Contemporary Educational Psychology 69 (2022) 102064

Contents lists available at ScienceDirect

Contemporary Educational Psychology

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

The evolution of metacognitive strategy use in an open-ended learning environment: Do prior domain knowledge and motivation play a role?

Yingbin Zhang [a,*], Luc Paquette [a], Nigel Bosch [b], Jaclyn Ocumpaugh [c], Gautam Biswas [d], Stephen Hutt [e], Ryan S. Baker [e]

[a] *Department of Curriculum & Instruction, University of Illinois at Urbana-Champaign, 1310 S Sixth Street, Champaign, IL 61820, United States*
[b] *School of Information Sciences and Department of Educational Psychology, University of Illinois at Urbana-Champaign, 501 E. Daniel St., Champaign, IL 61820, United States*
[c] *Penn Center for Learning Analytics, University of Pennsylvania, 3700 Walnut St., Philadelphia, PA 19104, United States*
[d] *Department of Computer Science/ISIS, Vanderbilt University, 1025 16th Ave South, Nashville, TN 37212, United States*
[e] *Graduate School of Education, University of Pennsylvania, 3700 Walnut St., Philadelphia, PA 19104, United States*

ARTICLE INFO

Keywords:
Metacognitive strategy
Self-regulated learning
Prior domain knowledge
Task value
Self-efficacy

ABSTRACT

There is a growing interest in viewing self-regulated learning as events unfolding over time, especially when students perform learning tasks in computer-based environments. Metacognitive activities are critical events in self-regulated learning. This study investigated the evolution of metacognitive strategy use in an open-ended computer-based learning environment, Betty's Brain. The data were from 93 sixth graders who used Betty's Brain to learn about climate change for four days. We extracted indicators of metacognitive strategy use from action logs. A knowledge test and self-report questionnaire were administrated before students started using Betty's Brain to assess prior domain knowledge and motivation, respectively. Results showed that metacognitive strategy use increased from the first to the second day and remained stable from the second to the fourth day of the study. The evolution of these behaviors varied across students. Task value and prior domain knowledge partially explained the individual differences in this evolution. Task value and prior domain knowledge also predicted the use of metacognitive strategies. Self-efficacy did not influence metacognitive strategy use. These results suggest the need for further investigation into the role of motivation and prior domain knowledge in the temporal evolution of metacognitive events.

1. Introduction

Open-ended, computer-based learning environments provide students with the opportunity to experience complex phenomena in authentic problem-solving scenarios and the freedom to learn by making their own decisions (Land, 2000; Lowyck, 2014). However, the freedom and complexity demand that learners actively monitor and manage their activities (Kinnebrew, Segedy, & Biswas, 2017; Segedy, Kinnebrew, & Biswas, 2015a). That is, learners need to self-regulate learning (SRL; Zimmerman, 1990). Metacognitive strategies are an essential component of SRL (Panadero, 2017), and their use can facilitate learning (Ohtani & Hisasaka, 2018). Studies have found that motivational factors and prior domain knowledge influence how learners use cognitive and metacognitive strategies (Liem et al., 2008; Ocak & Yamaç, 2013; Üner et al., 2020), but few studies have explored how these factors relate to changes in strategic behaviors over time.

Learning has temporal characteristics because it is the acquisition process of knowledge and skills (Molenaar, 2014). It takes time for these new skills to manifest in behavioral changes (Soderstrom & Bjork, 2015; Zimmerman, 2002). There has been increasing interest in understanding the temporal aspects of SRL (Azevedo, 2014; Molenaar & Järvelä, 2014; Winne & Baker, 2013), as they have both theoretical and practical implications. Understanding how strategic learning behaviors change over time—and how prior domain knowledge and motivation influence such change—can enrich SRL theories and lead to more informed decisions about when to provide scaffolding and to whom. As such, the present study investigated how metacognitive strategies use evolved over time in an open-ended computer-based learning environment, Betty's Brain. Specifically, we examined how prior knowledge, task value, and self-efficacy related to the temporal evolution of metacognitive strategies.

* Corresponding author at: University of Illinois at Urbana-Champaign, 1310 S Sixth Street, Champaign, IL 61820, United States.
E-mail address: yingbin2@illinois.edu (Y. Zhang).

https://doi.org/10.1016/j.cedpsych.2022.102064

Available online 30 March 2022

1.1. Self-regulated learning (SRL) and its temporal characteristics

SRL refers to how learners adaptively regulate their cognition, behaviors, motivation, and affect to achieve their learning goals (Schunk & Greene, 2017). SRL can be viewed as a series of cognitive and metacognitive events (Azevedo, Moos, et al., 2010) unfolding over time (Greene & Azevedo, 2010; Hadwin, 2021; Molenaar & Järvelä, 2014; Winne & Baker, 2013). Examples of cognitive events include memorization, elaboration, and reviewing, while examples of metacognitive events include goal setting, planning, self-monitoring, self-control, and self-evaluation (Dent & Koenka, 2016).

These broad categorizations are important, but SRL events, like learning, are highly contextually dependent (Azevedo et al., 2012; Li et al., 2020). An SRL event should not be classified into a specific category (cognitive vs. metacognitive) or subcategory (e.g., memorization and elaboration) without considering the specific context in which it occurs. For instance, in MetaTutor, an intelligent hypermedia environment for learning human body systems (e.g., the circulatory system; Azevedo, Johnson, et al., 2010), the action of taking notes may align with different subcategories of cognitive events (Azevedo et al., 2013). If students create notes by copying text verbatim from the learning material, taking notes may represent content reproduction (a shallow cognitive processing event). By contrast, if students create notes by integrating ideas from different sections, the note-taking action may represent elaboration (a deep cognitive processing event).

Researchers have investigated SRL events across two dimensions: (1) individual and (2) sequential characteristics (Knight et al., 2017; Molenaar, 2014). Individual characteristics may include the occurrence rate, timing, and duration of an event during learning processes. Sequential characteristics may include the sequential relations and transitions among events (Molenaar & Järvelä, 2014)—for example, the conditional probability of reading relevant book pages after taking a quiz. The sequential characteristics of SRL events can be discovered via data mining techniques, such as sequential pattern mining (Kinnebrew et al., 2014), process mining (Bannert et al., 2014), and epistemic network analysis (Paquette et al., 2021), or advanced statistical models, such as lag-sequential analyses (Kuvalja et al., 2014) and statistical discourse analyses (Molenaar & Chiu, 2014).

This paper focuses on the temporal change of individual characteristics, specifically, the rate at which students use metacognitive strategies and how the rate evolves during the learning process. We use the term *evolution* to refer to the temporal change of metacognitive strategy use based on the following considerations. Prior studies have used different terms such as evolution (de Backer et al., 2016; Kinnebrew et al., 2014), temporal variation (Paans et al., 2019; Zheng et al., 2019), and difference (Greene et al., 2021) to describe the change of SRL behaviors over time. Compared with the other terms, evolution implies that the behavioral change is systematic rather than random as the word evolution can refer to "a process of change in a certain direction" (Merriam-Webster, n.d.). Behavioral changes may demonstrate a learner's ability to adapt as they acquire new knowledge (Greene et al., 2021). However, this notion may only apply to systematic behavioral change as a random change does not require adaptivity. In summary, using evolution may match the SRL process better than other terms such as temporal variations and differences.

Studies have examined the evolution of SRL behaviors in various contexts, including classrooms without technology (de Backer et al., 2016), sheltered Internet-based learning environments (Paans et al., 2019), online collaborative inquiry environments (Zheng et al., 2019), and learning management systems (Greene et al., 2021). This research has found differences in the evolution of SRL behaviors between high and low-performing groups (Paans et al., 2019; Zheng et al., 2019), suggesting that finding ways to support SRL behaviors may be one path to enhancing learning. However, these studies did not investigate the relationship between how students' other characteristics related to the evolution of their SRL behaviors, which limits the inferences we can make about which SRL scaffolds were most effective to whom. In contrast, the current research investigates whether students' characteristics like prior domain knowledge, task value, and self-efficacy, explain the individual differences in the evolution of SRL behaviors within an open-ended learning environment.

1.2. Self-regulated learning and Open-ended learning environment

Open-ended learning refers to situations in which learners determine what, how, and when to learn based on their unique intentions and external goals (Hannafin et al., 1994). It contrasts with directed learning, where the environment and designers determine what is to be learned and the sequence of action. Open-ended learning environments provide learners with authentic contexts and rich resources to support the exploration of complex phenomena, the integration of new knowledge and daily experience, and learner-centered inquiries (Land, 2000). However, these environments demand that individuals actively monitor their understanding, evaluate their performance, and refine their strategies. As such, SRL is particularly critical in open-ended learning environments (Segedy et al., 2015a). Indeed, low-performing learners show fewer SRL activities and approach the task with less effective strategies (Kinnebrew et al., 2013; Roscoe et al., 2013; Sabourin et al., 2013). Nevertheless, how SRL strategy use evolves in such environments is underexplored. Kinnebrew et al. (2014) showed that students who received different scaffolding differed in the evolution of their strategic and ineffective behaviors. Segedy et al. (2015b) found that students' problem-solving strategies were relatively stable across days. The temporal characteristics of SRL events require further research in open-ended learning environments.

1.3. Self-regulated learning and prior domain knowledge

The information processing theory of SRL emphasizes the role of prior domain knowledge in learning (Winne & Hadwin, 2008). According to this model, domain knowledge is a cognitive condition that impacts how learners understand a task. Taub et al. (2014) further illustrate how domain knowledge may influence the cognitive and metacognitive activities in the four phases of SRL, i.e., task definition, goal setting and plan, execution, and adaptation. Indeed, studies have found a positive association between domain knowledge and SRL strategies (Li, 2019; Moos & Azevedo, 2008, 2009a; Taub et al., 2014; Taub & Azevedo, 2019). Moos and Azevedo (2008) as well as Taub and Azevedo (2019) have investigated the relationship between domain knowledge and SRL in undergraduates' learning about the human circulatory system. In Moos and Azevedo's (2008) research, learners with higher domain knowledge planned and monitored their learning more frequently than those with low domain knowledge. Learners with higher domain knowledge also engaged in more cognitive activities, such as note-taking, summarizing, and memorizing. Taub and Azevedo (2019) replicated these results, reinforcing the link between domain knowledge and SRL.

The increased use of SRL strategies by students with high prior knowledge is possibly related to working memory (Moos & Azevedo, 2008; Taub et al., 2014). Working memory has serious limitations in both capacity and duration when learners encounter novel information (Sweller, 2011). Learners with low domain knowledge need more working memory capacity for processing the information, while learners with high domain knowledge can allocate more of this capacity for regulation (Taub et al., 2014). As the learning process unfolds, domain knowledge increases, and some of the working memory capacity for novel information may be released. As these resources become available, learners engage in more regulatory activities.

In this account, learners who frequently use SRL strategies as part of their normal learning practices may be less likely to increase SRL strategy use. Similarly, increases in domain knowledge may vary, especially between novice and more expert learners. Thus, domain knowledge may influence both the overall SRL strategy use and the evolution of SRL strategy use. However, this assumption has not been examined.

1.4. Self-regulated learning and motivation

Motivation has been an essential component of several SRL models (Efklides, 2011; Pintrich, 2000; Winne & Hadwin, 2008; Zimmerman, 2000). Monitoring learning processes and enacting proper strategies consumes energy, time, and effort (Zimmerman, 2000). If learners are not motivated, they tend not to apply proper strategies or persist with the learning activity. Motivation may serve as a predictor, mediator, or outcome of self-regulation activities (Zimmerman & Schunk, 2008). The present study investigates two prominent sources of motivation: self-efficacy and task value. These motivational components are mutually correlated, but each has unique effects on SRL behaviors (Pintrich & de Groot, 1990; Üner et al., 2020; Zimmerman & Schunk, 2008).

1.4.1. Self-efficacy

Self-efficacy refers to an individual's beliefs about their capability to complete a task (Bandura, 1997). The social cognitive theory of SRL emphasizes the influential role of self-efficacy in SRL (Schunk & Ertmer, 2000; Zimmerman, 2000). Learners with higher self-efficacy reported more cognitive and metacognitive strategy use (Pintrich & de Groot, 1990). High self-efficacy learners tended to set challenging goals for themselves and persist when facing difficulty (Schunk & Ertmer, 2000; Zimmerman & Bandura, 1994). Similarly, high self-efficacy learners monitored working time more frequently than their low self-efficacy counterparts (Bouffard-Bouchard et al., 1991).

Research has used learning process data to investigate the association between self-efficacy and SRL. For instance, Moos and colleagues (Moos, 2014; Moos & Azevedo, 2009a) coded students' think-aloud data for metacognitive monitoring processes based on a well-developed SRL coding scheme (Azevedo & Cromley, 2004). The results showed that self-efficacy positively predicted the behaviors of monitoring understanding and progress toward goals. Hong et al. (2020) examined undergraduates' motivation and metacognition in a biology course. They extracted metacognitive behavior metrics from students' action logs in the Blackboard learning management system. Compared with groups reporting stronger performance-approach or performance-avoidance goals and psychological cost, the group characterized by higher self-efficacy, mastery-approach goal, and task value engaged more in activities related to planning their learning and monitoring their performance. Overall, the empirical evidence supports the claim that self-efficacy is positively related to SRL strategy use.

1.4.2. Task value

Task value refers to a student's perceptions about the importance, usefulness, enjoyment, and cost of a task (Wigfield & Eccles, 1992). Eccles and colleagues' expectancy-value model of achievement behavior claims that task-value beliefs, in addition to expectancy-related and ability-related beliefs, are critical determinants of task behaviors (Wigfield & Eccles, 1992; Eccles et al., 1983). If students perceive a high probability of success in a task but do not value it, they may not choose to engage in or put little effort into the task. Social cognitive models of SRL highlight the role of task value in SRL, especially the phase of goal setting and plan (Pintrich & Zusho, 2002; Zimmerman, 2000). If students highly value the task, they will spend more time on making and executing the plan.

Studies have found associations between task value and metacognitive strategy use. For instance, in Pintrich and de Groot's (1990) study, seventh-grade students' who perceived course work as important and interesting reported more use of SRL strategies, such as planning and comprehension monitoring. Task value has predicted cognitive and metacognitive strategy use in samples from adolescents and college students (Üner et al., 2020; Wolters & Pintrich, 1998), and the predicting effects were stable across the subjects of mathematics, English, and social studies (Wolters & Pintrich, 1998). Task value might indirectly impact the use of surface and in-depth learning strategies (e.g., memorization and questioning learning material) via achievement goals (Liem et al., 2008). It should be noted that measures of strategy use in these studies are self-reported. As Wigfield et al. (2008) suggest, using behavioral measures of SRL provides crucial supplementary information about associations between task value and SRL. Yet few studies have done so (Hong et al., 2020; Sabourin et al., 2013). As a result, it is unclear whether task value is also related to the temporal evolution of SRL strategy use.

1.5. The present study

The current study investigated how students' metacognitive strategy use changed from day to day in Betty's Brain, an open-ended learning environment, and examined the relationships that domain knowledge, task value, and self-efficacy had with metacognitive strategy use. Four research questions (RQs), based on the review of the literature, were investigated:

RQ 1: Does the use of metacognitive strategies increase across days?

Prior studies have found that, depending on the context, the frequency of metacognitive strategy use may decrease, increase, or remain stable over time (de Backer et al., 2016; Greene et al., 2021; Paans et al., 2019). In the current study, we expect the frequency would increase because students might become more familiar with Betty's Brain, and their knowledge of the studied topic may increase over time. Such increases in the knowledge about the environment and the domain might enhance students' capacity for applying metacognitive strategies.

RQ 2: If so, does the temporal evolution of the use of metacognitive strategies vary across students?

The temporal change of SRL behaviors has been found to differ between groups, such as students with high and low performance (Paans et al., 2019) and students receiving different scaffolding (Kinnebrew et al., 2014). Although this study did not compare specific groups, it is reasonable to assume that the evolution of metacognitive strategy use would vary across students due to individual differences.

RQ 3: Does students' prior domain knowledge predict their use of metacognitive strategies?

RQ 3.1: Is prior domain knowledge related to overall metacognitive strategy use?

RQ 3.2: Does prior domain knowledge explain the differences in the evolution of metacognitive strategy use across students?

For RQ 3.1, we expect that prior domain knowledge would be positively related to overall metacognitive strategy use, which is in line with previous studies (Li, 2019; Moos & Azevedo, 2008; Taub et al., 2014). For RQ 3.2, students who frequently apply metacognitive strategies as part of normal practices might be less likely to increase their strategy use. Besides, increases in domain knowledge may vary between novice and more experienced students. Thus, we expect that prior domain knowledge would explain the differences in the evolution of strategy use. However, there is currently a lack of theory or empirical evidence to guide the direction of the hypothesis, i.e., whether prior domain knowledge would be positively or negatively related to the increases in strategy use.

RQ 4: Do motivational factors (i.e., self-efficacy and task value) predict metacognitive strategy use?

RQ 4.1: Are motivational factors related to overall metacognitive strategy use?

RQ 4.2: Do motivational factors explain the differences in the evolution of metacognitive strategy use across students?

For RQ 4.1, prior studies have found that self-efficacy and task value are positively related to metacognitive strategy use (Moos, 2014; Moos & Azevedo, 2009a; Üner et al., 2020). Thus, we expect such associations in the current research. For RQ 4.2, students with higher task value and self-efficacy might engage in the task and adjust to the environment better than those with low task value and self-efficacy. However, their high motivation might not be sustained over multiple days, and decreases in motivation might cause smaller increases in metacognitive strategy use. Thus, we expect that task value and self-efficacy would explain some differences in the evolution of strategy use. Again, theory and empirical evidence are currently insufficient to guide the direction of the hypothesis.

2. Methods

2.1. Participants and procedure

The data were collected from 93 sixth-grade students in an urban middle school (grades 5–8) in the Southern U.S. This school serves around 700 students each year. For the 2018–2019 school year (when data collection took place), this school reported a student population that was 60% White, 25% Black, 9% Asian, and 5% Hispanic. Around 8% were enrolled in the free and reduced-price lunch program.

No demographic data were collected from individual students, but informal observations of the classes where research was conducted suggested that these classes appeared to reflect the school-level demographics for race, and sex was also well-balanced in this sample. Additionally, because this school was an academically competitive magnet school, students were not being tracked into high and low-performing groups, which sometimes leads to defacto segregation of meaningful demographic categories.

Students came from four classrooms of 21 to 23 students, each. Four or five students were seated at a table but worked independently on separate laptops. They occasionally talked to others, but there was not any sustained collaboration. The study lasted seven school days. On day 1, students spent 30 to 45 min completing a self-report questionnaire and a paper-based pretest. The self-report questionnaire measured students' motivation, and the pretest assessed their prior domain knowledge of climate change and causal relationships. On day 2, they received a 30-minute training about how to use Betty's Brain. Over the next four days, they spent 30 to 45 min per day learning about climate change within Betty's Brain. On the final day, students completed a posttest identical to the pretest.

2.2. Material

2.2.1. Betty's Brain

Betty's Brain is an open-ended computer-based learning environment that uses a learning-by-teaching approach (Biswas et al., 2016). Students learn about scientific phenomena, such as climate change and thermoregulation, by teaching Betty, a virtual student. Specifically, students build a causal map describing scientific phenomena, in which causal (cause-and-effect) relationships are represented by a set of concepts connected by directed links (see Fig. 1). To build this map, students can access hypermedia resource pages on relevant scientific concepts. Students can evaluate their causal modeling progress by asking Betty to take graded quizzes or by querying her on cause-and-effect questions related to what she has been taught. Betty's quiz grades or her explanations of her answers can help the student keep track of her progress (and thus their own). By looking at Betty's correct and incorrect answers, students can identify problems in their causal map. They can then improve their understanding of the topic by reading the resource pages and correcting those problems (e.g., missing links and incorrect links between concepts).

Students can also ask Mr. Davis, a virtual pedagogical agent

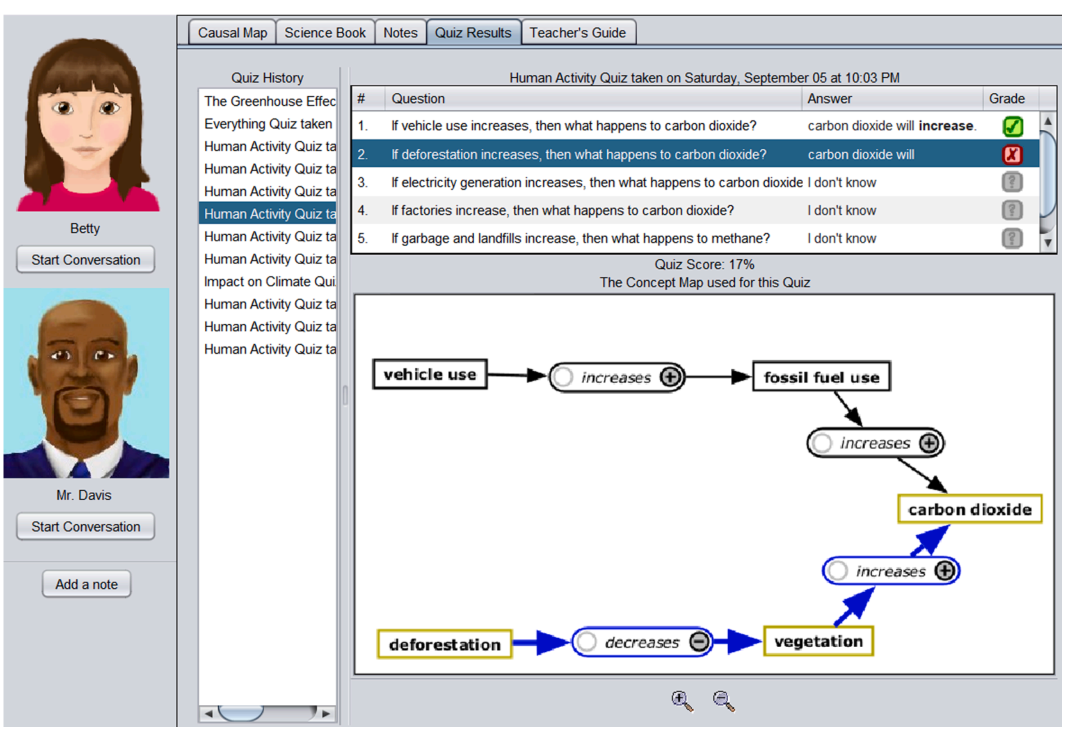

Fig. 1. *Screenshot of viewing quiz results and checking the chain of links Betty used to answer a quiz question.*
Note. The top right shows the quiz questions, answers, and grades. A gray grade means Betty could not answer the question because the question involved concepts or links that had not been added to the map. The second question, which was answered incorrectly, was selected, and the concepts and links that Betty used to answer this question were highlighted at the right bottom.

described to students as an experienced teacher, for help if they do not know how to use the system. In some situations, Mr. Davis may intervene if the student has difficulties and is not making progress in their map-building tasks. He may prompt students to read resource pages containing information that could improve the causal map (e.g., "You should go and read the page on Deforestation and Carbon Dioxide"). The prompts are triggered in conditions such as "quiz score has not improved in the students' last five attempts at updating their map".

The learning unit used in this study was on the topic of climate change. This unit was organized into four sections, including the introduction, greenhouse effect, human activities, and impacts on climate. It contained ten hypertext pages and covered 22 relevant scientific concepts and 27 causal relationships between these concepts.

2.2.2. Metacognitive strategy use

This study operationalizes metacognitive strategy use as coherent actions (Segedy et al., 2015a). Coherent actions are actions that support later actions or are based on prior actions. For instance, in Fig. 1, the quiz results could inform students that the causal links between deforestation and carbon dioxide were incorrect. After viewing these quiz results, if students read the resource pages that contained information about the correct relationship between the two concepts, the viewing and reading actions were coherent. Coherent actions imply the use of metacognitive strategies because a coherent action entails that a student monitors information generated by the prior actions (e.g., viewing quiz results) and adapt current actions (e.g., reading pages) based on the acquired information (Segedy et al., 2015a; Zhang et al., 2020). The two actions do not need to be consecutive, but it is necessary to restrict the time interval between them. Prior research in Betty's Brain found that students usually used information within 5 min of encountering the information (Segedy, 2014). The proportion of actions not supported by prior actions within 5 min was negatively correlated with students' map scores within Betty's Brain (the number of correct causal links minus the number of incorrect links; Segedy et al., 2015a). In contrast, the proportion of information that was used within 5 min of encountering the information was positively related to students' map scores and changes between pretest and posttest scores (Segedy et al., 2015a).

We analyzed students' action logs from Betty's Brain to identify coherent actions. There were five kinds of coherent actions in Betty's Brain: coherent *viewing*, *prompts*, *edits*, *reading,* and *marking*. Each type of coherent action had its incoherent counterpart, which might provide complementary information for understanding how metacognitive strategy use evolves. Thus, incoherent metrics were used in the analysis for RQ1 and defined here.

1. Coherent viewing was viewing quiz results actions that were coherent with later actions. It measured whether students used assessment results to support later activities and might indicate self-monitoring (Zhang et al., 2020).

Incoherent viewing was viewing quiz results actions that did not support later actions. In other words, students did not utilize information generated by viewing quiz results to guide reading or editing actions.

2. Coherent prompts. Mr. Davis might give prompts recommending students to read resource pages containing information that could improve the causal map. Coherent prompts were the prompts that students used, i.e., students read the resource pages recommended by the prompts. Coherent prompts assessed whether students utilized external feedback and might reflect self-control.

Incoherent prompts were prompts that students received but did not use.

3. Coherent edits were map edit actions that were based on reading actions or viewing quiz results actions. It measured whether students edited the concept map based on previously acquired information and might reflect self-control (Zhang et al., 2020).

Incoherent edits were map edit actions that were not supported by reading or viewing quiz results actions. For instance, if a student added a link between deforestation and carbon dioxide but did not read pages about their relationships, this edit would be incoherent no matter whether the link was correct or not.

4. Coherent reading was page reading actions that were based on quiz results or prompts from the system. It measured whether students intentionally sought relevant information to improve their understanding based on the quiz results or the prompt (Zhang et al., 2020). Thus, coherent reading might indicate self-control.

Incoherent reading was page reading actions that were not based on quiz results or reading prompts from the system. For example, in Fig. 1, after viewing the quiz results, if a student read the resource pages that contained information about the relation between deforestation and carbon dioxide, this reading action would be coherent. By contrast, if the student read resource pages that did not contain any information related to the quiz questions answered incorrectly, the reading action would be incoherent. Note that if students read a page that they did not read before, these reading actions would not be labeled incoherent even though they were not supported by the quiz results or reading prompts. Such reading might indicate that students did not know what to read next and opened a new page randomly or that they intentionally searched for information about causal links not in the map. We could not verify which reason drove these reading actions. Thus, they were excluded from the analysis.

5. Coherent marking was marking actions that were based on quiz results. This variable reflected how often, based on the quiz results, students understood what links on their map were correct or possibly incorrect and annotated them accordingly. Coherent marking might represent constructive monitoring behaviors because the marking action translates quiz results into systematic checking of the causal maps (Zhang et al., 2020).

Incoherent marking was marking actions that were not based on quiz results. For instance, students labeled a causal link correct or wrong without using quizzes to test its correctness.

Note that we did not examine incoherent metrics in the analyses for RQs 2 to 4 because these RQs focused on the use of metacognitive strategy, as measured through coherent actions (Segedy et al., 2015a; Zhang et al., 2020). No evidence supports the idea that incoherent actions may indicate metacognitive strategy use, although we do not assume that incoherent actions represent less effective or inadequate strategies.

In the current study, over half of the students did not use the marking functionality (86.6%, 56.5%, 47.7%, and 50.6% in the first, second, third, and fourth days of using Betty's Brain, respectively). Low marking usage provided limited information about students' differences in the application of metacognitive strategies. Thus, coherent and incoherent marking actions were not analyzed.

The number of coherent edits and the number of coherent prompts were computed per day per student. For coherent reading and viewing, we used the sum of duration per day per student as indicators rather than the action counts because the duration of reading and viewing actions could vary from seconds to minutes. Using the action counts as the indicators of coherent reading and viewing would imply that a coherent reading action with a span of 10 s is equivalent to a coherent reading action with a span of 100 s. The time on Betty's Brain varied across days and students, and thus, we divided the coherent action metrics by hours on Betty's Brain to make them comparable within and between students.

2.2.3. Prior domain knowledge

Domain knowledge was operationalized as pretest scores. The test assessed knowledge of climate change and causal relationships and contained seven multiple-choice and three short-answer questions. Each question involved both climate change and causal relationships. Each multiple-choice question had four choices, and students got one point if they answered a question correctly. Short-answer questions asked students to explain how one factor influenced another based on their understanding of the causal relations among concepts in the climate change domain. The correct answer to each question contained three or four successive causal links between a relevant set of concepts. A student got one point if their answers had one link that was the same or close to a link in the correct answer. The appendix presents two example questions. Students could get a total maximum score of 18 points. A posttest was administrated to check students' learning. The coefficient alpha was 0.75 and 0.84 for the pretest and posttest, respectively, indicating acceptable to satisfactory internal consistency, given the small number of items (Cortina, 1993).

2.2.4. Motivational factors

The self-report questionnaire measured two motivational factors: task value and self-efficacy. All items were scored on a 5-point Likert scale. Task value reflected students' perceived importance and utility of science in general and the learning topic (i.e., climate change). It was measured by three slightly modified items from the Science Learning Value subscale of Students' Motivation toward Science Learning questionnaire (SMTSL; Tuan et al., 2005). The coefficient alpha was 0.69, indicating acceptable internal consistency, given the few items (Cortina, 1993). The average item score was used as an indicator.

Self-efficacy represented the extent to which students thought they were able to learn science in general. Three items from the self-efficacy subscale of the SMTSL questionnaire measured it (Tuan et al., 2005). The coefficient alpha was 0.75, indicating acceptable internal consistency, given the few items (Cortina, 1993). The average item score was used as an indicator.

2.3. Data analyses

For RQ1, a one-way repeated analysis of variance (ANOVA) was conducted with each coherent metric as the dependent variable and day as the within-student independent variable. If a coherent metric varied across days, we used post-hoc pairwise comparisons to determine which pairs of consecutive days showed a significant difference. We applied a Bonferroni correction for multiple comparisons.

For RQ2, we fitted mixed models to the data using the *lme4* package in R (Bates et al., 2015), with each coherent metric as a response variable and day as the predictor. We compared models with and without the random effect of the day. When examining the random effect, the commonly used Wald test and the likelihood-ratio test (LRT) will be invalid if the random effect is zero (Stram & Lee, 1994). The mixture chi-square LRT is a better option (Stram & Lee, 1994). This test computes the LRT statistic and compares the statistic to two chi-square distributions. One chi-square distribution has a degree of freedom the same as the chi-square distribution in the commonly used LRT, while another has a degree of freedom one less than the commonly used LRT if only one random effect is examined. Then, it averages the two p-values returned by the comparison. We rejected the null hypothesis (e.g., the random effect of day on coherent reading equals zero) when the average p-value was$<$0.05, suggesting that the evolution of coherent reading may vary across students. We fitted linear mixed models for coherent reading and coherent viewing since their durations were continuous and log-linear mixed models with the Poisson distribution for coherent edits and coherent prompts because their frequencies were count variables (Snijders & Bosker, 2012). Taking coherent reading as an example, the linear mixed model without the random effect for the day was the following:

$$CR_{ij} = \gamma_{00} + \gamma_{10} \times day_{ij} + U_{0j} + R_{ij} \tag{1}$$

CR_{ij} refers to the duration of coherent reading per hour for student j on the i^{th} day. γ_{00} is the fixed effect of the intercept, and U_{0j} is the random effect of the intercept. If we let the value of day_{ij} in the first day equal to 0, i.e.,$day_{1j} = 0$, γ_{00} and U_{0j} can be interpreted as the mean and variance of the coherent reading duration across students in the first day, respectively. γ_{10} is the fixed effect of day. R_{ij} is residuals and independent of U_{0j}. With the random effect of day, the linear mixed model becomes:

$$CR_{ij} = \gamma_{00} + \gamma_{10} \times day_{ij} + U_{0j} + U_{1j} + R_{ij} \tag{2}$$

U_{1j} is the random effect of day, related to U_{0j} but independent ofR_{ij}.

For RQ3 and RQ4, we added domain knowledge, motivation, and their interaction with day to the mixed effect models. Taking coherent reading as an example, the linear mixed effect model is:

$$CR_{ij} = \gamma_{00} + \sum_k \gamma_{0k} \times Z_{kj} + \left(\gamma_{10} + \sum_k \gamma_{1k} \times Z_{kj}\right) \times day_{ij} + U_{0j} + U_{1j} + R_{ij} \tag{3}$$

Z_{kj} represents domain knowledge and motivation in student j. γ_{0k} is the main effect of these student-level predictors, and γ_{1k} is their interaction with day. The fixed effects were tested via 95% bootstrap confidence intervals with 1,000 resampling iterations (Davison & Hinkley, 1997). All predictors were grand-mean centered to mitigate the collinearity between the main and the interaction effects and ease the interpretation (Enders & Tofighi, 2007).

3. Results

3.1. Preliminary analyses

Table 1 displays the means, standard deviations, and correlations among pretest and posttest scores and motivational factors. The paired sample t-test showed that posttest scores were significantly higher than pretest scores ($t = 11.60$, $p < .001$), and the effect size was large (Cohen's $d = 1.20$), indicating that students learned from using Betty's Brain. Task value was related to neither pretest scores nor posttest scores. Self-efficacy was positively related to posttest scores. As in previous research (Liem et al., 2008; Üner et al., 2020; Wolters & Pintrich, 1998), the two motivational factors were positively related to each other.

Fig. 2 displays the means of coherent and incoherent metrics with 95% confidence intervals per day. There were remarkable increases in all coherent metrics from the first day to the second day. Daily increases in the remaining days were relatively weak, except for coherent viewing between the second and third days, the average duration of which rose from 6.54 min to 9.87 min per hour. For incoherent metrics, no noticeable change was found in the frequency of edits and the duration of viewing. There was a moderate decrease in the duration of incoherent reading from the first to the second day. Incoherent prompts increased gradually over the four days.

3.2. The evolution of metacognitive strategy use

Table 2 shows the results of the repeated ANOVA. The results indicated that all actions had significant variations across days. Thus, we conducted post-hoc pairwise comparisons to determine which pairs of consecutive days showed a significant difference. Table 2 displays the results. In line with Fig. 2, the post-hoc pairwise comparison indicated that all coherent actions differed significantly between days 1 and 2. Students edited the causal map more frequently based on collected information (coherent edit), adhered more often to the reading prompts

Table 1
Descriptive statistics and correlations of time-invariant variables.

	Pretest scores	Posttest scores	Task value	Self-efficacy
Posttest scores	0.63**	–	–	–
Task value	0.17	0.19	–	–
Self-efficacy	0.21	0.31*	0.44**	–
M	6.22	9.39	4.24	3.45
SD	2.62	3.16	0.55	0.68

Note: **p* < .05, ***p* < .01. Bonferroni correction was applied for multiple tests.

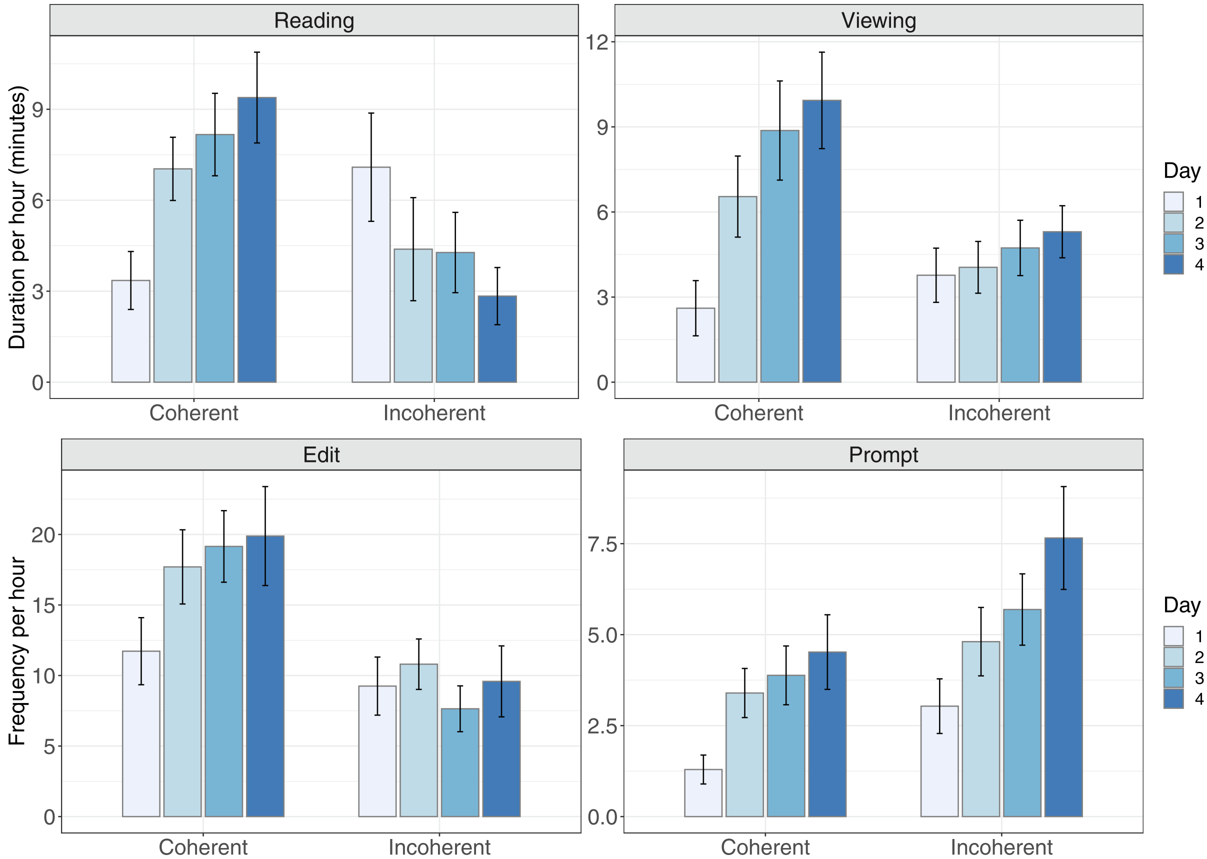

Fig. 2. The average frequency or duration of coherent actions with 95% confidence intervals.

Table 2
Results of repeated ANOVA and post hoc pairwise comparisons.

Dependent variable		ANOVA		Day 1 vs. 2		Day 2 vs. 3		Day 3 vs. 4	
		F	η^2	*t*	Cohen's *d*	*t*	Cohen's *d*	*t*	Cohen's *d*
Reading	Coherent	13.19	0.14***	6.57	0.84***	1.96	0.22	1.37	0.16
	Incoherent	5.48	0.06***	−2.59	0.33*	−0.58	0.06	−2.00	0.23
Viewing	Coherent	12.78	0.12***	5.22	0.67***	2.74	0.31*	1.14	0.13
	Incoherent	3.58	0.05*	0.48	0.06	1.04	0.12	1.05	0.12
Edit	Coherent	10.80	0.09***	6.89	0.88***	1.82	0.20	0.97	0.11
	Incoherent	3.21	0.04*	1.28	0.16	−3.27	0.37**	1.83	0.21
Prompt	Coherent	12.36	0.14***	5.95	0.76***	1.18	0.13	1.43	0.16
	Incoherent	9.79	0.10**	3.45	0.44**	2.58	0.29*	2.97	0.34*

Note: *$p<.05$; **$p<.01$; ***$p<.001$. The independent within student variable is day. Bonferroni correction was applied for multiple tests.

from the learning system (coherent prompt), and spent more time on reading pages containing information that could improve the causal map (coherent reading). They also spent more time checking quiz results that generated information supporting later actions (coherent viewing). Coherent viewing also differed between days 2 and 3.

For incoherent metrics, there were significant decreases in incoherent reading between days 1 and 2 and in incoherent edits between days 2 and 3. No significant daily change was found in incoherent viewing. Incoherent prompts increased significantly between all pairs of days. Overall, coherent actions increased over time, but incoherent actions decreased or did not change, except for incoherent prompts. Comparing the effect sizes of the change in coherent and incoherent prompts showed that the effect size for coherent prompts ($\eta^2 = 0.14$) was larger than for incoherent prompts ($\eta^2 = 0.10$).

Since the duration of coherent viewing increased from the first to second days and the second to third days, we were interested in whether the daily growth rate was constant from the first to third days. We created a new day-related variable, day_34, which grouped the third and fourth days together (i.e., the value of day_34 is 0, 1, 2, and 2 for the first, second, third, and fourth days, respectively). We fitted two linear mixed models with coherent viewing as the response variable and day_34 as the predictor[1]. The difference between the two models was that day_34 was categorical in one model but numeric in another. The model with categorical day_34 assumed that the daily growth rate from the first to the third days might vary, while the model with numeric day_34 supposed that the daily growth rate was constant. The latter was

[1] We considered linear and quadratic models with the raw day variable (i.e., day = 0, 1, 2, 3 for the first, second, third, and fourth days, respectively). With the raw day variable, the quadratic model fitted the data better than the linear model ($\chi^2 = 5.39$, $df = 1$, $p = .02$). However, the quadratic model with the raw day variable was not superior to the linear model with day_34 ($\chi^2 = 2.13$, $df = 1$, $p = .14$).

a restricted version of the former. Thus, the likelihood ratio test (LRT) could be used to determine whether the model with numeric day_34 fitted the data the same as the model with categorical day_34. The LRT revealed that the two models fitted the data the same ($\chi^2 = 0.70$, $df = 1$, $p = .40$), indicating that the daily growth rate in coherent viewing was constant from the first to the third day. Thus, in subsequent analyses, we used the numeric day_34.

To examine whether the increase in coherent metrics varied across students, we compared mixed models with and without the random effect of the day. There were increases from the first to the second days for coherent edit, read, and prompts, so we used a dummy day-related variable, not_first_day, which grouped the second, third, and fourth days together (i.e., for the first day, the value of not_first_day is 0, and for the other days, the value of not_first_day is 1). For coherent viewing, we used the day_34 variable described above. Table 3 shows the results of the mixture chi-square LRT. The evolution of coherent edits, read, and viewing varied across students, while the growth of coherent prompts did not.

3.3. *The influence of domain knowledge, task value, and self-efficacy*

Tables 4 to 7 present the results of mixed models for coherent reading, viewing, edits, and prompts, respectively. The base model only included the day as the explanatory variable. Models 1.1 and 1.2 contained domain knowledge and its interaction with the day. Models 2.1 and 2.2 had motivational factors and their interactions with the day. The final model only included variables that significantly predicted a coherent metric.

Domain knowledge predicted all coherent metrics. Students with higher domain knowledge had more coherent edits and prompts and spent more time on coherent reading and viewing than those with lower domain knowledge on the first day. Table 5 shows that the random effect of the day on coherent viewing decreased from 2.52 to 2.31 from Model 1.1 (without the interaction between domain knowledge and day) to Model 1.2 (containing the interaction), indicating that domain knowledge explained 8.3% of the random effect of day on coherent viewing. The moderating effect was statistically significant at the 0.01 significance level ($\gamma = 0.30$, 95% and 90% bootstrapped confidence intervals were [-0.02, 0.62] and [0.03, 0.57], respectively). On average, for students with one SD higher in domain knowledge, their increases in coherent viewing from the first to the third days was 3.14 min higher than the increase in students with one SD lower in domain knowledge[2]. Compared with the average increase in coherent viewing (6.00 min), the moderation effect of domain knowledge might be at a medium level.

Task value predicted coherent viewing (see Table 5). Students that valued the task spent more time on coherent viewing. In Table 6, the random effect of the day on coherent edits decreased from 0.42 to 0.36 from Model 2.1 (without the interaction between task value and day) to Model 2.2 (containing the interaction), indicating that task value explained 14.29% of the random effect. The moderating effect was statistically significant at the 0.05 significance level (γ = -0.39, 95% bootstrapped confidence intervals = [-0.76, -0.01]). The moderating effect indicates that students with high task value had fewer increases in coherent edits from the first to the second day. On average, for students with one SD higher in task value, their increase rate in coherent edits from the first to the second days was 65.12% of the increase rate in students with one SD lower in task value[3].

Self-efficacy neither predicted any coherent metric nor moderated their changes over days. Its coefficient estimates were much smaller than the coefficient estimates of task value.

4. Discussion

SRL is a dynamic process that unfolds over time (Azevedo, Moos, et al., 2010; Winne & Hadwin, 2008), and thus, researchers have been increasingly interested in the temporal characteristics of SRL events (Greene et al., 2021; Molenaar & Järvelä, 2014; Paans et al., 2019). This study investigated how metacognitive strategy use evolved daily within an open-ended learning environment and whether students showed different rates of change. It also examined whether these differences were related to students' domain knowledge, task value, and self-efficacy.

4.1. *The evolution of metacognitive strategy use (RQ1 + RQ2)*

The current study found that students' coherent actions increased over time, while their incoherent actions decreased, did not change, or experienced relatively fewer increases (depending on the type of action). The η^2 ranged from 0.04 to 0.14. The range of effect size matches prior work, where the η^2 of students' behavioral change over time fell between 0.01 and 0.22 (Paans et al., 2019). The increased use of metacognitive strategies (operationalized as coherent actions) might be related to the increases in familiarity with the learning content (climate change) and environment (Betty's Brain). As the learning process unfolded, students knew more about climate change (this is supported by significant gains from the pretest scores to the posttest scores). Both the present and prior studies have found that domain knowledge was positively related to metacognitive strategy use (Li, 2019; Moos & Azevedo, 2008; Taub et al., 2014). Increases in domain knowledge might enhance students' capacity for applying metacognitive strategies. In addition to becoming more familiar with the learning material, learners also became more familiar with the interface features of Betty's Brain over time. Familiarity with Betty's Brain features enabled students to utilize the tools offered by the system effectively, such as adopting the advice that recommended individual book pages, having Betty take a quiz, and analyzing the quiz results to identify incorrect causal links.

The growth of coherent actions lessened after the first two days. One possible explanation may be that, after using Betty's Brain for two days, students were now relatively familiar with the features of Betty's Brain and might have determined what strategies they found helpful towards making progress on the learning task. Thus, their coherent actions became stable. Alternatively, the students in this study might not possess sufficient working memory capacity to continue improving the frequencies of coherent actions. Sixth graders' cognitive and metacognitive skills are still under development (de Bruin et al., 2011), and the application of metacognitive strategy requires effortful control of behavior (Efklides, 2011). Moreover, students might become bored with the task (Roscoe et al., 2013) and less willing to exert effortful control of behavior. Decreases in motivation might mitigate the effect of increases in domain knowledge. Nevertheless, this hypothesis needs further examination.

Note that coherent prompts did not increase from the second to the fourth day, but incoherent prompts continued rising. Prompts recommended pages for reading, and incoherent prompts were those that students received but did not read the recommended pages. The results meant that students received an increased number of prompts from the system over days, but the number of prompts they followed each day remained stable. That is, they might select which prompts to follow. This may be a sign of more effective learning regulation compared to following all prompts provided by the system.

The second research question focused on whether the evolution of metacognitive strategy use varied across students. The increases in the frequency of adopting reading prompts showed no individual

[2] The 3.14 min was the product of 0.30 (the coefficient of the domain knowledge*day interaction in model 1.2 of Table 5), 2 (two days between the first and third days), 2.62 (the SD of domain knowledge in Table 1), and 2 (two SDs).

[3] The 65.12% was the natural exponential to the power of the product of -0.39 (the coefficient of the task value*day interaction in the final model of Table 6), 0.55 (the SD of task value in Table 1), and 2 (two SDs).

Table 3
Results of the mixture chi-square LRT for the random effect of the day.

	Model	No. τ	Deviance	Test statistics	p from χ^2_2	χ^2_1	average
Coherent reading	Null	1	2026.3				
	Full	2	2020.9	5.4	0.069	0.020	0.043
Coherent viewing	Null	1	2129.9				
	Full	2	2106.8	23.1	< 0.001	< 0.001	< 0.001
Coherent edit	Null	1	2360.8				
	Full	2	2300.8	60.0	< 0.001	< 0.001	< 0.001
Coherent prompt	Null	1	1288.7				
	Full	2	1284.5	4.2	0.120	0.040	0.080

Note: Null model, only a random effect for intercept. Full model, containing random effects for intercept and day. No. τ, the number of random effects.

Table 4
The linear mixed model for coherent reading.

	Base model	Model 1.1	Model 1.2	Model 2.1	Model 2.2	Final Model
Fixed effects						
Intercept	6.93 [6.19, 7.70]*	6.91 [6.16, 7.67]*	6.91 [6.13, 7.65]*	6.94 [6.09, 7.72]*	6.91 [6.16, 7.72]*	6.91 [6.21, 7.59]*
not_first_day	4.89 [3.44, 6.31]*	4.86 [3.34, 6.26]*	4.85 [3.38, 6.28]*	4.88 [3.39, 6.29]*	5.00 [3.54, 6.44]*	4.86 [3.44, 6.25]*
domain knowledge		0.44 [0.12, 0.72]*	0.44 [0.17, 0.73]*			0.44 [0.15, 0.73]*
not_first_day*domain knowledge			0.06 [−0.49, 0.62]			
Task value				−0.57 [−2.22, 0.89]	−0.65 [−2.13, 1.01]	
Self-efficacy				0.23 [−1.00, 1.57]	0.28 [−0.99, 1.48]	
not_first_day*Task value					−1.93 [−4.84, 0.80]	
not_first_day*Self-efficacy					−0.15 [−2.57, 2.27]	
Random effects						
Intercept	6.33	5.21	5.21	6.15	6.30	5.21
not_first_day	3.78	3.88	3.87	3.66	3.23	3.88

Note: not_first_day = 0 if the first day; else not_first_day = 1. *, the 95% confidence intervals do not contain zero.

Table 5
The linear mixed model for coherent viewing.

	Base model	Model 1.1	Model 1.2	Model 2.1	Model 2.2	Final Model
Fixed effects						
Intercept	6.25 [5.30, 7.26]*	6.22 [5.30, 7.16]*	6.21 [5.22, 7.17]*	6.19 [5.15, 7.14]*	6.23 [5.29, 7.25]*	6.18 [5.22, 7.11]*
day_34	2.99 [2.15, 3.82]*	2.97 [2.09, 3.78]*	2.96 [2.13, 3.81]*	3.05 [2.17, 3.87]*	3.00 [2.15, 3.87]*	3.00 [2.12, 3.87]*
domain knowledge		0.57 [0.16, 0.93]*	0.56 [0.21, 0.93]*			0.52 [0.18, 0.85]*
day_34*domain knowledge			0.30 [−0.02, 0.62]			
Task value				2.04 [0.04, 3.88]*	1.99 [0.17, 4.04]*	2.06 [0.20, 3.77]*
Self-efficacy				0.66 [−0.87, 2.36]	0.63 [−0.89, 2.21]	
day_34*Task value					0.33 [−1.37, 1.95]	
day_34*Self-efficacy					0.60 [−0.83, 2.05]	
Random effects						
Intercept	12.14	9.89	10.09	10.47	10.62	9.09
day_34	2.90	2.52	2.31	2.78	2.67	2.31

Note: day_34 = 0 if the first day; day_34 = 1 if the second day; else day_34 = 2. *, the 95% confidence intervals do not contain zero.

differences, but we found variation across students in the growth of coherent edits, reading, and viewing. The individual differences in the growth of coherent edits and viewing were partially explained by students' domain knowledge and task value. We discuss these in the following sections.

4.2. The impact of prior domain knowledge on metacognitive strategy use (RQ3)

Prior domain knowledge was positively associated with all types of metacognitive strategy use. This was in line with prior studies (Moos & Azevedo, 2008, 2009a; Taub & Azevedo, 2019). High domain knowledge allows learners to process novel information with less working memory capacity (Taub et al., 2014). Students with high domain knowledge might be able to apply more working memory resources to monitor and control activities and execute more coherent actions.

The moderation effect of domain knowledge on the day was statistically significant for coherent viewing at the 0.10 significance level. The random effect of the day on coherent viewing decreased 8.3% after adding the moderation effect of domain knowledge to the model. The reason for this moderation effect might be that domain knowledge is associated with efficient learning (Alexander et al., 1994; Beier & Ackerman, 2005). High domain knowledge students might learn faster than low domain knowledge students, and the greater daily knowledge acquisition in climate change, in turn, led to greater increases in coherent viewing. Overall, the result suggests that domain knowledge may contribute to the evolution of metacognitive strategy use. Future research can examine this assumption by using larger sample sizes, measuring daily learning gains, and relating it to the evolution of metacognitive strategy use.

4.3. The impact of motivation on metacognitive strategy use (RQ4)

In line with existing findings (Üner et al., 2020; Wolters & Pintrich, 1998), task value positively predicted the application of one

Table 6
The log-linear mixed models with the Poisson distribution for coherent edits.

	Base model	Model 1.1	Model 1.2	Model 2.1	Model 2.2	Final Model
Fixed effects						
Intercept	2.60 [2.46, 2.74]*	2.60 [2.46, 2.73]*	2.60 [2.47, 2.74]*	2.60 [2.45, 2.74]*	2.59 [2.46, 2.74]*	2.60 [2.47, 2.72]*
not_first_day	0.68 [0.47, 0.89]*	0.66 [0.47, 0.86]*	0.68 [0.48, 0.90]*	0.68 [0.49, 0.88]*	0.70 [0.50, 0.92]*	0.68 [0.49, 0.87]*
domain knowledge		0.09 [0.04, 0.14]*	0.10 [0.05, 0.15]*			0.09 [0.04, 0.14]*
not_first_day*domain knowledge			−0.05 [−0.13, 0.03]			
Task value				0.01 [−0.26, 0.29]	0.08 [−0.21, 0.36]	0.05 [−0.17, 0.29]
Self-efficacy				0.08 [−0.16, 0.31]	0.08 [−0.14, 0.31]	
not_first_day*Task value					-0.38 [-0.76, -0.01]*	-0.39 [-0.76, -0.02]*
not_first_day*Self-efficacy					-0.00 [-0.28, 0.31]	
Random effects						
Intercept	0.43	0.36	0.36	0.42	0.42	0.35
not_first_day	0.42	0.42	0.42	0.42	0.36	0.37

Note: not_first_day = 0 if the first day; else not_first_day = 1. *, the 95% confidence intervals do not contain zero. The coefficient is at the log scale of coherent edits. Taking the coefficient of domain knowledge in the final model as an example, holding other variables, one unit increase in domain knowledge means that coherent edits per hour were expected to increase $e^{0.09} = 1.09$ times.

Table 7
The log-linear mixed models with the Poisson distribution for coherent prompts.

	Base model	Model 1.1	Model 1.2	Model 2.1	Model 2.2	Final Model
Fixed effects						
Intercept	0.91 [0.74, 1.06]*	0.90 [0.72, 1.04]*	0.89 [0.73, 1.03]*	0.90 [0.73, 1.06]*	0.89 [0.72, 1.05]*	0.90 [0.74, 1.04]*
not_first_day	1.12 [0.86, 1.43]*	1.12 [0.87, 1.42]*	1.17 [0.88, 1.52]*	1.13 [0.89, 1.45]*	1.16 [0.88, 1.55]*	1.12 [0.88, 1.42]*
domain knowledge		0.11 [0.05, 0.16]*	0.12 [0.06, 0.18]*			0.11 [0.06, 0.17]*
not_first_day*domain knowledge			−0.07 [−0.18, 0.05]			
Task value				0.15 [−0.17, 0.47]	0.18 [−0.16, 0.50]	
Self-efficacy				0.09 [−0.14, 0.36]	0.11 [−0.14, 0.39]	
not_first_day*Task value					−0.15 [−0.84, 0.46]	
not_first_day*Self-efficacy					−0.10 [−0.66, 0.36]	
Random effects						
Intercept	0.38	0.32	0.32	0.37	0.37	0.32

Note: not_first_day = 0 if the first day; else not_first_day = 1. *, the 95% confidence intervals do not contain zero. The coefficient is at the log scale of coherent prompts. Taking the coefficient of domain knowledge in the final model as an example, holding other variables, one unit increase in domain knowledge means that coherent prompts per hour were expected to increase $e^{0.11} = 1.12$ times.

metacognitive strategy: coherent viewing. Students who thought the task and science were important spent more time on coherent viewing than those with lower task value on the first day. The result suggests that starting the task with high value may drive students to evaluate their understanding and decide the next step based on the evaluation results.

Students with higher task value had smaller increases in the frequency of coherent edits than those with lower task value. The reason may be their high value toward the task is not maintained through the learning process (Wigfield et al., 1997). When students devalue the activity, they may put less effort into monitoring and controlling their behavior and cognition (Wigfield et al., 2008). Thus, decreases in the value for the task may lead to smaller increases in coherent edits. It is unexpected that self-efficacy was not a significant predictor of metacognitive strategy use given prior work, as discussed below in sections 4.4 and 4.5.

4.4. Implications

This study contributes to our understanding of SRL as a series of events unfolding over time. We found that students evolved toward increased use of metacognitive strategies and decreased use of less effective behaviors. This result suggests that learners as young as sixth-grade students may be able to regulate behaviors adaptively in an open-ended learning environment. However, their adaptivity may be limited since the increases in metacognitive strategy use lessened after the first two days. The change in the magnitude of increases or decreases of SRL behaviors implies opportunities for adaptive scaffolding (De Backer et al., 2016). For example, in Betty's Brain, the number of coherent actions increased little after the first two days, but one-fourth to one-third of reading, viewing, and editing actions were still incoherent (see Fig. 2). Thus, during these days, students may need scaffolding that differs from the first two days to enhance metacognitive strategy use. However, further investigation is needed to determine why increases in coherent actions tapered off and what scaffolding could help.

This study found that task value explained the individual differences in the evolution of metacognitive strategy use, but self-efficacy did not. Moreover, self-efficacy was not related to the overall frequency of any metacognitive strategy use. Motivation have been viewed as factors that are worth adapting instruction to (Aleven et al., 2017; Shute & Zapata-Rivera, 2012). The current findings suggest that adapting metacognitive strategy scaffolding to task value may be better than adapting to self-efficacy. At least, it may not be beneficial to adapt scaffolding to task-independent self-efficacy, which was measured in this study.

Prior domain knowledge positively predicted the use of metacognitive strategy use, no matter which type of metacognitive strategy. This effect implies that it is necessary to provide learners with tailored scaffolding based on their domain knowledge in open-ended learning environments (Land, 2000). Learners with low prior knowledge need more support on utilizing the functions offered by open-ended learning environments. For example, in Betty's Brain, learners with low prior knowledge may need longer training on how to use this system. After they start learning activities in Betty's Brain, the system may provide them with scaffolding to interpret quiz results, extract causal relationships between concepts from the resource page, and transfer their understanding to the causal map. Furthermore, we found a positive association between prior knowledge and the use of prompts (see Table 7), which indicates that learners with low prior knowledge may be less likely to utilize the support of Betty's Brain. Thus, instructors and learning systems may particularly demonstrate and emphasize the usefulness of the support to learners with low prior knowledge. For high prior knowledge learners, the system may reduce the spontaneous support and provide support only when they request, given that the expertise reversal effect suggests that superfluous support may hamper high-domain knowledge learners' learning (Kalyuga, 2007).

The prompt in Betty's Brain was adaptive because it was triggered in particular conditions, such as missing a quiz question multiple times (Biswas et al., 2016). Clarebout and Elen (2008) found that students were more inclined to take adaptive advice on using tools than random advice during open-ended learning. The current study found that students only used a stable number of prompts, even though they received increasing prompts over days. Thus, offering more advice to learners does not necessarily increase the frequency of advice adoption, even though the advice is adaptive. Moreover, the increasing number of unused prompts raises the concern that excessive prompts may frustrate students and interfere with learning. If so, the prompt triggering mechanism should ensure that learners will not receive excessive prompts. This may be achieved by restricting the number of prompts that the system offers to learners within a period and avoiding duplicate prompts.

4.5. Limitations and further research

One limitation of this study is the lack of student-level demographics in this sample. We collected school-level demographics, which we expected to resemble the sample demographics closely, but the sample's specific race, age, and gender distributions were unknown. This is an acknowledged weakness since prior research has shown that demographic differences relate to metacognition and SRL. Specifically, prior research has shown that metacognitive skills are still developing in middle school students (de Bruin et al., 2011), but that gendered differences in these constructs have already begun to emerge (see discussion in Pajares, 2002). Velayutham et al.'s (2012) research suggests that task value shows gendered differences in SRL, for example, but Schnell et al.'s (2015) work on self-efficacy and SRL strategies did not find a moderating effect of gender. Although gender was well-balanced in this study, our results may still be obscuring important differences that might be relevant to identifying which SRL interventions are most likely to help.

Similarly, previous research has shown differences in self-regulation strategies by students from different ethnic backgrounds. For help-seeking behaviors, these demographic differences have been studied at both the student and the school level (Karumbaiah et al., 2021; Schenke et al., 2015). While the driving forces for these differences are not fully understood, they do suggest the importance of incorporating demographic variables into future research.

Future research should make greater efforts to incorporate demographics into analyses. For example, researchers may examine whether demographics explain individual differences in the evolution of metacognitive strategy use and moderate the relationships between other variables and metacognitive strategy use. Such results may deepen the understanding of the role of demographic factors in the dynamic SRL process and generate insights about adapting support to these factors. Collecting this demographic information is also vital to other kinds of analyses, for example, predictive modeling, as it is the only way to ensure against algorithmic biases (Paquette et al., 2020).

Studies have found that self-efficacy is one of the most critical factors that impact learning (Chin-Chung et al., 2011; Moos & Azevedo, 2009b; Pajares, 1996a). However, in the present study, neither its main effect nor its interaction with the day was significant on any coherence metric. A possible explanation might be that the measure of self-efficacy in this study was toward science in general rather than task-specific. Compared with task-specific self-efficacy, domain self-efficacy measures have no or weaker predictive power to task performance (Liu et al., 2020; Pajares, 1996b; Ramos Salazar & Hayward, 2018). When students respond to the science self-efficacy items, they may not have knowledge about the learning topic (e.g., climate change) in mind (Pajares, 1996a). Further research may compare the associations between metacognitive strategy use and domain-specific as well as task-specific self-efficacy. Another possible explanation is that students' self-efficacy expectations may not fit the Betty's Brain context. Betty's Brain utilizes a learning-by-teaching paradigm and is open-ended (Biswas et al., 2016). It supports the development of SRL skills by requiring students to plan, regulate, and manage their activities. Self-efficacy is built on previous experiences (Bandura, 1997). In this study, previous experiences might not fit the current task because students might have little experience with learning environments similar to Betty's Brain. Thus, when investigating the role of self-efficacy in SRL, researchers need to consider how self-efficacy items touch on the learning content and environment.

This study did find associations between task value and metacognitive strategy use, but the overall effect of task value was small. The reason may be that sixth-grade students' metacognitive skills are under development (de Bruin et al., 2011). The relations between task value and regulating activities may gradually grow as children learn how to regulate their behavior. We expect that a more substantial effect of task value may be found in older populations, such as college students. Moreover, this study did not decompose task value into different facets, such as incentive and attainment value, utility value, and cost (Wigfield & Eccles, 2000), because of limited task value items. Different value facets have different associations with SRL (Wigfield et al., 2008). The effect of task value on the evolution of metacognitive strategy use may also depend on the exact facet.

Another limitation is that this study viewed motivation as static during the task. However, motivation may fluctuate as the learning process unfolds (Bernacki et al., 2015; Kalyuga, 2007). Furthermore, motivation is also a target of regulation and serves as both predictors and outcomes of self-regulation activities (Zimmerman & Schunk, 2008). The association between metacognitive strategy use and motivation may also be bidirectional during learning (Pavlik Jr, 2013). Further studies may measure both metacognitive strategy use and motivation at multiple points and investigate their mutual influence in evolution.

Metacognitive strategies generally include goal setting, planning, self-monitoring, control, and evaluation (Dent & Koenka, 2016). Metacognitive strategy use in this study mainly covered monitoring and control. The evolution of other metacognitive strategy use may be different from self-monitoring and control activities. For instance, learners' metacognitive behaviors evolved toward deep-level goal setting and monitoring activities but not planning and evaluation in one prior study (de Backer et al., 2016). Consequently, the role of domain knowledge and motivation in the evolution of metacognitive strategy use may depend on the type of metacognitive strategies.

4.6. Conclusion

This study supports the continued calls for temporal analyses of SRL events (Azevedo, 2014; Hadwin, 2021; Molenaar & Järvelä, 2014; Winne, 2010). Specifically, it shows how metacognitive strategy use evolved daily in an open-ended learning environment. Students used metacognitive strategies more frequently on the second day of learning than on the first day, implying students' adaptivity in metacognitive strategy use. The evolution of metacognitive strategy use varied across students, and task value and prior domain knowledge partially explained this variation. Task value and prior domain knowledge positively predicted the overall metacognitive strategy use. The results suggest further investigations into the role of motivation and prior domain knowledge in the temporal evolution of SRL events. The findings imply that learners may need different scaffolding at different task-solving phases. Adapting scaffolding to task-independent self-efficacy may not be useful. It may be beneficial to nudge those with low prior knowledge to utilize the scaffolding.

Funding

This work was supported by National Science Foundation [grant numbers 1561676] and the China Scholarship Council [grant numbers 201806040180].

CRediT authorship contribution statement

Yingbin Zhang: Conceptualization, Methodology, Formal analysis, Writing – original draft. **Luc Paquette:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition. **Nigel Bosch:** . **Jaclyn Ocumpaugh:** Methodology, Investigation, Writing – review & editing. **Gautam Biswas:** Methodology, Writing – review & editing, Resources, Supervision, Funding acquisition. **Stephen Hutt:** Methodology, Writing – review & editing. **Ryan S. Baker:** Methodology, Supervision, Funding acquisition.

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.

Acknowledgement

None.

Appendix. Examples of knowledge test questions

An example of the multiple-choice question:

Question: Light from the sun comes to the earth and its energy is absorbed by the atmosphere. What is the relation between this absorbed light energy and heat energy absorbed by the earth?

a. Absorbed light energy increases the amount of absorbed heat energy.

b. Absorbed light energy decreases the amount of absorbed heat energy.

c. Absorbed light energy does not change the amount of absorbed heat energy.

d. Absorbed light energy is not related to absorbed heat energy.

(Correct answer is a.)

An example of the short answer question:

Question: We know that deforestation (cutting down a large number of trees) increases the earth's absorbed heat energy. Explain, step-by-step, how deforestation increases the earth's absorbed heat energy.

Step 1: Deforestation reduces the number of trees on the earth, so more deforestation would decrease vegetation.

Step 2: When vegetation decreases, _____

Step 3: _________

Step 4: _________

References

Alexander, P. A., Kulikowich, J. M., & Jetton, T. L. (1994). The role of subject-matter knowledge and interest in the processing of linear and nonlinear texts. *Review of Educational Research, 64*(2), 201–252. https://doi.org/10.2307/1170694

Aleven, V., McLaughlin, E. A., Glenn, R. A., & Koedinger, K. R. (2017). Instruction based on adaptive learning technologies. In R. E. Mayer & P. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 522-560). Routledge.

Azevedo, R. (2014). Issues in dealing with sequential and temporal characteristics of self-and socially-regulated learning. *Metacognition and Learning, 9*(2), 217–228. https://doi.org/10.1007/s11409-014-9123-1

Azevedo, R., Behnagh, R., Duffy, M., Harley, J., & Trevors, G. (2012). Metacognition and self-regulated learning in student-centered leaning environments. In D. Jonassen & S. Land (Eds.), *Theoretical foundations of student-centered learning environments* (2nd ed., pp. 171-197). Routledge. http://doi.org/10.4324/9780203813799.

Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-Behnagh, R., Bouchet, F., & Landis, R. (2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 427-449). Springer. http://doi.org/10.1007/978-1-4419-5546-3_28.

Azevedo, R., Johnson, A., Chauncey, A., & Burkett, C. (2010). Self-regulated learning with metatutor: Advancing the science of learning with metacognitive tools. In M. S. Khine & I. M. Saleh (Eds.), *New science of learning: Cognition, computers and collaboration in education* (pp. 225-247). Springer. http://doi.org/10.1007/978-1-4419-5716-0_11.

Azevedo, R., Moos, D. C., Johnson, A. M., & Johnson, A. M. (2010). Measuring cognitive and metacognitive regulatory processes during hypermedia learning: Issues and challenges. *Educational Psychologist, 45*(4), 210–223. https://doi.org/10.1080/00461520.2010.515934

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

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman/Times Books/ Henry Holt & Co.

Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning, 9*(2), 161–185. https://doi.org/10.1007/s11409-013-9107-6

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beier, M. E., & Ackerman, P. L. (2005). Age, ability, and the role of prior knowledge on the acquisition of new domain knowledge: Promising results in a real-world learning environment. *Psychology and Aging, 20*(2), 341–355. https://doi.org/10.1037/0882-7974.20.2.341

Bernacki, M. L., Nokes-Malach, T. J., & Aleven, V. (2015). Examining self-efficacy during learning: Variability and relations to behavior, performance, and learning. *Metacognition and Learning, 10*(1), 99–117. https://doi.org/10.1007/s11409-014-9127-x

Biswas, G., Segedy, J. R., & Bunchongchit, K. (2016). From design to implementation to practice a learning by teaching system: Betty's brain. *International Journal of Artificial Intelligence in Education, 26*(1), 350–364. https://doi.org/10.1007/s40593-015-0057-9

Bouffard-Bouchard, T., Parent, S., & Larivée, S. (1991). Influence of self-efficacy on self-regulation and performance among junior and senior high-school age students. *International Journal of Behavioral Development, 14*(2), 153–164. https://doi.org/10.1177/016502549101400203

Chin-Chung, T., Shih-Chyueh, C., Jyh-Chong, L., & Meng-Jung, T. (2011). Self-efficacy in internet-based learning environments: A literature review. *Journal of Educational Technology & Society, 14*(4), 222–240.

Clarebout, G., & Elen, J. (2008). Advice on tool use in open learning environments. *Journal of Educational Multimedia and Hypermedia, 17*(1), 81–97.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge University Press.

de Backer, L., van Keer, H., Moerkerke, B., & Valcke, M. (2016). Examining evolutions in the adoption of metacognitive regulation in reciprocal peer tutoring groups. *Metacognition and Learning, 11*(2), 187–213. https://doi.org/10.1007/s11409-015-9141-7

de Bruin, A. B. H., Thiede, K. W., Camp, G., & Redford, J. (2011). Generating keywords improves metacomprehension and self-regulation in elementary and middle school children. *Journal of Experimental Child Psychology, 109*(3), 294–310. https://doi.org/10.1016/j.jecp.2011.02.005

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

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Expectancies, values, and academic behaviors* (pp. 75-146). Freeman.

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist, 46*(1), 6–25. https://doi.org/10.1080/00461520.2011.538645

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*(2), 121–138. https://doi.org/10.1037/1082-989X.12.2.121

Greene, J. A., Plumley, R. D., Urban, C. J., Bernacki, M. L., Gates, K. M., Hogan, K. A., … Panter, A. T. (2021). 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., & Azevedo, R. (2010). The measurement of learners' self-regulated cognitive and metacognitive processes while using computer-based learning environments. *Educational Psychologist, 45*(4), 203–209. https://doi.org/10.1080/00461520.2010.515935

Hadwin, A. F. (2021). Commentary and future directions: What can multi-modal data reveal about temporal and adaptive processes in self-regulated learning? *Learning and Instruction, 72*, Article 101287. https://doi.org/10.1016/j.learninstruc.2019.101287

Hannafin, M. J., Hall, C., Land, S., & Hill, J. (1994). Learning in open-ended environments: Assumptions, methods, and implications. *Educational Technology, 34*(8), 48–55.

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), 1409–1430. https://doi.org/10.1037/edu0000445

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*(4), 509–539. https://doi.org/10.1007/s10648-007-9054-3

Karumbaiah, S., Ocumpaugh, J., & Baker, R. S. (2021). Context matters: Differing implications of motivation and help-seeking in educational technology. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-021-00272-0

Kinnebrew, J. S., Loretz, K. M., & Biswas, G. (2013). A contextualized, differential sequence mining method to derive students' learning behavior patterns. *Journal of Educational Data Mining, 5*(1), 190–219. https://doi.org/10.5281/zenodo.3554617

Kinnebrew, J. S., Segedy, J. R., & Biswas, G. (2014). Analyzing the temporal evolution of students' behaviors in open-ended learning environments. *Metacognition and Learning, 9*(2), 187-215. http://doi.org/ 10.1007/s11409-014-9112-4.

Kinnebrew, J. S., Segedy, J. R., & Biswas, G. (2017). Integrating model-driven and data-driven techniques for analyzing learning behaviors in open-ended learning environments. *IEEE Transactions on Learning Technologies, 10*(2), 140–153. https://doi.org/10.1109/TLT.2015.2513387

Knight, S., Wise, A. F., & Chen, B. (2017). Time for change: Why learning analytics needs temporal analysis. *Journal of Learning Analytics, 4*(3), 7–17. https://doi.org/10.18608/jla.2017.43.2

Kuvalja, M., Verma, M., & Whitebread, D. (2014). Patterns of co-occurring non-verbal behaviour and self-directed speech; a comparison of three methodological approaches. *Metacognition and Learning, 9*(2), 87–111. https://doi.org/10.1007/s11409-013-9106-7

Land, S. M. (2000). Cognitive requirements for learning with open-ended learning environments. *Educational Technology Research and Development, 48*(3), 61–78. https://doi.org/10.1007/BF02319858

Li, L. (2019). Effect of prior knowledge on attitudes, behavior, and learning performance in video lecture viewing. *International Journal of Human-Computer Interaction, 35*(4–5), 415–426. https://doi.org/10.1080/10447318.2018.1543086

Li, S., Du, H., Xing, W., Zheng, J., Chen, G., & Xie, C. (2020). Examining temporal dynamics of self-regulated learning behaviors in STEM learning: A network approach. *Computers & Education, 158*, Article 103987. https://doi.org/10.1016/j.compedu.2020.103987

Liem, A. D., Lau, S., & Nie, Y. (2008). The role of self-efficacy, task value, and achievement goals in predicting learning strategies, task disengagement, peer relationship, and achievement outcome. *Contemporary Educational Psychology, 33*(4), 486–512. https://doi.org/10.1016/j.cedpsych.2007.08.001

Liu, Q., Liu, J., Cai, J., & Zhang, Z. (2020). The relationship between domain- and task-specific self-efficacy and mathematical problem posing: A large-scale study of eighth-grade students in China. *Educational Studies in Mathematics, 105*(3), 407–431. https://doi.org/10.1007/s10649-020-09977-w

Lowyck, J. (2014). Bridging learning theories and technology-enhanced environments: A critical appraisal of its history. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 3-20). Springer. http://doi.org/10.1007/978-1-4614-3185-5_1.

Merriam-Webster. (n.d.). Evolution. In *Merriam-Webster.com dictionary*. Retrieved November 8, 2021, from https://www.merriam-webster.com/dictionary/evolution.

Molenaar, I. (2014). Advances in temporal analysis in learning and instruction. *Frontline Learning Research, 2*(4), 15–24. https://doi.org/10.14786/flr.v2i4.118

Molenaar, I., & Chiu, M. M. (2014). Dissecting sequences of regulation and cognition: Statistical discourse analysis of primary school children's collaborative learning. *Metacognition and Learning, 9*(2), 137–160. https://doi.org/10.1007/s11409-013-9105-8

Molenaar, I., & Järvelä, S. (2014). Sequential and temporal characteristics of self and socially regulated learning. *Metacognition and Learning, 9*(2), 75–85. https://doi.org/10.1007/s11409-014-9114-2

Moos, D. C. (2014). Setting the stage for the metacognition during hypermedia learning: What motivation constructs matter? *Computers & Education, 70*, 128–137. https://doi.org/10.1016/j.compedu.2013.08.014

Moos, D. C., & Azevedo, R. (2008). Self-regulated learning with hypermedia: The role of prior domain knowledge. *Contemporary Educational Psychology, 33*(2), 270–298. https://doi.org/10.1016/j.cedpsych.2007.03.001

Moos, D. C., & Azevedo, R. (2009a). Self-efficacy and prior domain knowledge: To what extent does monitoring mediate their relationship with hypermedia learning? *Metacognition and Learning, 4*(3), 197. https://doi.org/10.1007/s11409-009-9045-5
Moos, D. C., & Azevedo, R. (2009b). Learning with computer-based learning environments: A literature review of computer self-efficacy. *Review of Educational Research, 79*(2), 576–600. https://doi.org/10.3102/0034654308326083
Ocak, G., & Yamaç, A. (2013). Examination of the relationships between fifth graders' self-regulated learning strategies, motivational beliefs, attitudes, and achievement. *Educational Sciences: Theory and Practice, 13*(1), 380–387.
Ohtani, K., & Hisasaka, T. (2018). Beyond intelligence: A meta-analytic review of the relationship among metacognition, intelligence, and academic performance. *Metacognition and Learning, 13*(2), 179–212. https://doi.org/10.1007/s11409-018-9183-8
Paans, C., Molenaar, I., Segers, E., & Verhoeven, L. (2019). Temporal variation in children's self-regulated hypermedia learning. *Computers in Human Behavior, 96*, 246–258. https://doi.org/10.1016/j.chb.2018.04.002
Pajares, F. (1996a). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543–578. https://doi.org/10.2307/1170653
Pajares, F. (1996b). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemporary Educational Psychology*, *21*(4), 325-344. http://doi.org/https://doi.org/10.1006/ceps.1996.0025.
Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory Into Practice, 41*(2), 116–125. https://doi.org/10.1207/s15430421tip4102_8
Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*, 422. https://doi.org/10.3389/fpsyg.2017.00422
Paquette, L., Grant, T., Zhang, Y., Biswas, G., & Baker, R. (2021). Using epistemic networks to analyze self-regulated learning in an open-ended problem-solving environment. In: A.R. Ruis, S.B. Lee (Eds). *Communications in computer and information science: Vol. 1312. Advances in quantitative ethnography. ICQE 2021* (pp. 185-201). Springer. https://doi.org/10.1007/978-3-030-67788-6_13.
Paquette, L., Ocumpaugh, J., Li, Z., Andres, J. M. A. L., & Baker, R. S. (2020). Who's learning? Using demographics in EDM research. *Journal of Educational Data Mining, 12*(3), 1–30. https://doi.org/10.5281/zenodo.4143612
Pavlik, P. I., Jr (2013). Mining the Dynamics of Student Utility and Strategy Use during Vocabulary Learning. *Journal of Educational Data Mining, 5*(1), 39–71. https://doi.org/10.5281/zenodo.3554609
Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Academic Press. https://doi.org/10.1016/B978-012109890-2/50043-3.
Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33. https://doi.org/10.1037/0022-0663.82.1.33
Pintrich, P. R., & Zusho, A. (2002). The development of academic self-regulation: The role of cognitive and motivational factors. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 249-284). Academic Press. http://doi.org/10.1016/B978-012750053-9/50012-7.
Ramos Salazar, L., & Hayward, S. L. (2018). An examination of college students' problem-solving self-efficacy, academic self-efficacy, motivation, test performance, and expected grade in introductory-level economics courses. *Decision Sciences Journal of Innovative Education, 16*(3), 217–240. https://doi.org/10.1111/dsji.12161
Roscoe, R. D., Segedy, J. R., Sulcer, B., Jeong, H., & Biswas, G. (2013). Shallow strategy development in a teachable agent environment designed to support self-regulated learning. *Computers & Education, 62*, 286–297. https://doi.org/10.1016/j.compedu.2012.11.008
Sabourin, J. L., Shores, L. R., Mott, B. W., & Lester, J. C. (2013). Understanding and predicting student self-regulated learning strategies in game-based learning environments. *International Journal of Artificial Intelligence in Education, 23*(1–4), 94–114. https://doi.org/10.1016/j.compedu.2012.11.008
Schenke, K., Lam, A. C., Conley, A. M., & Karabenick, S. A. (2015). Adolescents' help seeking in mathematics classrooms: Relations between achievement and perceived classroom environmental influences over one school year. *Contemporary Educational Psychology, 41*, 133–146. https://doi.org/10.1016/j.cedpsych.2015.01.003
Schnell, K., Ringeisen, T., Raufelder, D., & Rohrmann, S. (2015). The impact of adolescents' self-efficacy and self-regulated goal attainment processes on school performance — Do gender and test anxiety matter? *Learning and Individual Differences, 38*, 90–98. https://doi.org/10.1016/j.lindif.2014.12.008
Schunk, D. H., & Ertmer, P. A. (2000). Self-regulation and academic learning: Self-efficacy enhancing interventions. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 631–649). Academic Press. https://doi.org/10.1016/B978-012109890-2/50048-2.
Schunk, D. H., & Greene, J. A. (2017). Historical, contemporary, and future perspectives on self-regulated learning and performance. *Routledge*. https://doi.org/10.4324/9781315697048.ch1
Segedy, J. R. (2014). *Adaptive Scaffolds in Open-Ended Computer-Based Learning Environments* (Publication No. 3674132) [Doctoral dissertation, Vanderbilt University]. ProQuest Dissertations & Theses Global.
Segedy, J. R., Kinnebrew, J. S., & Biswas, G. (2015a). Using coherence analysis to characterize self-regulated learning behaviours in open-ended learning environments. *Journal of Learning Analytics, 2*(1), 13–48.
Segedy, J. R., Kinnebrew, J. S., & Biswas, G. (2015b). Coherence over time: Understanding day-to-day changes in students' open-ended problem solving behaviors. In Conati C., Heffernan N., Mitrovic A., Verdejo M. (Eds), *Lecture notes in computer science: Vol. 9112. Artificial Intelligence in Education. AIED 2015* (pp. 449-458). Springer. https://doi.org/10.1007/978-3-319-19773-9_45.
Shute, V. J., & Zapata-Rivera, D. (2012). Adaptive educational systems. In P. J. Durlach & A. M. Lesgold (Eds.), *Adaptive Technologies for Training and Education* (pp. 7-27). Cambridge University Press.
Snijders, T. A., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Sage Publishers.
Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance. *Perspectives on Psychological Science, 10*(2), 176–199. https://doi.org/10.1177/1745691615569000
Stram, D. O., & Lee, J. W. (1994). Variance components testing in the longitudinal mixed effects model. *Biometrics, 1171–1177*. https://doi.org/10.2307/2533455
Sweller, J. (2011). Cognitive load theory. *Psychology of Learning and Motivation*, *55*, 37-76. http://doi.org/https://doi.org/10.1016/B978-0-12-387691-1.00002-8.
Taub, M., Azevedo, R., Bouchet, F., & Khosravifar, B. (2014). Can the use of cognitive and metacognitive self-regulated learning strategies be predicted by learners' levels of prior knowledge in hypermedia-learning environments? *Computers in Human Behavior, 39*, 356–367. https://doi.org/10.1016/j.chb.2014.07.018
Taub, M., & Azevedo, R. (2019). How does prior knowledge influence eye fixations and sequences of cognitive and metacognitive SRL processes during learning with an Intelligent Tutoring System? *International Journal of Artificial Intelligence in Education, 29*(1), 1–28. https://doi.org/10.1007/s40593-018-0165-4
Tuan, H. L., Chin, C. C., & Shieh, S. H. (2005). The development of a questionnaire to measure students' motivation towards science learning. *International Journal of Science Education, 27*(6), 639–654. https://doi.org/10.1080/0950069042000323737
Üner, A., Mouratidis, A., & Kalender, O. (2020). Study efforts, learning strategies and test anxiety when striving for language competence: The role of utility value, self-efficacy, and reasons for learning English. *Educational Psychology, 40*(6), 1–19. https://doi.org/10.1080/01443410.2020.1736989
Velayutham, S., Aldridge, J. M., & Fraser, B. (2012). Gender differences in student motivation and self-regulation in science learning: A multi-group structural equation modeling analysis. *International Journal of Science and Mathematics Education, 10*(6), 1347–1368. https://doi.org/10.1007/s10763-012-9339-y
Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*(3), 265–310. https://doi.org/10.1016/0273-2297(92)90011-P
Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015
Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J. A., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology, 89*(3), 451–469. https://doi.org/10.1037/0022-0663.89.3.451
Wigfield, A., Hoa, L. W., & Klauda, S. L. (2008). The role of achievement values in the regulation of achievement behaviors. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 169-195). Lawrence Erlbaum Associates Publishers.
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., & Baker, R. S. (2013). The potentials of educational data mining for researching metacognition, motivation, and self-regulated learning. *Journal of Educational Data Mining, 5*(1), 1–8. https://doi.org/10.5281/zenodo.3554619
Winne, P. H., & Hadwin, A. F. (2008). The weave of motivation and self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and application*. Lawrence Erlbaum Associates Publishers.
Wolters, C. A., & Pintrich, P. R. (1998). Contextual differences in student motivation and self-regulated learning in mathematics, English, and social studies classrooms. *Instructional Science, 26*(1), 27–47. https://doi.org/10.1023/A:1003035929216
Zhang, Y., Paquette, L., Baker, R. S., Ocumpaugh, J., Bosch, N., Munshi, A., & Biswas, G. (2020). The relationship between confusion and metacognitive strategies in Betty's Brain. In C. Rensing & H. Drachsler (Eds.), *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 276-284). Association for Computing Machinery. http://doi.org/10.1145/3375462.3375518.
Zheng, J., Xing, W., & Zhu, G. (2019). Examining sequential patterns of self- and socially shared regulation of STEM learning in a CSCL environment. *Computers & Education, 136*, 34–48. https://doi.org/10.1016/j.compedu.2019.03.005
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. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press. https://doi.org/10.1016/B978-012109890-2/50031-7.
Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2
Zimmerman, B. J., & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal, 31*(4), 845–862. https://doi.org/10.3102/00028312031004845
Zimmerman, B. J., & Schunk, D. H. (2008). Motivation: An essential dimension of self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 1-30). Lawrence Erlbaum Associates Publishers. http://doi.org/10.4324/9780203831076.