), 745–765. DOI: 10.1037/pspp0000075

[16] Judith M. Harackiewicz, Christopher S. Rozek, Chris S. Hulleman, and Janet S. Hyde. 2012. Helping Parents to Motivate Adolescents in Mathematics and Science: An Experimental Test of a Utility-Value Intervention. *Psychol Sci* 23, 8 (August 2012), 899–906. DOI: 10.1177/0956797611435530

[17] Avi Hofstein, Ingo Eilks, and Rodger Bybee. 2011. Societal issues and their importance for contemporary science education—a pedagogical justification and the state-of-the-art in Israel, Germany, and the USA. *Int J of Sci and Math Educ* 9, 6 (December 2011), 1459–1483. DOI: 10.1007/s10763-010-9273-9

[18] Robin B. Howse, Garrett Lange, Dale C. Farran, and Carolyn D. Boyles. 2003. Motivation and Self-Regulation as Predictors of Achievement in Economically Disadvantaged Young Children. *The Journal of Experimental Education* 71, 2 (January 2003), 151–174. DOI: 10.1080/00220970309602061

[19] Jared Keeley, Ryan Zayac, and Christopher Correia. 2008. Curvilinear relationships between statistics anxiety and performance among undergraduate students: Evidence for optimal anxiety, SERJ 7, 1 (May 2008), 4–15. DOI: 10.52041/serj.v7i1.477

[20] Oenardi Lawanto, Harry Santoso, Kevin Lawanto, and Wade Goodridge. 2014. Self-Regulated Learning Skills and Online Activities Between Higher and Lower Performers on a Web Intensive Undergraduate Engineering Course. *JEO* 11, 3 (2014). DOI: 10.9743/JEO.2014.3.2

[21] Daeyeoul Lee, Sunnie Lee Watson, and William R Watson. 2020. The Relationships Between Self-Efficacy, Task Value, and Self-Regulated Learning Strategies in Massive Open Online Courses. *IRRODL* 21, 1 (January 2020), 23–39. DOI: 10.19173/irrodl.v20i5.4389

[22] Qiujie Li, Rachel Baker, and Mark Warschauer. 2020. Using clickstream data to measure, understand, and support self-regulated learning in online courses. *The Internet and Higher Education* 45, (2020), 100727. DOI: 10.1016/j.iheduc.2020.100727

[23] Alex Lishinski, Aman Yadav, Jon Good, and Richard Enbody. 2016. Learning to Program: Gender Differences and Interactive Effects of Students' Motivation, Goals, and Self-Efficacy on Performance. In *Proceedings of the 2016 ACM Conference on International Computing Education Research*, ACM, Melbourne VIC Australia, 211–220. DOI: 10.1145/2960310.2960329

[24] Peter McPartlan, Teomara Rutherford, Fernando Rodriguez, Justin F. Shaffer, and Amanda Holton. 2021. Modality motivation: Selection effects and motivational differences in students who choose to take courses online. *The Internet and Higher Education* 49, (April 2021), 100793. DOI: 10.1016/j.iheduc.2021.100793

[25] Michael G. Moore and Greg Kearsley. 2012. *Distance education: A systems view of online learning* (3rd ed ed.). Wadsworth Cengage Learning, Belmont, CA.

[26] Ernesto Panadero. 2017. A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Front. Psychol.* 8, (April 2017), 422. DOI: 10.3389/fpsyg.2017.00422

[27] Crystal L. Park, Michelle K. Williams, Paul R. Hernandez, V. Bede Agocha, Lauren M. Carney, Andrea E. DePetris, and Sharon Y. Lee. 2019. Self-regulation and STEM persistence in minority and non-minority students across the first year of college. *Soc Psychol Educ* 22, 1 (2019), 91–112. DOI: 10.1007/s11218-018-9465-7

[28] L. Price. 2006. Gender differences and similarities in online courses: challenging stereotypical views of women: Gender differences and similarities in online courses. *Journal of Computer Assisted Learning* 22, 5 (September 2006), 349–359. DOI: 10.1111/j.1365-2729.2006.00181.x

[29] A. Phippen, L. Sheppard, and S. Furnell. 2004. A practical evaluation of Web analytics. *Internet Research* 14, 4 (2004), 284–293. DOI: 10.1108/10662240410555306

[30] Paul R. Pintrich. 2000. The Role of Goal Orientation in Self-Regulated Learning. In the *Handbook of Self-Regulation*. Elsevier, 451–502. DOI: 10.1016/B978-012109890-2/50043-3

[31] Paul R. Pintrich and Akene Zusho. 2002. The Development of Academic Self-Regulation. In *Development of Achievement Motivation*. Elsevier, 249–284. DOI: 10.1016/B978-012750053-9/50012-7

[32] Fernando Rodriguez, Renzhe Yu, Jihyun Park, Mariela Janet Rivas, Mark Warschauer, and Brian K. Sato. 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*, ACM, Tempe AZ USA, 456–460. DOI: 10.1145/3303772.3303841

[33] Ido Roll and Philip H. Winne. 2015. Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Learning Analytics* 2, 1 (July 2015), 7–12. DOI: 10.18608/jla.2015.21.2

[34] Joy R. Rudland, Clinton Golding, and Tim J. Wilkinson. 2020. The stress paradox: how stress can be good for learning. *Med Educ* 54, 1 (January 2020), 40–45. DOI: 10.1111/medu.13830

[35] Fatemeh Salehian Kia, Marek Hatala, Ryan S. Baker, and Stephanie D. Teasley. 2021. Measuring Students' Self-Regulatory Phases in LMS with Behavior and Real-Time Self Report. In *LAK21: 11th International Learning Analytics and Knowledge Conference*, ACM, Irvine CA USA, 259–268. DOI: 10.1145/3448139.3448164

[36] Phil Sands and Aman Yadav. 2020. Self-Regulation for High School Learners in a MOOC Computer Science Course. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, ACM, Portland OR USA, 845–851. DOI: 10.1145/3328778.3366818

[37] Michael Scharkow. 2016. The Accuracy of Self-Reported Internet Use—A Validation Study Using Client Log Data. *Communication Methods and Measures* 10, 1 (January 2016), 13–27. DOI: 10.1080/19312458.2015.1118446

[38] Frieder L. Schillinger, Jochen A. Mosbacher, Clemens Brunner, Stephan E. Vogel, and Roland H. Grabner. 2021. Revisiting the Role of Worries in Explaining the Link Between Test Anxiety and Test Performance. *Educ Psychol Rev* 33, 4 (2021), 1887–1906. DOI: 10.1007/s10648-021-09601-0

[39] Dale H. Schunk and Peggy A. Ertmer. 2000. Self-Regulation and Academic Learning. In *Handbook of Self-Regulation*. Elsevier, 631–649. DOI: 10.1016/B978-012109890-2/50048-2

[40] Kaye Stacey. 2005. The place of problem solving in contemporary mathematics curriculum documents. *The Journal of Mathematical Behavior* 24, 3–4 (2005), 341–350. DOI: 10.1016/j.jmathb.2005.09.004

[41] Ming-Te Wang and Jacquelynne S. Eccles. 2013. School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction* 28, (2013), 12–23. DOI: 10.1016/j.learninstruc.2013.04.002

[42] Allan Wigfield and Jacquelynne S. Eccles. 2000. Expectancy–Value Theory of Achievement Motivation. *Contemporary Educational Psychology* 25, 1 (2000), 68–81. DOI: 10.1006/ceps.1999.1015

[43] Allan Wigfield and Kathryn R. Wentzel. 2007. Introduction to Motivation at School: Interventions That Work. *Educational Psychologist* 42, 4 (September 2007), 191–196. DOI: 10.1080/00461520701621038

[44] Allan Wigfield, Susan L. Klauda, and Jenna Cambria. 2011. Influences on the Development of Academic Self-Regulatory Processes: Department of Human Development, University of Maryland. In the *Handbook of Self-Regulation of Learning and Performance*. Routledge.

[45] Philip H. Winne. 2006. How Software Technologies Can Improve Research on Learning and Bolster School Reform. *Educational Psychologist* 41, 1 (2006), 5–17. DOI: 10.1207/s15326985ep4101_3

[46] Philip H. Winne. 2010. Improving Measurements of Self-Regulated Learning. *Educational Psychologist* 45, 4 (October 2010), 267–276. DOI: 10.1080/00461520.2010.517150

[47] Robert M. Yerkes and John D. Dodson. 1908. The relation of strength of stimulus to rapidity of habit-formation. *J. Comp. Neurol. Psychol.* 18, 5 (November 1908), 459–482. DOI: 10.1002/cne.920180503

[48] Mingming Zhou and Philip H. Winne. 2012. Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction* 22, 6 (2012), 413–419. DOI: 10.1016/j.learninstruc.2012.03.004

[49] Barry J. Zimmerman. 1990. Self-Regulated Learning and Academic Achievement: An Overview. *Educational Psychologist* 25, 1 (January 1990), 3–17. DOI: 10.1207/s15326985ep2501_2

[50] Barry J. Zimmerman. 2000. Attaining Self-Regulation. In *Handbook of Self-Regulation*. Elsevier, 13–39. DOI: 10.1016/B978-012109890-2/50031-7

[51] Barry J. Zimmerman. 2008. Investigating Self-Regulation and Motivation: Historical Background, Methodological Developments, and Future Prospects. *American Educational Research Journal* 45, 1 (2008), 166–183. DOI: 10.3102/0002831207312909

[52] Barry J. Zimmerman and Dale H. Schunk (Eds.). 2011. In *Handbook of self-regulation of learning and performance.* Routledge/Taylor & Francis Group, New York, NY, US.