

Examining the critical role of evaluation and adaptation in self-regulated learning[☆]

Mladen Raković^{a,*}, Matthew L. Bernacki^b, Jeffrey A. Greene^b, Robert D. Plumley^b, Kelly A. Hogan^c, Kathleen M. Gates^d, Abigail T. Panter^c

^a Centre for Learning Analytics, Faculty of Information Technology, Monash University, Australia

^b School of Education, University of North Carolina at Chapel Hill, USA

^c College of Arts & Sciences, University of North Carolina at Chapel Hill, USA

^d Quantitative Psychology, University of North Carolina at Chapel Hill, USA

ARTICLE INFO

Keywords:

Self-regulated learning
Metacognition
Evaluation
Adaptation
Reflective writing

ABSTRACT

Researchers and many educators agree that the ability to self-regulate learning is important for academic success. Yet, many students struggle to anticipate learning difficulties and adjust accordingly. Further, despite theorizing that self-regulated learning involves adaptation across learning cycles, few researchers have examined students' evaluative judgments, their implications for students' behavior in a subsequent learning cycle, or their effects on achievement. Utilizing data from a large, introductory college biology course, we examined how struggling students' evaluative judgments made after a first unit exam predicted changes in learning behaviors as well as how those changes predicted performance on a subsequent exam. We used natural language processing to analyze data from a reflective essay written after a first unit exam, identifying language that reflected evaluation of prior studying and plans to adapt learning. Then, we utilized digital traces of learning behaviors to measure students' actual adaptation of their use of learning resources. Results from a path analysis revealed students' evaluations predicted how extensively they discussed plans to adapt their learning process. Plans to adapt described in written reflections predicted an increase in the frequency of desirable learning behaviors, which in turn predicted higher subsequent exam scores, after controlling for previous exam performance. These findings provide empirical evidence of multiple theorized self-regulated learning processes, including how evaluations of learning at the end of a learning cycle can inform planning and behavior changes in a subsequent learning cycle, and that increases in the enactment of effective learning strategies predict improved performance in complex learning tasks.

1. Introduction

The high dropout rate in STEM courses hinders efforts to expand the STEM workforce in the United States (Hamm et al., 2020; Olson & Riordan, 2012). The path towards dropout begins early in students' undergraduate studies (Tinto, 1993), and early performances in gateway courses serve as critical inputs that shape students' perceptions of their ability to succeed in their chosen STEM degree and career (Perez, Cromley, & Kaplan, 2014). This can determine whether they will continue to pursue their degree in turn (Flanders, 2017). To help

students succeed in those courses and complete their STEM degrees, many instructors utilize *active learning*, a student-centered educational practice that promotes engagement (Bonwell & Eison, 1991; Braxton et al., 2000; Prince, 2004). Active learning is a broadly applied term used by instructors across STEM disciplines and is generally understood to involve a proportional reduction in passive lecture pedagogies and shift towards activities that require students to engage more intentionally in learning activities as a way of constructing their own knowledge (Lombardi et al., 2021). Examples of these type of activities include answering questions while reading the textbook to ensure understanding

[☆] This research was supported by the U.S. National Science Foundation Award DUE-1821594. The opinions, findings, and conclusions, or recommendations expressed are those of the authors and do not necessarily reflect the views of the National Science Foundation.

* Corresponding author at: Research Fellow, Centre for Learning Analytics, Faculty of Information Technology, Monash University, 25 Exhibition Walk, Clayton, VIC 3800, Australia.

E-mail address: mladen.rakovic@monash.edu (M. Raković).

<https://doi.org/10.1016/j.cedpsych.2021.102027>

(as opposed to more passive reading), and rehearsing, explaining, and elaborating on their knowledge during homework, quizzes, practice exams, or inquiry-based group projects. Researchers have found that implementation of active learning in undergraduate STEM courses improves overall examination performance and narrows achievement gaps across STEM disciplines and class sizes (Freeman et al., 2014; Theobald et al., 2020).

However, courses that require active learning are rigorous and present a challenge to learners. Students sometimes respond negatively to the instructional design (Shekhar et al., 2020), and many struggle to determine how to best make use of the abundant learning activities and resources available to them (Miller & Bernacki, 2019). As a result, in active learning STEM courses, many students score lower than what they expected on early examinations, and many of these students then struggle to overcome these initial poor performances. This can undermine students' motivation and feelings of self-efficacy, which in turn negatively influence their approach to future learning (Bruinsma, 2004; Cook et al., 2013), including whether and how they self-regulate their learning (Zimmerman, 2013). Students' success at self-regulating their learning is positively associated with their academic performance (Dent & Koenka, 2016). More research is needed to understand how students in active learning STEM courses reflect upon poor performance on an exam, and whether and how that reflection relates to subsequent learning behaviors and performance.

In this article, we focus on students who fail to meet their own expectations on an initial course exam in order to understand how these students (1) evaluate their initial engagement in the course and (2) plan to adapt their learning. We then examine (3) how these students execute that plan and (4) how their future learning behaviors relate to future performance. This research can provide insight for practitioners who aim to support students' adaptation after initial struggles in their academic program. This research can also provide empirical evidence of the occurrence and importance of a critical phase that gives process models of self-regulated learning (SRL; Winne & Hadwin, 1998) their cyclical structure: the periodic evaluation of one's current approach to learning and the adaptation of that approach moving forward.

To succeed in an active learning course, students who underestimate their early performance will need to self-regulate their subsequent learning (Virtanen et al., 2017). They will not only need to re-appraise the demands of their course and to select those materials and activities that help them master learning objectives, but also, they will have to evaluate their past learning behaviors (i.e., judge what studying approaches did and did not work well for them; Winne & Hadwin, 1998) and make adaptations to ensure future learning behaviors will advance their learning and improve their performance. Both theory (Winne & Hadwin, 1998) and research (Binbasaran Tuysuzoglu & Greene, 2015; Pardo, Han, & Ellis, 2017; Wolters, Won, & Hussain, 2017) have suggested that students who initially underestimate their learning and performance will struggle to engage productively in SRL (Winne & Hadwin, 1998), often because of failures to make accurate evaluative judgments about their learning, and difficulty planning and enacting productive adaptations to future learning behaviors. In this study, we leverage an unequivocal, external evaluative judgment – the first exam grade in the course – and use students' reflections on their prior approach, exam performance, and future plans to understand how they described their initial execution (i.e., the first learning cycle corresponding to the unit 1 content on the exam), and how they plan to adapt their learning during the second cycle.

Further, whether and how evaluations of the results of prior studying relate to planning and learning enactment in subsequent learning cycles is a critical aspect of SRL that has not yet been thoroughly investigated (Greene & Schunk, 2017). Research to date has mostly focused on discrete actions or judgments reflecting cognitive or metacognitive processes within a single learning cycle (e.g., learning and studying behaviors in preparation for a course unit exam; Binbasaran Tuysuzoglu & Greene, 2015). Theory suggests that outcomes received at the end of

one learning cycle (e.g., an exam grade for the first unit in a course) should be evaluated and lead to productive adaptations at the start of the next learning cycle (e.g., making plans to learn and studying differently during the second unit in a course; Winne & Hadwin, 1998), but there is a lack of research supporting this presumption.

In this study, we sought to understand whether and how students adapt the way they learn after they receive an exam grade that is lower than what they expected. To do so, we focused our investigation on students who performed worse than they hoped on an exam, collecting their evaluations of their past learning approach, their plans for adaptation, and their subsequent execution of learning behaviors in the next learning cycle, all of which we used to predict the students' next exam performance. We were particularly interested in studying students whose first exam performance fell below their expectations because this initial poor performance provided them with a clear prompt to reflect upon and change their learning. After receiving their Unit 1 exam score, these students responded to prompts to compose written reflections comprising evaluation of their learning during the previous unit and plans to adapt their learning behavior in the subsequent content unit. We appraised their reflections by applying a novel and scalable natural language processing (NLP) method to compute a set of linguistic features that reflected their metacognitive evaluations and plans to adapt their learning. Digital trace data reflecting students' use of learning resources from the course learning management system (LMS) were used to measure adaptations to their learning behaviors during the second content unit. Then we assessed how evaluations, proposed adaptations, and actual changes in learning behaviors from the first learning cycle (i.e., prior to the first exam) to the second learning cycle (i.e., after the first exam through the second exam) predicted second exam performance.

Research on students' evaluations of their learning, how they plan to adapt that learning, and those evaluations' and plans' role in guiding future adaptation can benefit practitioners who aim to support students who underperform on their early exams in a college course. Early poor performances in gateway STEM courses pose significant challenges to the students (Cook et al., 2013). Our findings can inform future instructional efforts to help those students engage in productive SRL that positively affect subsequent achievements. Such efforts can include the design of interventions to help struggling students engage in more effective metacognitive evaluation of their learning and make more effective adaptations to their subsequent learning behaviors, allowing them to realize the powerful affordances of active learning pedagogies in undergraduate, introductory STEM courses.

2. Literature review

2.1. Self-regulated learning

There is ample empirical evidence that students SRL knowledge, skills, and dispositions predict their academic performance across a number of disciplines (Dent & Koenka, 2016), including STEM courses (Binbasaran Tuysuzoglu & Greene, 2015). SRL is a cyclical process (Winne & Hadwin, 1998) where learning cycles are marked by the evaluative judgements that can emerge from the learner or that can be imposed externally by the environment. For self-regulated learners who monitor for the effectiveness of their learning, i.e., generate metacognitive evaluations to determine whether the strategies are helpful, the learning cycle can be as short as the time span between initiation and evaluation, and this time span can vary across learners. The closure and next iteration of the learning cycle is often marked externally, typically by a criterion event such as an exam.

What constitutes a learning cycle is often idiosyncratic and specific to individual learners. However, in formal undergraduate educational contexts, learners proceed through a series of externally imposed learning cycles defined by a course schedule. In each content unit in a course, the learner is introduced to a topic and the learning objectives that guide instruction. Course sessions then proceed, until mastery of the

learning objectives is assessed by a summative unit exam.

Each of those cycles, at varying grain sizes, are comprised of processing that can be characterized across four loosely ordered stages: task definition, goal-setting and planning, strategy enactment, and reflection and adaptation (Winne & Hadwin, 1998). Self-regulated learners hence appraise the affordances provided and constraints imposed by learning tasks, set goals and create plans on how to accomplish them, enact learning strategies they believe will help them achieve their goals and monitor for the effectiveness of learning strategies. After a period of engagement, a learning cycle ends with a period of more summative evaluation of the learning process to this point, and learners cycle back to revisit their perception of their task, and whether a goal, plan, or enactment strategy needs to be revised in the next learning cycle. For instance, when an exam score is returned, self-regulated learners articulate and evaluate prior studying and make forward-reaching plans to improve their future studying in the next learning cycle that aligns to the next course unit.

Successful adaptations to learning behaviors are therefore contingent upon students' accurate evaluative judgements of prior learning and their ability to translate those judgments into productive subsequent learning behaviors (Winne & Hadwin, 1998). For this reason, examining learning over multiple rather than within a single learning cycle can shed light on theorized positive effects that evaluative judgements (e.g., metacognitive evaluation) and adaptation have on learning performance in subsequent learning cycles. The results of this analysis can add to limited research on SRL across multiple learning cycles and, in turn, support understanding and promoting of productive SRL behaviors to benefit students in active learning courses, particularly those students who, despite their best efforts, do not achieve in line with their expectations.

2.2. Metacognitive evaluations

Researchers who investigate cognitive and metacognitive processes of SRL have reported positive effects of many self-regulated learning processes, including students' time management, planning, achievement goal orientation, cognitive and metacognitive strategy use on performance outcomes (cf., Broadbent & Poon, 2015; Dent & Koenka, 2016). Most research that examines metacognitive processes observe events that reflect planning prior to task engagement, or monitoring during engagement in the task (e.g., Ariel et al., 2009; Azevedo & Cromley, 2004; Hadwin & Webster, 2013; Moos & Azevedo, 2008; Wolters et al., 2017). Less common are studies that focus on metacognitive processes that occur at a discretely observed adaptation stage of SRL (but see Binbasaran Tuysuzoglu & Greene, 2015; Pieschl, Bromme, Porsch, & Stahl, 2009). A substantial amount of the extant research on metacognitive processes has demonstrated the critical roles that processes during the planning phase of the SRL cycle have on subsequent processes and outcomes (Eilam & Aharon, 2003). Similarly, metacognitive monitoring during enactment of strategies and students' decisions to shift strategies have been key predictors of learning (Binbasaran Tuysuzoglu & Greene, 2015). Less commonly observed are the metacognitive processes at the end of the learning cycle where retrospective evaluations influence the planning and enactment that might be observed – and responsible for improved learning – during future cycles of self-regulated learning. When students evaluate learning and propose to adapt their future approach, prompting them to formalize these reflections in writing can reveal the details of their monitoring processes and the way these guide their selection and implementation of learning strategies. Learning theorists propose that this process is critical to optimize learning, and that such an evaluative practice is essential for learners to consolidate lessons learned about effective learning, so that they can be transferred across subsequent learning cycles, and also to future learning tasks that may span different domains (Hattie & Donoghue, 2016; Winne, 2018; Winne & Hadwin, 1998). In this study, we leverage the naturally occurring performance feedback that students

receive on their exam, which can act as a metacognitive monitoring prompt, as an opportunity to capture evidence of metacognitive evaluation.

2.3. Adaptation in Self-Regulated learning

Productive self-regulated learners consciously exercise cognitive and metacognitive processes to accomplish goals for learning (Winne, 2018), and over time, learners optimize their learning efforts by cyclically appraising and improving their approach en route to achieving their goal (Greene & Schunk, 2017). Learners' metacognitive monitoring and control (i.e., modifications to learning behaviors) are critical to each stage of the SRL cycle (Greene & Azevedo, 2007; Winne & Hadwin, 1998), where appraisals of task conditions cue are conducted in light of metacognitive knowledge and experience, and inform goal setting and planning. Metacognitive control is further exercised when this plan is executed, and thereafter adapted as learners adjust learning behaviors (Winne, 2014), contingent upon metacognitive monitoring judgements that appraise whether the strategies they have enacted are appropriately advancing them towards their stated learning goals (Livingston, 2003). In classroom contexts, when learning products such as exam grades indicate that students' efforts have failed to meet students' academic expectations, self-regulated learners may modify levels and types of learning engagement in the next learning cycle to ensure they achieve performance that satisfies their academic aspirations (Greene, Muis, & Pieschl, 2010).

Binbasaran Tuysuzoglu and Greene (2015) investigated this contingency between metacognitive monitoring and control and how students' adaptation after making negative judgments of learning predicted their achievement during a science learning task (Binbasaran Tuysuzoglu & Greene, 2015). They found support for the theoretical assumption that exerting metacognitive control and adapting learning strategy selection in response to a negative metacognitive monitoring judgment is critical to successful learning. Students who metacognitively judged their learning strategy to be insufficient and subsequently chose to adopt a new strategy (i.e., they exercised a negative metacognitive judgment and then adaptive metacognitive control) achieved higher scores on a posttest to students who made similar negative judgments of learning, but chose instead to sustain the same metacognitive control strategy that had provided insufficient progress towards their learning goal to that point.

Pieschl et al. (2009) also found that students who adapted their learning by increasing desirable learning behaviors in order to deal with complex learning task outperformed their less adaptive counterparts. Both studies demonstrate the importance of adaptation within a reasonably brief learning task conducted in laboratory settings (Pieschl et al., 2009). To date, little effort has been made to investigate how students adapt their learning in authentic, complex tasks that are common to challenging undergraduate coursework. The observation of student adaptation when self-regulating learning during authentic learning tasks provides an opportunity to establish the relevance of SRL theory to authentic contexts, and further, to observe how the cycles imposed by the design of learning tasks such as college courses can cue evaluation, adaptation, and self-regulated learners' refinement of their learning process. We elected to investigate students' evaluation and adaptation using evidence from their reflective writing when prompted to consider their learning and achievement in the first unit of their biology course, and how these reflections related to the behaviors they conducted in a future learning cycle: the second unit.

2.4. Reflective writing in education

Educators in many academic disciplines rely upon reflective writing tasks to prompt and document students' metacognition (Gibson, Kitto, & Bruza, 2016). For instance, in a reflective writing assignment, students may be asked to recall the course resources they used in previous

learning cycle, judge what worked or did not work well for them, and propose changes to their future learning (e.g., [Tanner, 2012](#)). The reflective writing assignments are often administered after exams to help students increase awareness of their learning ([Medina et al., 2017](#)). Previously, researchers have found that students who engaged in reflective writing were more accurate in evaluating their learning and achieved a better understanding on what should be improved in the future ([Allan & Driscoll, 2014](#); [O'Loughlin & Griffith, 2020](#)), motivating more future research on the relationship between reflective writing and learning performance. In these reflective writing assignments, students produce samples of written content that is a valuable resource for researchers and instructors who want to examine students' metacognition on different topics. However, the manual examination of students' reflections by researchers and instructors has usually been a labor-intensive process ([Cui et al., 2019](#); [Ullmann, 2017](#)). With recent advancements of NLP computational methods, the opportunities arise to scale up the analysis of reflective writing and efficiently understand students' reflective texts.

A few computational models have been developed in recent years to analyze metacognitive activity in reflective writing. They revealed the potential of NLP in supporting research on metacognition. [Luo and Litman \(2016\)](#) designed a rubric-based approach to automatically analyze the quality of students' metacognitive reflective responses with respect to active and constructive mode of [Chi's \(2009\)](#) ICAP framework. This method involved the identification of text content through a comprehensive list of keywords that relate to course content and organization, and the development of a prediction model that demonstrated accuracy and scalability across different lectures, topics and courses. [Gibson et al. \(2016\)](#) created an NLP algorithm based on predefined linguistic rules to identify metacognitive knowledge, regulation and experience, and compute metacognitive activity in reflection. The algorithm relied upon the multifaceted model of metacognition ([Efklides, 2008](#)) and has been shown to successfully discriminate between instances of strong and weak metacognitive activity in reflective writing.

[Raković, Winne, Marzouk, and Chang \(2021\)](#) developed a computational approach using content and rhetorical features of text to detect knowledge transforming (i.e., elaborating and integrating source knowledge through successive planning and revising) in argumentative essays developed from multiple sources. The algorithm was accurate in identifying knowledge transforming sentences in undergraduate argumentative essays, offering potential for future research. This approach was based upon a theoretical work on psychological processes that co-occur in writing ([Bereiter & Scardamalia, 1987](#); [Scardamalia et al., 1984](#)). According to [Scardamalia et al. \(1984\)](#), analytical and reflective thought in writing develops through interaction between operations in content and rhetoric problem spaces in the text. In this study, we captured student metacognitive processes based on the same theoretical principle, but in a different type of writing task (reflective instead of analytical).

We thus focused our algorithmic approach on identifying relevant content in students' reflections (e.g., students' recollection of learning activities and resources they utilized in a previous learning cycle) and also on capturing how this content was rhetorically presented (e.g., student's justification why they engaged in a particular activity), a methodological benefit to SRL research. We measured two kinds of metacognitive processes from students' written responses to a series of prompts in the reflective writing assignment: metacognitive evaluation of prior studying and intentions to adapt to future studying. We defined a set of rhetorical rules specific to each prompt to accurately measure metacognition in a response provided and distinguish between evaluative and adaptive language. For example, in response to a prompt asking a student to evaluate their learning activities before class, an utterance qualifying as an event of metacognitive evaluation of prior studying will not only contain references to relevant course resources and/or learning tactics, but will also provide a rhetoric that reflects a reason for choices a student made (e.g., by using phrases for reasoning such as *due to*,

because...). In the context of prompt asking a student to describe how they intend to study for the next exam, metacognitive adaptation would be an utterance that arouse when a student made a rhetorical move in their writing by using e.g., a sequential phrase *next*, while describing how do they plan to engage with course resources and tactics in the future. Details about our approach, including reflective prompts and predefined content and rhetorical features, are provided in the Method section.

2.5. Measuring SRL behaviors using student data from online platforms

In addition to measuring students' evaluations of their learning and intentions to adapt, we used digital traces of learning events to observe the frequency and variety of learning events in the first and second units of the course, and the ways they reflected behavioral adaptations in students' learning process. In alignment with recommendations from learning theorists regarding the varieties of data necessary to substantiate and investigate self-regulated learning events ([Winne, 2018](#)). [Bernacki \(2018\)](#) explains how digital traces are created when students interact with technology-enhanced learning platforms (e.g., click the link/button, download a document, input a value into a text field) and how such interactions reflect learning events. The interactions are logged by the software as timestamped events, which are stored in a database for use and can be enriched with appropriate metadata about the digital learning objects that students utilize, and the kinds of self-regulated learning processes those objects are designed to support ([Bernacki, 2018](#)).

The digital traces of student learning in a biology course that hosts learning resources on an LMS course site might include events such as use of resources for planning (e.g., accessing study guides or sample exams), strategy use (e.g., attempting practice quizzes for self-assessment and retrieval practice) or monitoring (e.g., accessing learning objectives or quiz feedback). Students who complete training that promotes self-regulated learning and strategy use tend to make greater use of these materials ([Bernacki, Vosicka, & Utz, 2019](#)), and those who use resources specifically designed to support metacognitive processes tend to perform better in their undergraduate science coursework ([Hong, Bernacki, & Perera, 2020](#)).

[Greene et al. \(2019\)](#) provide corroborating evidence that the temporality of observed engagement in self-regulated learning behaviors including course management (e.g., accessing announcements, lecture notes, syllabus), information acquisition (e.g., accessing guided reading questions, reviewing current semester exam) and metacognitive activities (accessing personal gradebook) changes from one unit to the next, and when events are conducted early in the semester, they predict a students' performance in the course ([Greene et al., 2019](#)). Whereas these studies demonstrate the external validity of digital traces with respect to academic performance and their variance between learners and over units of learning, these data have yet to be examined alongside data that reflect students' reflections upon and evaluations of their prior learning, or their intentions to sustain or adapt their approach to learning in future course units.

2.6. The current study

Discrepancies between products of learning and learners' expectations should trigger subsequent metacognitive monitoring through evaluation of prior studying ([Winne, 2018](#); [Winne & Azevedo, 2014](#); [Winne & Hadwin, 1998](#)). As a response to the evaluation, self-regulated learners engage in the development of specific, forward-reaching plans for similar learning tasks in the future ([Winne & Azevedo, 2014](#)). Consequently, self-regulated learners modify their learning behaviors in future learning cycles ([Zimmerman, 1990](#)) by choosing to engage in behaviors that are more likely to boost exam performance, while avoiding behaviors that may not benefit achievement. Hence, from one learning cycle to the next, self-regulated learners make a shift to more

productive learning behaviors (Winne, 2018). In order to investigate the common and problematic trend of poor initial achievement in early undergraduate coursework that contributes to dropout from the STEM pipeline (e.g., Flanders, 2017), we sought to add to the sparse research on what happens after students whose first exam performance fell below their expectations, including how these students evaluated their learning and made plans for adaptation, how they actually adapted their learning behaviors in a subsequent learning cycle, and how the contingency between evaluation and adaptation affected future achievement.

We investigated SRL in a large-scale introductory, active learning STEM course. The course included three unit exams and one final, cumulative exam. After the Unit 1 exam, students were asked to reflect, in writing, on their learning behaviors during the previous learning cycle (i.e., evaluation) and describe how they planned to study for the Unit 2 exam during the next learning cycle (i.e., adaptation). From the responses we gathered, we computed the volume of metacognitive evaluation and plans to adapt. We also collected LMS trace data about students' learning behaviors during the first and second learning cycle to identify theory-aligned learning behaviors that were likely to have a positive (i.e., desirable behaviors) and not substantially positive (i.e., other behaviors) relationship with subsequent achievement, and computed adaptation to those behaviors during the second learning cycle, during Unit 2. Last, we gathered achievement data on Unit 1 and Unit 2 exams. Achievement and learning behaviors in the remainder of the course were not analyzed in this study.

Motivated by previous research, we posited the relations represented in the path model in Fig. 1, with the following five hypotheses to guide our study.

H1: Exam 1 score will have a negative relationship with the volume of student's metacognitive evaluation in reflective writing.

H2: The volume of metacognitive evaluation will be positively related to the volume of plans to adapt.

H3: The volume of plans to adapt will be positively related to increase in frequency of desirable behaviors from the Unit 1 learning cycle to the Unit 2 learning cycle.

H4: Increase in frequency of desirable behaviors from the Unit 1 to the Unit 2 learning cycle will be positively related to Exam 2 score after controlling for Exam 1 score.

H5: There will be a positive indirect effect of volume of metacognitive evaluation on the frequency of desirable behaviors through the volume of plans to adapt.

We additionally posed a general research question (RQ1) and examined how additional traces of behavioral engagement (other behaviors) would associate with the volume of metacognitive evaluation, plans to adapt, and subsequent learning outcomes.

3. Methods

3.1. Participants

A total of 468 students enrolled in a first-year undergraduate introductory biology course at a Southeastern public university consented to participate in research projects with their course. Of these, 128 students participated in an alternate experiment, conducted before Exam 1 that trained students how to use particular learning strategies (e.g., self-testing, self-explaining, help seeking) to promote achievement in the remainder of the course. We excluded these students from consideration in order to avoid any influence of those activities on a subset of our sample. Thereafter, we included all of the remaining 340 students who (1) completed the activities from which our focal variables were drawn (i.e., a survey prior to Exam 1, Exam 1, a reflective writing assignment after Exam 1, and Exam 2) and (2) earned the Exam 1 grade that was lower than the grade they predicted they would achieve on the exam. This produced a focal sample of 109 students.

Our final analytic sample of students represented 23% of the class. Members' mean age was 18.62, 88 were women, 19 were men, 2

students did not declare their gender, and 39 students identified as members of under-represented minority groups (14 African American, 13 Hispanic, 14 multiracial, 1 unknown ethnicity). Twenty-seven were first-generation college students. Those retained in the sample did not differ from those excluded on any demographic variable, including mean age or proportional representation by gender or by minority or first generation status. Those included in the sample scored lower on a measure of prior knowledge, which may be expected given the selection criteria¹.

Of the 109 students in our focal sample, 49 students were in the course section 1 and 60 students in the course section 2. No differences were observed between sections on demographic variables (i.e., age, gender, proportion of students from underrepresented groups and first generation status) or on the prior knowledge measure (Tables 1–3).²

3.2. Context

Biology 101 (Bio101) was an introductory course designed to be an undergraduate's first course in the sequence of courses covering biology content. Course topics included Biochemistry, Cell Biology, Genetics, Molecular Biology, Anatomy, Physiology and Biodiversity. This course was required for Biology majors and served as a requirement for most other students enrolled in the course who were not Biology majors. Students were required to achieve a final course grade of at least 76.9% to proceed in their major.

Multiple sections of the course are offered in each semester, and the instructional team who delivers these courses have developed a high structure active learning design that includes face-to-face sessions and a digitally rich learning management system course site. All sections of the course follow a parallel schedule that maintains the scope and sequence of course topics. All sections used the same master design of the LMS course site, and all exams were written by the instructional team as a group, who co-wrote items that corresponded to the course learning objectives drawn from a shared table of specifications for exam design.

The instructors integrated many course activities into the LMS, and the trace data students produce when they engage with digital resources provide an opportunity to observe some of their learning behaviors. In each learning cycle corresponding to a lecture topic, students were provided with opportunities to access guided reading questions they would answer while reading the textbook prior to lecture, complete a chapter quiz to self-assess their knowledge before attending the lecture, and download an outline of the topic to be covered in class. Students could return to the LMS site to download complete notes after the lecture session, and in preparation for the exam, students could access their syllabus with learning objectives, documentation on the exam design and coverage, past exams they could use as guides and a practice exam they could use to rehearse their knowledge and test-taking. Students also had opportunities to engage in online discussion forums and use tools to schedule appointments with the instructor, or academic coaching or

¹ For students who underachieved vs. those who did not: age, $t(176) = 0.93$, $p = 0.35$, and proportion of students from underrepresented groups, $\chi^2(1, N = 172) = 2.53$, $p = 0.11$. The differences in gender, $\chi^2(1, N = 178) = 18.62$, $p = 0.00$, first generation status, $\chi^2(1, N = 178) = 14.12$, $p = 0.00$, and prior knowledge, $t(175) = 2.75$, $p = 0.00$ between students who underachieved and those who did not were statistically significant (Table 3). For completed vs. did not complete all activities: age, $t(328) = -0.80$, $p = 0.43$, prior knowledge (test administered on the first day of classes), $t(322) = -1.77$, $p = 0.08$, proportion of students from underrepresented groups $\chi^2(1, N = 318) = 0.32$, $p = 0.57$, first-generation status, $\chi^2(1, N = 330) = 0.67$, $p = 0.41$ and gender, $\chi^2(1, N = 330) = 1.90$, $p = 0.17$ (Tables 1 and 2).

² For section 1 vs section 2 students: age, $t(107) = -0.93$, $p = 0.36$, prior knowledge (test administered on the first day of classes), $t(107) = 1.74$, $p = 0.08$, proportion of students from underrepresented groups $\chi^2(1, N = 107) = 1.40$, $p = 0.24$, first-generation status, $\chi^2(1, N = 107) = 0.11$, $p = 0.75$ and gender, $\chi^2(1, N = 107) = 0.21$, $p = 0.65$.

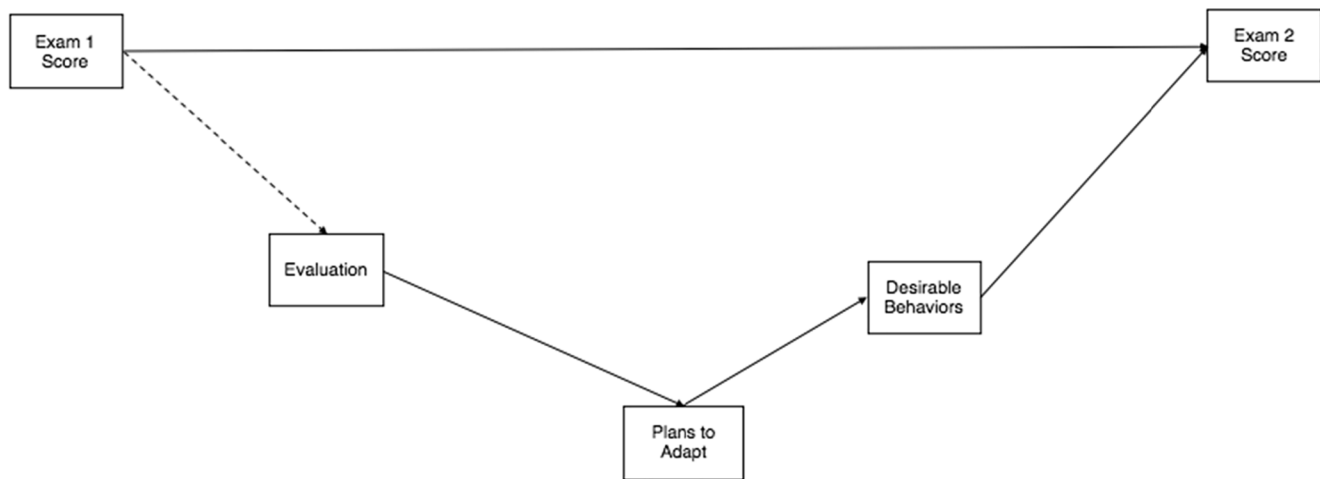


Fig. 1. The Hypothesized Path Model. Note. Solid lines depict positive and dashed lines depict negative relationships. Exam 1 is Unit 1 exam; Exam 2 is Unit 2 exam; Evaluation is volume of metacognitive evaluation in student reflection; Plans to Adapt is volume of plans to adapt described in student reflection; Desirable behaviors are learning behaviors that, if attempted more in Unit 2, positively affect Exam 2 scores.

Table 1
Sociodemographic Characteristics of Participants.

Baseline Characteristic	Completed all activities		Did not complete all activities		Full Sample	
	n	%	n	%	n	%
Gender						
Female	124	68.5	94	59.1	218	64.1
Male	54	29.8	58	36.5	112	32.9
Not provided	3	1.7	7	4.4	10	3.0
From underrepresented group						
Yes	30	16.6	30	18.9	60	17.6
No	142	78.5	116	72.9	258	75.9
Not provided	9	4.9	13	8.2	22	6.5
First generation status						
Yes	29	16.0	31	19.5	60	17.6
No	149	82.3	121	76.1	270	79.4
Not provided	3	1.7	7	4.4	10	3.0

Note. Participants who completed all activities were on average 18.8 years old ($SD = 1.5$) and participants who did not complete all activities were on average 18.9 years ($SD = 1.9$). Average prior knowledge test score of students who completed all activities was 57.5 ($SD = 12.5$) and of students who did not complete all activities was 60.0 ($SD = 12.7$) out of 100 points.

Table 2
Overview of gender, underrepresented group status, first generation status, age and pretest score between students who did and those who did not complete all the activities.

Baseline Characteristic	Completed all activities	Did not complete all activities	Statistical Test	df
Gender				
Female	124	94	$\chi^2 = 1.90^{ns}$	1
Male	54	58		
From underrepresented group				
Yes	30	30	$\chi^2 = 0.32^{ns}$	1
No	142	116		
First generation status				
Yes	29	31	$\chi^2 = 0.67^{ns}$	1
No	149	121		
Age	18.75	18.90	$t = -0.80^{ns}$	328
Pretest score (out of 100)	57.51	60.00	$t = -1.77^{ns}$	322

Significance codes: $ns\ p > 0.05$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Table 3
Overview of gender, underrepresented group status, first generation status, age and pretest score between students who underachieved and those who did not.

Baseline Characteristic	Underachieved	Did not underachieve	Statistical Test	df
Gender				
Female	88	36	$\chi^2 = 18.62^{***}$	1
Male	19	35		
From underrepresented group				
Yes	23	7	$\chi^2 = 2.53^{ns}$	1
No	84	58		
First generation status				
Yes	27	2	$\chi^2 = 14.12^{***}$	1
No	80	69		
Age	18.69	18.91	$t = 0.93^{ns}$	176
Pretest score (out of 100)	55.86	61.38	$t = 2.75^{***}$	175

Significance codes: $ns\ p > 0.05$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

tutoring services to obtain guidance on how to learn (though no direct instruction on the science of learning was offered in the course itself). Students' interactions with the LMS features produced the trace data used in this study.

3.3. Measurement

3.3.1. Exam 1 Grade Expectation

Prior to Exam 1, we emailed eligible students a short survey asking what grade they expected to earn on the approaching exam. Students that earned a grade below what they expected were included in the sample.

3.3.2. Exam Scores

Unit 1 and 2 scores were obtained from the respective exams. The exams consisted of multiple-choice and short answer questions and tested for recall, understanding, and application of biochemistry and cell biology concepts (i.e., Unit 1 material) and genetics and molecular biology concepts (i.e., Unit 2 material). The exam content was aligned across the two course instructors, for both exams. Each exam was scored out of 100 points and each was worth 25% of the final grade in the course. We found statistically significant differences in Exam 1, $t(107) = 5.82$, $p = 0.00$; $d = 1.13$, and Exam 2 scores, $t(107) = 2.76$, $p = 0.00$; $d = 0.53$, between the two class sections. Therefore, we controlled for differences by section in our analyses.

Table 4

Prompts, SRL Constructs and Linguistic Features of Evaluation and Adaptation in Reflective Writing.

Prompt	SRL construct	Linguistic features	
		Content	Rhetoric
1. List the activities you consider most important in preparing for class. In a few sentences, justify the activities you selected.	Metacognitive evaluation	Learning tactics, resources	purpose, reason, causal-general, causal-causal-conditional
2. Reflect on class time. What strategies are you using during class time that you consider to be most helpful? Why?	Metacognitive evaluation	Learning tactics, resources	purpose, reason, causal-general, causal-causal-conditional
3. Reflect on your post-class reviewing. In a few sentences, describe how and when you review learning materials covered before and during class.	Metacognitive evaluation	Learning tactics, resources	All
4. Reflect on your exam studying. In a few sentences, describe the best strategies you used to study for the exam.	Metacognitive evaluation	Learning tactics, resources	purpose, reason, causal-general, causal-causal-conditional
5. Describe how you intend to study for the next exam. What about your approach will stay the same and what will change?	Metacognitive adaptation	Learning tactics, resources	Simultaneous, additives-emphatic, additives-appositional, sequential

Table 5

The Dictionary of Content and Rhetorical Features Used in the Reflection Scoring Algorithm.

Content features		Rhetorical features
Strategies	Resources	
read, reread, re-read, self test, self-test, test myself, quiz myself, self-question, question myself, question, create question, review, revise, monitor, comprehend, understand, do, submit, complete task, seek help, gather course information, collect, plan, reflect	syllabus, schedule, guided reading question, GRQ, textbook, course textbook, instructions, resources, mastering question, question, note, slide, lecture note, lecture slide, lecture, problem, mastering biology, ebook, learning catalytics, outline, power point slides, old exams, supplemental material, announcements, email, additional reading, handouts, study guides, supplemental instruction, peer mentor, peer mentoring, peer, piazza, online discussion, discussion, forum, office hours, classmate, mate, colleague, friend, required reading, homework, assignment, animation, tutorial, internet, web, quiz, MB, LC, learning catalytics, practice exams, group work, team work, group, team, answer, group activities, in class, in-class, exam, multiple choice, short answer, class discussions, bio tutoring, BioCell, tutor, mentor	• Purpose (so that, in order that, in order to) • Reason (since, as, because, now that) • Simultaneous (while, as) • Conditional (if, in case, provided that) • Concessive (although, though, even though) • Substitives (instead, rather, rather than) • Additives-emphatic (in addition, moreover, furthermore, besides, also) • Additives-appositional (that is, in other words, for instance, for example) • Additives-comparative (likewise, similarly, equally) • Causal-general (therefore, consequently, for that reason, thus) • Causal-causal-conditional (then, in that case, otherwise) • Sequential (next, first, second, last, finally, up to now, to sum up)

3.3.3. Reflections on Prior Studying and Plans to Adapt

A reflective writing survey was administered one week after receiving Exam 1 grade. The five open ended prompts were included in the survey. Students were asked to reflect on the most important learning activities and strategies they had chosen to 1) prepare for class, 2) do their work in class, 3) review course materials after class and 4) study for the Exam 1. In prompt 5, students were asked to describe how they planned to study for Exam 2. The minimal allowed text entry per prompt was 500 characters. In Table 4, we list reflection prompts, related SRL constructs, and linguistic features relevant to measuring SRL constructs, given each prompt.

3.3.4. SRL Behaviors

We obtained an extensive record of student trace data from the course LMS, which we used to model SRL behaviors. The LMS platform was a major repository of course resources, such as the syllabus, course calendar, announcements, class outlines, lecture notes and slides, guided reading questions (GRQ), additional readings, practice exams, information about supplemental instruction and links to external resources (e.g., discussion forums or booking system for office hours). We obtained a count of each behavior (e.g., the number of times a student clicked the announcement, downloaded syllabus, submitted GRQ), in units 1 and 2,

separately.

3.4. Data Preparation

3.4.1. Exam 1 Underachievement

We obtained a difference between expected and earned grade on the Exam 1 for each eligible student by subtracting a score earned on the Exam 1 from expected grade converted into a percentage, following the course grading plan, A = 93, B = 83, C = 73, D = 63, F = 53. For example, a difference of -4 indicates that the student earned 4 percentage points lower score than they expected. We determined the sample for the study by selecting students who had negative values for this measure.

3.4.2. Automatically Scoring a Volume of Metacognitive Evaluation and Plans to Adapt

Building on previous computational approaches (Gibson et al., 2016; Luo & Litman, 2016; Raković et al., 2021), we developed a system to automatically compute the volume (i.e., normalized frequency of words matching a predefined set of words) of metacognitive evaluations and plans to adapt in student responses to the reflective prompts. The system examined content and rhetorical features of the reflective responses.

First, relevant content features were captured by the algorithm using a dictionary of learning tactics and course resources we predefined for this study (Table 5). We utilized the Natural Language Toolkit (Loper & Bird, 2002) of the Python programming language to identify content features, including their different lexical forms that may occur due to flexibility in open-ended responding (e.g., “read instructions” was equivalent in our analysis to “reading instruction” or to “read instruction”). Second, we accounted for rhetorical features in reflections by capturing rhetorical connectives, i.e., phrases that bind together different pieces of content information in a cohesive and focused discourse (Carlsen, 2010). The rhetorical connectives we looked at in this analysis were categorized in 12 groups defined by Celce-Murcia and Larsen-Freeman (1983; see Tables 4 and 5). We assigned each prompt a set of rhetorical features given the expected rhetorical characteristics of an answer.

To provide an example of our method, imagine a highly metacognitive student responding to prompt 1 (i.e., List the activities you consider most important in preparing for class. In a few sentences, justify the activities you selected.). This prompt was designed to elicit metacognitive evaluation (i.e., “most important” “justify the activities”), so rhetorical connectives for purpose, reason, or cause would be expected to occur in their answer (e.g., *I find that doing the GRQs and reading the textbook were the most helpful, because even if I sometimes skipped around to find an answer, I still got an understanding of the bigger concepts needed for the classes*). Of course, appropriate rhetoric alone is not a sufficient indicator of a student’s metacognitive activity when reflecting on their studying. As shown in the example above, the student must also entwine relevant content into this rhetorical shell. In this example, they spoke about learning tactics and resources they opted for to study before class (i.e., *reading textbook* and *doing GRQs*).

We relied upon sentences as the unit of analysis. Compared to paragraphs, sentences as informational blocks afford the opportunity for more fine-grained analysis of information in texts (Bransford et al., 1972). For each answer, we computed the incidence of content and rhetorical features, regardless of the relationships among them, in each sentence and added them together to calculate the metacognitive activity score for the entire answer.

As learners’ reflections varied in their length, and the length of written passages has implications for the interpretation of raw amount of language that reflects evaluation and adaptivity, we chose to use a normalized interpretation in order to avoid inflated estimates of such a language due to differences in students’ verbosity. Because evaluation responses were longer than the adaptation responses, normalization also enabled us to adopt the same scaling and eased interpretation across these two variables. The metacognitive activity score was thus normalized by calculating a proportion of content and rhetorical features relative to the number of sentences in the answer. We obtained five such scores. We add scores together for prompts 1–4 to compute a volume of metacognitive evaluation variable. The score obtained for prompt 5 represented the volume of a student’s plans to adapt. To be sure normalization did not unduly influence our results, we ran competing models that employed raw versions of the same variables. The pattern of significant direct relationships was sustained across models, and we report using normalized values accordingly.

The automated scoring approach we used in this study was the same as one tested previously with undergraduate students completing a similar task (Raković et al., 2021). In that study, automated scoring of content and rhetorical linguistic features demonstrated considerable accuracy (i.e., 72.6%) compared to human scorers in terms of identifying whether a student told (i.e., restates information) or transformed information (i.e., provides a deeper account of restated information, for example through reasoning or applying). Based on this accuracy and the similarity in tasks between this study and the previous one, we were confident the automated scoring algorithm would produce findings with sufficient validity evidence to warrant interpretation.

3.4.3. Adaptation to behaviors

From among all the trace data collected, we identified SRL behaviors a student would be more likely to benefit from if they increased their engagement in them during Unit 2. We also identified SRL behaviors that may not be as helpful for improving achievement if the student additionally engaged in them during Unit 2. To identify these behaviors, we utilized (Greene & Azevedo, 2009) framework of SRL micro-level processes (i.e., specific SRL activities that emerge when students learn using hypermedia environments). The SRL micro-level activities are categorized into five higher-order macro-level groups: planning, monitoring, strategy use, handling of task difficulty and demands, and interest activities. We adapted (Greene & Azevedo, 2009) micro-level codes to fit the affordances of course LMS. Thus, we developed 16 categories of SRL micro-processes and mapped LMS behaviors accordingly. For example, if a student opened the document with assignment instructions, we inferred they were engaged in the micro-level process of task knowledge acquisition; if they submitted a request for office hours, we inferred they were engaged in help seeking; see Appendix A for complete list of SRL micro-processes and related LMS behaviors. Behaviors that could not fit this theoretical scheme (e.g., *submitting consent to participate in the research study*), were not included in the analysis.

Next, we situated the SRL micro-level processes into the specifics of course design (Bernacki et al., 2019). Given the learning objectives instructors defined for Unit 2 and the learning resources instructors provided to encourage Unit 2 studying, we deemed the following five groups of SRL micro-processes critical for success in Unit 2: task knowledge acquisition, content knowledge acquisition, submitting assigned work, self-testing and metacognitive monitoring. In other words, these categories defined the behaviors that students should enact more frequently to meet Unit 2 learning objectives (i.e., desirable SRL behaviors). On the other hand, additional engagement in some

Table 6

Behaviors Mapped to SRL Micro-Processes and Bivariate Correlations with Exam 2 Score.

Behavioral Engagement	Behavior	SRL micro-level process, articulated from Greene and Azevedo (2009)	Correlation with Exam 2 score
Desirable behaviors	View assignment instructions	task knowledge acquisition	0.15
	Review results of a submitted practice exam	metacognitive monitoring	0.12
	Submit practice exam answers	self-testing	0.07
	Review slides presented during lecture (posted after class)	content knowledge acquisition	0.06
	Finalize and submit a response	submitting assigned work	0.02
	Download the GRQ word doc	task knowledge acquisition	0.02
Other behaviors	Download additional readings provided for some lessons	content knowledge acquisition	−0.33**
	View course calendar in Sakai	course information gathering	−0.13
	Download file for supplemental instruction	help seeking	−0.09
	Download course syllabus	course information gathering	−0.08
	Click on a Lessons page in Sakai (this page provides the lesson learning objectives, not the lesson content)	course information gathering	−0.07

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

micro-processes of SRL may not necessarily benefit or may even harm Unit 2 achievement (i.e., other SRL behaviors). For instance, it may be too late to additionally engage in course information gathering or help seeking behaviors in Unit 2 as, obviously, more of these behaviors should be exercised in Unit 1.

We obtained a count of each behavior, per student, per unit. We computed a volume of adaptation to behavior by subtracting the Unit 1 count from the Unit 2 count. We conducted a bi-variate Pearson correlation analysis between adaptation to behaviors in Unit 2 and Exam 2 score to confirm the theoretical and instructional expectations for desirable and other SRL behaviors in Unit 2. The signs of the correlational coefficients we obtained for listed behaviors matched most of our expectations (Table 6). The only exception was the behavior *Downloading additional readings provided for some lessons* which, although classified as content knowledge acquisition (i.e., desirable SRL behavior), negatively correlated with Exam 2 score. This somewhat counterintuitive finding may have occurred because students infrequently downloaded additional readings during Unit 2 (i.e., $M = 0.109$, $SD = 0.045$) and because these additional readings did not specifically align to the learning objectives within that unit, and may have also included access of readings that corresponded to topics from other units. Further, the students who download and engage with these readings are likely spending more time and attention on unassessed topics than their peers, and this shift in their effort on additional course materials that extend beyond the scope of unit 2 exam could lead to lower performance on topics that are assessed on the exam.

Behaviors that correlated with Exam 2 score with values close to 0 (i.e., within ± 0.02) were excluded from analyses. We summed the volumes of Unit 2 adaptations across desirable and other behaviors and created two corresponding, normalized variables for analysis.

4. Results

Table 7 shows descriptive statistics and correlations for major study variables. The Shapiro-Wilk's normality test was statistically significant, indicating non-normal distributions. To mitigate the violation of normality, robust maximum likelihood estimation with Satorra-Bentler corrections (Satorra, 1992) was used when fitting the path model. The hypothesized model (Fig. 1) was computed using *lavaan* (Rosseel, 2012) package implemented in R language.

We used Mahalanobis distance to evaluate our data for multivariate outliers, finding five outliers that exceeded the 95th percentile. We winsorized values for these cases by setting their values to the value at the 95th percentile. Our initial data screening also warranted controlling for class section, as we found unexpected differences in exam scores across sections [Exam 1: $t(107) = 5.82$, $p < 0.001$, $d = 1.13$; Exam 2: $t(107) = 2.76$, $p = 0.006$, $d = 0.53$]. We tested our model with the Exam 1 variable, a section dichotomous variable, and an interaction term comprised of the two variables as additional predictors of each endogenous variable (i.e., evaluation, plans to adapt, desirable behaviors, and Exam 2 score). These additional predictors were statistically significantly related to desirable behaviors, only. Including those predictors in our model did not change the focal relations of interest (i.e., our hypotheses). We chose to retain the additional predictor variables in the final path model presented, for completeness. The final path model

obtained good data-model fit $\chi^2 = 13.08$ ($df = 14$), $p = 0.52$; CFI = 1.00, RMSEA = 0.00, SRMR = 0.06. The final path model is presented in Fig. 2.

The hypothesized relationship between Exam 1 score and volume of metacognitive evaluation (H1) was not statistically significant in final model, ($\beta = 0.08$, $p = 0.41$), indicating that Exam 1 score did not predict the volume of metacognitive evaluation in reflective writing. The hypothesized relationship between volume of metacognitive evaluation and volume of plans to adapt (H2) was statistically significant and positive ($\beta = 0.49$). In other words, students whose text included more metacognitive evaluation about their prior studying were more likely to write specific, forward-reaching plans about future studying. The volume of plans to adapt was positively related to an increase in the frequency of desirable behaviors from Unit 1 learning cycle to Unit 2 learning cycle ($\beta = 0.22$), supporting hypothesis H3. An increase in the frequency of desirable behaviors positively predicted Exam 2 score ($\beta = 0.24$) after controlling for Exam 1 score, supporting hypothesis H4. The results for hypothesized direct effects are provided in Table 8. The indirect path from evaluation to desirable behaviors through the volume of plans to adapt was positive and statistically significant, ($\beta = 0.11$, $p = 0.02$), supporting hypothesis H5 (Table 9).

The results we obtained for RQ1 (Table 10) indicate that the volume of plans to adapt was not statistically significantly related to the frequency of other behaviors ($\beta = 0.12$, $p = 0.18$), and an increase in frequency of other behaviors was not statistically significantly related to Exam 2 score ($\beta = -0.03$, $p = 0.68$). In addition, the indirect path from evaluation to other behaviors through the volume of plans to adapt was not statistically significant ($\beta = 0.06$, $p = 0.17$). We retained the Other Behaviors variable in the model for completeness.

5. Discussion

Despite ample research into many of the claims of SRL theory (Winne & Hadwin, 1998), few researchers have investigated the key assumption that reflection after completing a task should inform changes to SRL processing in subsequent tasks. To address this gap, we aimed to understand how students who performed worse than they expected in the first unit exam evaluated their studying and how those evaluations related to plans to adapt and subsequent learning behaviors and achievements in an authentic context of a challenging undergraduate STEM course. We computed metacognitive evaluation and plans to adapt from students' reflective essays and captured changes in SRL processing during the subsequent unit using LMS trace data. Our path analysis results supported the key assumption of SRL theory that evaluation predicts adaptation, with effects upon subsequent performance. Students' evaluation during their written reflections following Unit 1 exam predicted how extensively they discussed plans to adapt their learning process in subsequent learning cycle (Unit 2). Plans to adapt described in written reflections predicted an increase in the frequency of desirable learning behaviors during Unit 2, which in turn predicted higher Unit 2 exam scores, after controlling for previous exam performance.

Table 7
Descriptive Statistics and Correlations for Major Study Variables.

Variable	M	SD	Skew	Kurtosis	1	2	3	4	5	6
1. Exam 1 score (of 100)	63.30	20.80	-0.68	0.26	—					
2. Exam 2 score (of 100)	71.0	17.0	-1.30	3.77	0.41***	—				
3. Evaluation	0.28	0.14	2.45	10.34	-0.01	0.10	—			
4. Adaptation	0.31	0.15	1.22	2.96	-0.03	0.03	0.57***	—		
5. Desirable behaviors	0.53	0.18	-0.66	0.17	0.30**	0.29**	-0.08	0.11	—	
6. Other behaviors	0.81	0.15	-1.87	6.21	-0.11	-0.10	0.03	0.04	-0.16	—

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

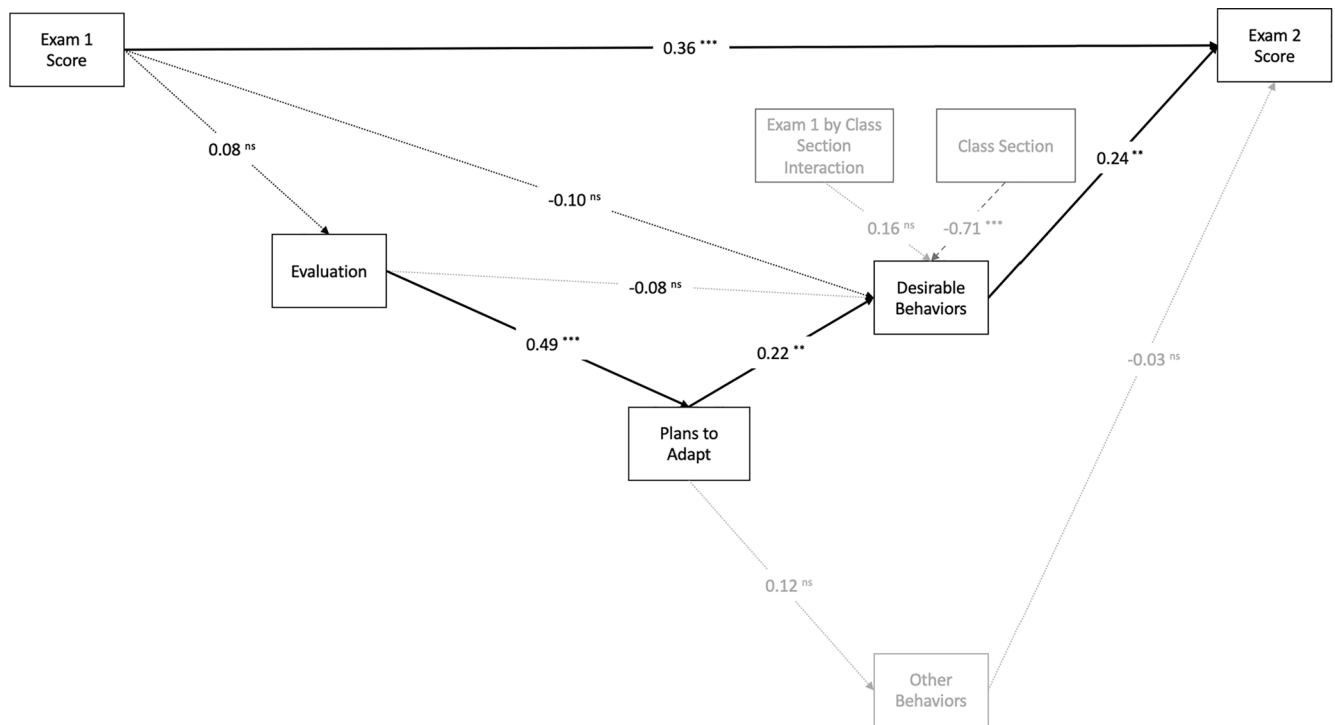


Fig. 2. The Final Path Model with Standardized Coefficients. Note. Solid lines depict positive statistically significant relations, dashed lines depict negative statistically significant relations, and dotted lines depict statistically non-significant relationships. Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ns $p > 0.05$.

Table 8
Hypothesized Direct Effects

Hypothesis	Estimate Unstandardized (SE)	Estimate Standardized (SE)	p-value
H1 (Exam 1 Score predicting Evaluation)	0.02 (0.02)	0.08 (0.09)	0.41
H2 (Evaluation predicting Plans to Adapt)	0.70 (0.11)	0.49 (0.08)	0.00
H3 (Plans to Adapt predicting Desirable Behaviors)	0.28 (0.11)	0.22 (0.08)	0.009
H4 (Desirable Behaviors predicting Exam 2 Score)	0.62 (0.23)	0.24 (0.09)	0.008

Table 9
Hypothesized Indirect Effect.

Hypothesis	Estimate Unstandardized (SE)	Estimate Standardized (SE)	p-value
H5 (Evaluation predicting Desirable Behaviors through Plans to Adapt)	0.20 (0.09)	0.11 (0.04)	0.02

5.1. Theoretical implications

The relationships observed provide empirical evidence supporting SRL as a contingent process (Ben-Eliyahu & Bernacki, 2015; Winne, 2010; Winne & Hadwin, 1998; Zimmerman, 1990) where metacognitive evaluation of learning at the end of learning cycle can inform planning and behavior changes in subsequent learning cycles, and the contingency between metacognitive evaluation and plans to adapt can predict enactment of effective learning strategies that lead to improved performance in complex learning tasks, as Binbasaran Tuysuzoglu and Greene (2015) demonstrated. Additionally, our model results included a

Table 10
Other Behaviors in a Relation to Metacognitive Evaluation, Plans to Adapt and Achievements.

Relationship	Estimate Unstandardized (SE)	Estimate Standardized (SE)	p-value
Plans to Adapt predicting Other Behaviors	0.13 (0.10)	0.12 (0.09)	0.18
Other Behaviors predicting Exam 2 Score	-0.11 (0.26)	-0.03 (0.08)	0.68
Evaluation predicting Other Behaviors through Plans to Adapt	0.09 (0.07)	0.06 (0.04)	0.17

novel finding that previous exam performance did not predict the volume of metacognitive evaluation. This finding clashes with theoretical positions that discrepancies between products of learning and learners' expectations trigger subsequent metacognitive monitoring (Winne, 2018; Winne & Azevedo, 2014; Winne & Hadwin, 1998). We speculate the reason for this could be a change in a student's perception of their own underachievement during the period between receiving their Exam 1 score and submitting responses to reflective writing assignment. For instance, the reflective writing assignment was due approximately a week after the Exam 1 scores were released. During that period, students could have compared their scores to the class mean or to their classmates' scores, and reframed their underachievement (e.g., "Even though I expected an A and earned a B, given the class mean was B-, I actually did not perform that poorly"). In a future research study, researchers could administer the reflective writing assignment immediately after exam scores were announced, to determine whether the volume of metacognitive evaluation, before comparison to peers, is related to exam score expectations. Further, we acknowledge there may be other plausible explanations for the lack of a statistically significant relationship between these two variables, including a possible restriction of range due to our focus on only those students whose exam scores were below what they expected.

5.2. Practical implications

Our findings add to the growing body of research on the role of adaptive SRL processes in undergraduate STEM courses, e.g., [Greene et al. \(2019\)](#). Specifically, these findings suggest that prompting and guiding students to compose accurate accounts of prior performance and provide specific plans for future studying are critical to helping students who underachieved make a shift towards more productive SRL behaviors in the next learning cycles. Further, the analytical procedure we utilized in this study could be a powerful and scalable way to help students in need of support early in semester. The reflective writing prompts seemed to prompt useful reflection at the end of learning cycle. In addition, our computational approach to automatically measure metacognition in reflective writing may provide instructors with an efficient means of computing the volume of metacognitive evaluation and plans to adapt in reflective essays, allowing for intervention in a timely manner (e.g., by suggesting that students who performed worse than they expected revise their reflections and more deeply elaborate on their plans to adapt). Moreover, by customizing the dictionary of terms utilized in our algorithm, instructors may be able to adjust the algorithm to a specific course.

6. Limitations and recommendations for future research

We noted the following major limitations in this study that may be addressed in future research. First, our sample of students who performed worse than they expected came from a single course in a single context. This study should be replicated in similar and different contexts. Next, we acknowledge that asking students to respond to metacognitive prompts may increase the likelihood of students' engagement in evaluation and planning. Furthermore, even though reflective responses were reasonably short, differences in writing style and language ability among students may challenge some of the assumptions embedded in the NLP algorithm. For instance, instead of using sequential rhetorical connectives (e.g., *first*, *next*, *finally*...) to describe their plans to adapt, some students tended to provide a bulleted list. It is also possible that more verbose writers would create responses with higher scores. Moreover, some students may have included evaluative statements in their responses to adaptation prompts and/or plans in their responses to evaluative prompts, regardless instructions they received for each prompt. Such variability in writing and language could affect the computational accuracy and researchers should consider developing ways to identify and reclassify such language into the proper category. Last, our inferences regarding how students used digital objects in the course LMS and how such use approximated their SRL behaviors may not necessarily be correct. For example, we classified *posing a question in the discussion forum* as a help seeking behavior, however, we acknowledge it is possible that a student could ask a question unrelated to the course, which would not be help seeking. Also, we inferred that a student who accessed the assignment instructions page multiple times throughout the semester was more self-regulated than their peer who accessed it less, however, we are aware it is possible that self-regulated learners would keep track of the assignment instructions and not need to access them multiple times. The traced behaviors we observed, though mostly of an expected sign relative to theoretical and instructional expectations, weakly correlated with Exam 2 scores and were mostly statistically non-significant. This is not surprising as individual trace behaviors may not correlate with outcomes on their own until aggregated with similar trace behaviors, given that idiosyncratic differences across participants exist (e.g., one participant may review powerpoint

slides frequently, another may review class notes frequently, resulting in neither trace being correlated on its own, but their aggregate being correlated). As such, inferences about these individual trace data, and what they indicate, should be validated in future studies. Overall, more research is needed to validate inferences made from digital traces, e.g., by interviewing students or by collecting verbal protocol data ([Greene, Copeland, Deekens, & Freed, 2018](#)) concurrently with digital trace data, to validate the latter with the former.

7. Conclusion

Increasingly, STEM instructors are utilizing active learning pedagogies to help students succeed in their courses and complete their degrees. To take advantage of the powerful affordances of active learning pedagogies, students need self-regulatory knowledge and skills. In particular, to recover from unexpectedly poor performance early in the semester, students need to engage in effective metacognitive evaluation of their learning and make effective adaptations to their behaviors in a subsequent learning cycle. We drew upon [Winne and Hadwin's \(1998\)](#) four-stage model of SRL to investigate metacognitive monitoring and control processes during adaptation, an under-researched stage of SRL. We applied natural language processing to identify language reflecting evaluation of prior studying and plans to adapt learning as stated in a reflective essay after a first unit exam and utilized digital traces of learning behaviors to measure students' adaptation from their prior use of learning resources. Path model results largely supported our hypotheses. The substantive and positive relationship detected between volume of metacognitive evaluation of prior learning and volume of plans to adapt is consistent with previous theoretical and empirical work. Metacognitive evaluation and plans to adapt can predict an increase in desirable learning behaviors and further relate to improved performance in complex learning tasks, such as those in an introductory STEM college course based on active learning pedagogies. Our findings indicate that composing accurate accounts of prior and specific plans for future studying may help underachieving students make a shift towards more productive SRL behaviors in subsequent learning cycles and increase their chances for success in an active learning STEM course. We also note that, because SRL processes are nuanced and the sufficiency of a self-regulated learning process such as a metacognitive control strategy is dependent on (1) the metacognitive monitoring process that precipitates it (e.g. a negative judgment of learning about one's mastery of a concept), (2) the accuracy of this judgment, and (3) the appropriateness of the strategy selected to address it (e.g., "rehearse the concept using self-explanation" is likely to be superior to "keep re-reading about the concept"; [Winne & Hadwin, 1998](#)), additional research might focus at this more granular view of such processes. This would require thorough instrumentation of learning environments, and include both digital tracing of events reflecting strategy choices, as well as self-report prompts to clarify what precipitated them (e.g., [Salehian Kia et al., 2021](#)). These approaches will also require innovations in research designs, including the use of mixed methods to derive valid (per [Winne, 2020](#)) and contextualized observation of SRL processes at a fine grain size in authentic contexts ([Ben-Eliyahu & Bernacki, 2015](#)).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

LMS behaviors mapped to SRL processes

SRL Process Enacted	LMS Behavior
course information gathering	Viewing calendar events Loading of a Lessons page Read announcement in Sakai Viewing page in Sakai where syllabus is hosted Downloading the syllabus from syllabus page or resources Course documents like group assignments and accessing MasteringBio Downloading the GRQ word doc View assignment instructions Download study guide Download example exam from previous semesters Downloading the reflection prompt file Embedded tool for viewing course reserves
diligence keeping up with course information	Class outlines are questions intended to be answered during class
task knowledge acquisition	Editing calendar events Attending a session of class as recorded in Learning Catalytics Slides presented during lecture that are posted after class Supplemental readings provided for some lessons Download example exam key from previous semesters
Resources knowledge acquisition	Starting the assessment Submitting assessment answers Starting a new submission Starting a submission of the reflection assignment Save a pending submission Finalize and submit a response Submit a correct answer to an item within MasteringBio coursework Submit an incorrect answer to an item within MasteringBio coursework Provided the solution of an item during a MasteringBio quiz Submit a correct answer to an item during a MasteringBio quiz Submit an incorrect answer to an item during a MasteringBio quiz Submitting self-reflection answers
goals	Resuming a previously started assessment Resuming a previously started self-reflection submission
time and effort planning	Starting the practice exam Resuming a previously started practice exam Submitting practice exam answers
content knowledge acquisition	Clicking the link to Learning Center page Use of the discussion forum to pose a question Requesting a time slot during office hours Use hint within MasteringBio coursework View an item within MasteringBio coursework Provided the solution of an item within MasteringBio coursework Course specific review session provided by the Learning Center Review sessions provided by bio department TAs and Peer Mentors Attending an office hours appointment with the instructor Attending a one-on-one peer mentoring session with bio department tutors Attending one-on-one academic coaching at the learning center Attending one-on-one writing coaching at the learning center
doing assigned work	Revise an existing pending submission Review results of a submitted assessment View the gradebook page in Sakai Review results of a submitted practice exam
submitting assigned work	Review a previously submitted response Download current semester exam file Download key to current semester exam
completing started work	
self-testing	
help seeking	
revising	
metacognitive monitoring	
monitoring for accuracy	
reflection on past performance	

References

- Allan, E. G., & Driscoll, D. L. (2014). The three-fold benefit of reflective writing: Improving program assessment, student learning, and faculty professional development. *Assessing Writing*, 21, 37–55. <https://doi.org/10.1016/j.asw.2014.03.001>
- Ariel, R., Dunlosky, J., & Bailey, H. (2009). Agenda-based regulation of study-time allocation: When agendas override item-based monitoring. *Journal of Experimental Psychology: General*, 138(3), 432. <https://doi.org/10.1037/e566962012-072>
- Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of educational psychology*, 96(3), 523–535. <https://doi.org/10.1037/0022-0663.96.3.523>
- Ben-Eliyahu, A., & Bernacki, M. L. (2015). Addressing complexities in self-regulated learning: A focus on contextual factors, contingencies, and dynamic relations. *Metacognition and Learning*, Article 10. <https://doi.org/10.1007/s11409-015-9134-6>
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Erlbaum Associates.
- Bonwell, C. C., & Eison, J. A. (1991). *Active Learning: Creating Excitement in the Classroom*. 1991 ASHE-ERIC Higher Education Reports. ERIC Clearinghouse on Higher Education, The George Washington University, One Dupont Circle, Suite 630, Washington, DC 20036-1183.
- Bernacki, M. L. (2018). Examining the cyclical, loosely sequenced, and contingent features of self-regulated learning: Trace data and their analysis. *Handbook of self-regulation of learning and performance*. Routledge/Taylor & Francis Group.
- Bernacki, M. L., Vosicka, L., & Utz, J. C. (2019). Can a brief, digital skill training intervention help undergraduates “learn to learn” and improve their STEM achievement? *Journal of Educational Psychology*, 112(4). <https://doi.org/10.1037/edu0000405>

- Binbasaran Tuysuzoglu, B., & Greene, J. A. (2015). An investigation of the role of contingent metacognitive behavior in self-regulated learning. *Metacognition and Learning*, 10(1). <https://doi.org/10.1007/s11409-014-9126-y>
- Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive psychology*, 3(2), 193–209. [https://doi.org/10.1016/0010-0285\(72\)90003-5](https://doi.org/10.1016/0010-0285(72)90003-5)
- Braxton, J. M., Milem, J. F., & Sullivan, A. S. (2000). The Influence of Active Learning on the College Student Departure Process. *The Journal of Higher Education*, 71(5), 569–590. <https://doi.org/10.1080/00221546.2000.11778853>
- Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, 27, 1–13. <https://doi.org/10.1016/j.iheduc.2015.04.007>
- Bruinsma, M. (2004). Motivation, cognitive processing and achievement in higher education. *Learning and instruction*, 14(6), 549–568. <https://doi.org/10.1016/j.learninstruc.2004.09.001>
- Carlsen, C. (2010). Discourse connectives across CEFR-levels: A corpus based study. *Communicative Proficiency and Linguistic Development: Intersections between SLA and language testing research*, 191–210.
- Celce-Murcia, M., Larsen-Freeman, D., & Williams, H. A. (1983). *The grammar book: An ESL/EFL teacher's course* (p. 299). Newbury House.
- Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in cognitive science*, 1(1), 73–105. <https://doi.org/10.1111/j.1756-8765.2008.01005.x>
- Cook, E., Kennedy, E., & McGuire, S. Y. (2013). Effect of teaching metacognitive learning strategies on performance in general chemistry courses. *Journal of Chemical Education*, 90(8), 961–967. <https://doi.org/10.1021/ed300686h>
- Cui, Y., Wise, A. F., & Allen, K. L. (2019). Developing reflection analytics for health professions education: A multi-dimensional framework to align critical concepts with data features. *Computers in Human Behavior*, 100, 305–324. <https://doi.org/10.1016/j.chb.2019.02.019>
- Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, 28(3), 425–474. <https://doi.org/10.1007/s10648-015-9320-8>
- Eklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, 13(4), 277–287. <https://doi.org/10.1027/1016-9040.13.4.277>
- Eilam, B., & Aharon, I. (2003). Students' planning in the process of self-regulated learning. *Contemporary Educational Psychology*, 28(3), 304–334. [https://doi.org/10.1016/S0361-476X\(02\)00042-5](https://doi.org/10.1016/S0361-476X(02)00042-5)
- Flanders, G. R. (2017). The effect of gateway course completion on freshman college student retention. *Journal of College Student Retention: Research, Theory & Practice*, 19(1), 2–24. <https://doi.org/10.1177/1521025115611396>
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410–8415. <https://doi.org/10.1073/pnas.1319030111>
- Gibson, A., Kitto, K., & Bruza, P. (2016). Towards the discovery of learner metacognition from reflective writing. *Journal of Learning Analytics*, 3(2), 22–36. <https://doi.org/10.18608/jla.2016.3.2>
- Greene, J. A., & Azevedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research*, 77(3). <https://doi.org/10.3102/003465430303953>
- Greene, J. A., & Azevedo, R. (2009). A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. *Contemporary Educational Psychology*, 34(1). <https://doi.org/10.1016/j.cedpsych.2008.05.006>
- Greene, J. A., Copeland, D. Z., Deekens, V. M., & Freed, R. (2018). Self-regulated learning processes and multiple source use in and out of school. *Handbook of multiple source use*. Routledge.
- Greene, J. A., Muis, K. R., & Pieschl, S. (2010). The role of epistemic beliefs in students' self-regulated learning with computer-based learning environments: Conceptual and methodological issues. *Educational Psychologist*, 45(4). <https://doi.org/10.1080/00461520.2010.515932>
- Greene, J. A., Plumley, R. D., Urban, C. J., Bernacki, M. L., Gates, K. M., Hogan, K. A., ... Panter, T. A. (2019). Modeling temporal self-regulatory processing in a higher education biology course. *Learning and Instruction*, 72, Article 101201. <https://doi.org/10.1016/j.learninstruc.2019.04.002>
- Greene, J. A., & Schunk, D. H. (2017). Historical, contemporary, and future perspectives on self-regulated learning and performance. *Handbook of self-regulation of learning and performance*. Routledge.
- Hadwin, A. F., & Webster, E. A. (2013). Calibration in goal setting: Examining the nature of judgments of confidence. *Learning and Instruction*, 24, 37–47. <https://doi.org/10.1016/j.learninstruc.2012.10.001>
- Hamm, J. M., Perry, R. P., Chipperfield, J. G., Hladkyj, S., Parker, P. C., & Weiner, B. (2020). Reframing achievement setbacks: A motivation intervention to improve 8-year graduation rates for students in science, technology, engineering, and mathematics (STEM) fields. *Psychological Science*, 31(6), 623–633. <https://doi.org/10.1177/0956797620904451>
- Hattie, J. A. C., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. *npj Science of Learning*, 1, 16013. <https://doi.org/10.1038/npjscilearn.2016.13>
- Livingston, J. A. Metacognition: An Overview. *Psychology*, 13, 259–266.
- Loper, E., & Bird, S. (2002). NLTK: the natural language toolkit. *arXiv preprint cs/0205028*. 10.3115/1118108.1118117.
- Hong, W., Bernacki, M. L., & Perera, H. N. (2020). A latent profile analysis of undergraduates' achievement motivations and metacognitive behaviors, and their relations to achievement in science. *Journal of Educational Psychology*, 112(7). <https://doi.org/10.1037/edu0000445>
- Lombardi, D., Shipley, T. F., Astronomy Team, Biology Team, Chemistry Team, Engineering Team, ... Physics Team. (2021). The curious construct of active learning. *Psychological Science in the Public Interest*, 22(1). <https://doi.org/10.1177/1529100620973974>
- Luo, W., & Litman, D. (2016). Determining the quality of a student reflective response. *The Twenty-Ninth International FLAIRS Conference*.
- Medina, M. S., Castleberry, A. N., & Persky, A. M. (2017). Strategies for improving learner metacognition in health professional education. *American Journal of Pharmaceutical Education*, 81(4), 78. <https://doi.org/10.5688/ajpe81478>
- Miller, C. J., & Bernacki, M. L. (2019). Training preparatory mathematics students to be high ability self-regulators: Comparative and case-study analyses of impact on learning behavior and achievement. *High Ability Studies*, 30(1–2). <https://doi.org/10.1080/13598139.2019.1568829>
- Moos, D. C., & Azevedo, R. (2008). Monitoring, planning, and self-efficacy during learning with hypermedia: The impact of conceptual scaffolds. *Computers in Human Behavior*, 24(4), 1686–1706. <https://doi.org/10.1016/j.chb.2007.07.001>
- O'Loughlin, V. D., & Griffith, L. M. (2020). Developing Student Metacognition through Reflective Writing in an Upper Level Undergraduate Anatomy Course. *Anatomical Sciences Education*, 13(6), 680–693. <https://doi.org/10.1002/ase.v13.610.1002/ase.1945>
- Olson, S., & Riordan, D. G. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Executive Office of the President: Report to the President.
- Pardo, A., Han, F., & Ellis, R. A. (2017). Combining university student self-regulated learning indicators and engagement with online learning events to predict academic performance. *IEEE Transactions on Learning Technologies*, 10(1), 82–92. <https://doi.org/10.1109/TLT.462007610.1109/TLT.2016.2639508>
- Perez, T., Cromley, J., & Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *Journal of Educational Psychology*, 106(1). <https://doi.org/10.1037/a0034027>
- Pieschl, S., Bromme, R., Porsch, T., & Stahl, E. (2009). Does adaptation to task complexity guarantee success in learning? processes and conditions of beneficial adaptations in self-regulated learning. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2638–2643).
- Prince, M. (2004). Does active learning work? A review of the research. *Journal of engineering education*, 93(3), 223–231. <https://doi.org/10.1002/j.2168-9830.2004.tb00809.x>
- Raković, M., Winne, P. H., Marzouk, Z., & Chang, D. (2021). Automatic identification of knowledge transforming content in argument essays developed from multiple sources. *Journal of Computer Assisted Learning*, 37(4). <https://doi.org/10.1111/jcal.12531>
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of statistical software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Salehian Kia, F., Hatala, M., Baker, R. S., & Teasley, S. D. (2021). Measuring Students' Self-Regulatory Phases in LMS with Behavior and Real-Time Self Report. In *LAK21: 11th International Learning Analytics and Knowledge Conference* (pp. 259–268).
- Satorra, A. (1992). Asymptotic robust inferences in the analysis of mean and covariance structures. *Sociological methodology*, 22, 249. <https://doi.org/10.2307/270998>
- Scardamalia, M., Bereiter, C., & Steinbach, R. (1984). Teachability of reflective processes in written composition. *Cognitive science*, 8(2), 173–190. <https://doi.org/10.1207/s15516709cog0802.4>
- Shekhar, P., Borrego, M., DeMonbrun, M., Finelli, C., Crockett, C., & Nguyen, K. (2020). Negative Student Response to Active Learning in STEM Classrooms: A Systematic Review of Underlying Reasons. *Journal of College Science Teaching*, 49(6).
- Tanner, K. D. (2012). Promoting student metacognition. *CBE—Life Sciences Education*, 11(2), 113–120. <https://doi.org/10.1187/cbe.12-03-0033>
- Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., ... Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, 117(12), 6476–6483. <https://doi.org/10.1073/pnas.1916903117>
- Tinto, V. (1993). Leaving college: Rethinking the causes and cures of student attrition (2nd ed.). University of Chicago Press. 10.7208/chicago/9780226922461.001.0001.
- Ullmann, T. D. (2017). March). Reflective writing analytics: Empirically determined keywords of written reflection. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 163–167). 10.1145/3027385.3027394.
- Virtanen, P., Niemi, H., & Nevgi, A. (2017). Active learning and self-regulation enhance student teachers' professional competences. *Australian Journal of Teacher Education*, 1–20. <https://doi.org/10.14221/ajte.10.14221/ajte/vol42/iss12/10.14221/ajte.2017v42n12.1>
- Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational psychologist*, 45(4), 267–276. <https://doi.org/10.1080/00461520.2010.517150>
- Winne, P. H. (2014). Issues in researching self-regulated learning as patterns of events. *Metacognition and Learning*, 9(2), 229–237. <https://doi.org/10.1007/s11409-014-9113-3>
- Winne, P. H. (2018). *Cognition and metacognition within self-regulated learning*. In *Educational psychology handbook series. Handbook of self-regulation of learning and performance*, 10.4324/9781315697048-3 (pp. 36–48). Routledge/Taylor & Francis Group.

- Winne, P. H. (2020). Construct and consequential validity for learning analytics based on trace data. *Computers in Human Behavior*, 112, 106457. <https://doi.org/10.1016/j.chb.2020.106457>
- Winne, P. H., & Azevedo, R. (2014). Metacognition. *The Cambridge handbook of the learning sciences* (pp. 63–87). Cambridge University Press. <https://doi.org/10.1017/CBO9781139519526.006>
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Erlbaum.
- Wolters, C. A., Won, S., & Hussain, M. (2017). Examining the relations of time management and procrastination within a model of self-regulated learning. *Metacognition and learning*, 12(3), 381–399. <https://doi.org/10.1037/0022-0663.95.1.179>
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25(1), 3–17. https://doi.org/10.1207/s15326985ep2501_2
- Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist*, 48(3), 135–147. <https://doi.org/10.1080/00461520.2013.794676>