Studying Student use of Self-Regulated Learning Tools in an Open-Ended Learning Environment

John S. Kinnebrew, Brian C. Gauch, James R. Segedy, and Gautam Biswas

Department of EECS and ISIS, Vanderbilt University, Nashville TN 37212, USA
{john.s.kinnebrew,brian.gauch,james.segedy,gautam.biswas}@vanderbilt.edu

Abstract. This paper discusses a design-based research study that we conducted in a middle school science classroom to test the effectiveness of SimSelf, an open-ended learning environment for science learning. In particular, we evaluated two tools intended to help students develop and practice the important regulatory processes of planning and monitoring. Findings showed that students who used the supporting tools as intended demonstrated effective learning of the science topic. Conversely, students who did not use the tools effectively generally achieved minimal success at their learning tasks. Analysis of these results provides a framework for redesigning the environment and highlights areas for additional scaffolding and guidance.

Keywords: Open-ended learning environments, Self-regulated learning, Learning environment design

1 Introduction

Cognitive scientists have established that the ability to regulate one's own learning is critical for developing effective learning practices. Self-regulated learning (SRL) is an active theory of learning that describes how learners set goals, create plans to achieve their goals, and continually monitor their progress in completing the plan. A realization of inadequate performance may lead to revising plans and goals. SRL is a multi-faceted construct that involves emotional and behavioral control, management of one's learning environment and cognitive resources, perseverance in the face of difficulties, and social interactions to achieve effective learning [12]. *Open-ended computer-based learning environments* (OELEs; [4]) provide students with opportunities to develop and practice their SRL processes; they provide students with a learning task and a set of tools for exploring, hypothesizing, and building solutions to authentic and complex problems.

In this paper, we present our recent work in developing *SimSelf*, an OELE for learning SRL strategies in the context of science learning tasks. *SimSelf* challenges students to learn science by creating models of scientific processes, and our goals in developing it are to research techniques for supporting students' understanding of specific SRL processes. The discussion in this paper focuses specifically on two tools in *SimSelf*:(1) a *planning interface* that support students' development and practice of strategies for *making plans*; and (2) a *monitoring interface* for students to analyze their own cognitive, metacognitive, affective, and motivational processes [1].

The results from this study were mixed. The analysis showed that some students utilized the planning and self-monitoring tools in a manner consistent with our expectations,and these students tended to have higher pre-post test gains and science modeling task performance. However, others did not use the tools as intended and often had correspondingly lower learning and task performance. A deeper evaluation and interpretation of these results provides us with a framework for redesigning some of the tools and interfaces, as well as identifying areas for additional scaffolding and guidance in the next version of *SimSelf*.

2 OELEs that support Self-Regulated Learning

Winne and Hadwin [10] have developed a conceptual framework called COPES (Conditions, Operations, Products, Evaluations and Standards) for modeling and analysis of SRL processes. They posit that learning occurs in four basic phases: (1) task definition, (2) goal-setting and planning, (3) use of studying tactics and methods, and (4) adaptations to metacognition. Winne and Hadwin further operationalize this SRL model by hypothesizing sets of information-processing operations that govern behaviors in each step. Thus, the model complements other more conceptual SRL models (e.g., [8]) by introducing an operational description of the processes underlying each phase of SRL.

A variety of computer-based learning environments have been developed that support the development of SRL processes. Winne and his colleagues developed *gStudy*, a system for assessing SRL processes as described by their COPES model [6]. Another example is the hypermedia environment called *MetaTutor*, which adapts the COPES model to detect, model, trace, and foster students' SRL about human body systems [1].

More recently, Chi and VanLehn [11]) have developed *Pyrenees*, a system that requires students to construct models using explicit strategies (*e.g.*, goal reduction for complex problem solving). In work with the *Betty's Brain* OELE [3, 5], students learned science concepts by building causal models of science phenomena, such as interdependence and balance in ecological systems. Studies with this system have shown that teaching students open-ended problem solving strategies can lead to their constructing higher-quality causal models [9]. However, the use of better strategies often has not translated to better learning of domain content. Given the importance of SRL processes in developing independent learners and the fact that there is not much conclusive evidence of how best to teach SRL to middle school students in OELEs, we have developed a system called *SimSelf* that explicitly focuses on SRL instruction.

3 SimSelf

SimSelf is an OELE that presents students with a complex array of tasks united in a single context: creating accurate models of scientific systems and processes. Students demonstrate their understanding by creating concept maps of the *structure* and *behavior* of the science topic under study. Both structure and behavior maps represent the system as a set of entities connected by directed links. The structure map captures the connections (*e.g.*, the hypothalamus *sends signals to* skeletal muscles) and hierarchical relationships (*e.g.*, the hypothalamus is *a part of* the brain) among entities. The behavior map captures causal relationships among entities (*e.g.*, skeletal muscle contractions *increase* friction in the body), and they may either describe increase (+) or decrease (-) relationships [5].

SimSelf includes tools for acquiring information, applying that information to map building, and assessing the quality of constructed maps. Students can acquire domain knowledge by reading hypertext resources.As students read, they need to identify structural and behavioral relationships and use this learned information to build their maps. Learners can assess their maps by having *SimSelf* automatically reason with the maps to complete *quizzes*. The software can then grade the generated answers and show how they were derived from the maps by highlighting the entities and links that were used to generate the answer. The system also includes a strategy guide that discusses the declarative, procedural, and conditional knowledge of important SRL strategies.

3.1 Planning and Monitoring Tools in SimSelf

To support students' SRL processes, *SimSelf* includes planning and monitoring tools that students may choose to use to help regulate their learning. Students can use the planning tool (Figure 1) to set a learning goal and then specify the steps they will take to achieve that goal. Students can add and delete activities and SRL processes to each step from lists on the right-hand side of the interface. For example, in Step 2 of the plan shown in Figure 1, the student has specified that she will read the science book page on skin contraction, evaluate the material, and check her learning as she constructs her structure map. Students can also mark steps as completed. Ideally, thoughtful planning and keeping track of those plans will help students monitor their progress more effectively, reflect on any difficulties they experience, and take action to overcome those difficulties.

The monitoring tool allows students to evaluate and record their use of various learning strategies and their cognitive, affective, and motivational states during learning. Students monitor by answering "yes or no" questions about themselves (or selecting a "not sure" option). Ideally, answering these questions will allow students to practice monitoring and make them more aware of their own internal states, an important step in effectively regulating their learning.

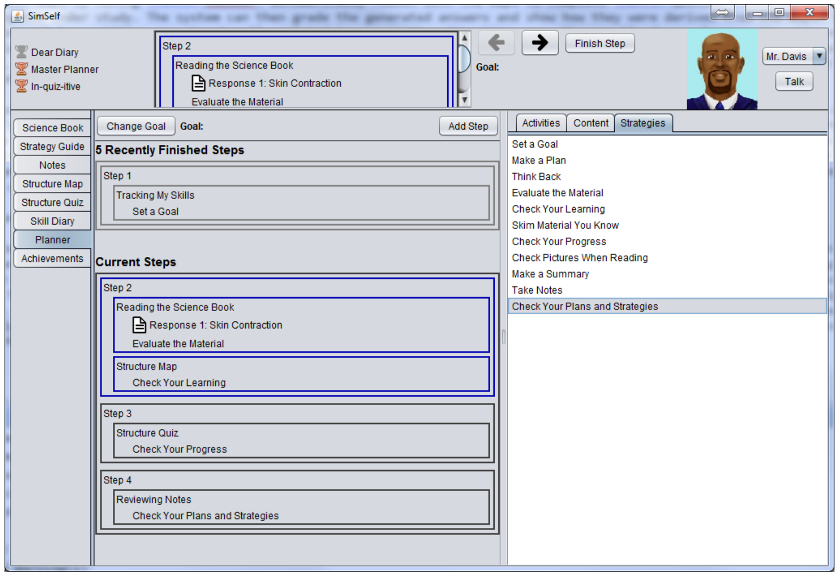

Fig. 1. The SimSelf planning interface.

4 Experimental Study

We report an initial study of students' use of the *SimSelf* planning and monitoring tools and the extent to which their use of the tools was productive for learning. To ensure that all students used the tools, we instructed them to: (1) use the planner at the beginning of each day to plan their approach; and (2) use the monitoring tool at the end of each day to reflect on their approach. Further, we told students that careful use of these tools throughout their work in *SimSelf* would facilitate their learning. We investigated the following questions:

1. Did students revise and/or update their plans regularly?
2. Did students execute the steps that they put in their plans?
3. Did students accurately monitor their own abilities, as reflected in their behaviors while using the system?
4. Did these behaviors relate to learning and task performance?

4.1 Participants, Topic Unit and Text Resources

Twenty-five 5$^{\text{th}}$-grade students from a middle Tennessee classroom used *SimSelf* to learn about human thermoregulation when exposed to cold temperatures. The expert structure map contained 10 concepts, 9 hierarchical links, and 3 connection links covering skin and the nervous, muscular, and circulatory systems. The expert behavior map contained 10 concepts and 11 causal links, and students constructed this map in three parts: cold detection (cold temperatures, heat loss, body temperature, cold detection, hypothalamus response); vasoconstriction (blood vessel constriction, blood flow to the skin, heat loss); and shivering (skeletal muscle contractions, friction in the body, heat in the body).

The thermoregulation resources were organized into two introductory pages discussing the nervous system and homeostasis, one page discussing the structure, behavior, and function method of understanding scientific systems[2], as well as the structure of the thermoregulatory system. These pages were followed by one page each discussing cold detection, skin contraction, vasoconstriction, and shivering. Additionally dictionary pages defined the main concepts. The text was 16 pages (2,682 words) with a Flesch–Kincaid reading grade level of 8.1.

4.2 Learning Assessments

Learning was assessed using a pre- and post-test design. Each test consisted of 7 causal reasoning questions and 10 science content questions. The causal reasoning questions presented students with an abstract causal map and asked students to reason with the map to answer questions like "If concept A increases, what would happen to concept B?" Students were awarded one point for each correct answer. Science content questions included 6 multiple choice questions and 4 short answer questions. The multiple choice questions, each with four choices, tested students' understanding of primary concepts and simple relations among them. One point was awarded for each correct answer. Short answer questions asked students to consider a given scenario (*e.g.*, alcohol consumption) and explain its causal impact on thermoregulation. These questions were coded by identifying the causal relationships in learners' answers, which were scored by comparing them to the causal relationships in the expert map. One point was awarded for each causal relationship in the student's answer that was the same as or closely related to a relation specified in the expert map. Two coders independently scored five of the pre- and post-tests with over 85% inter-rater reliability, at which point one of the coders individually coded the remaining answers. The maximum combined score for all science questions was 17.

4.3 Log File Analysis

SimSelf automatically generates *event logs* that capture every time-stamped *action* taken by the student (*e.g.*, deleting a concept) and interface view that was displayed while the system was running. We used these log files to calculate measures of students' behaviors and task performance while using the system. The *map score* for a structure or behavior map is calculated as the number of correct links (*i.e.*, links that appear in the expert map) minus the number of incorrect links in the map. A student's *best map score* for a particular topic ua the highest map score they attained while working on that topic. [1]. Our behavior analyses focused on students' use of the planning and monitoring interfaces. We calculated the following measures:

[1] We use best map scores because students sometimes delete their entire map.

1. *Planning Activities*: when and how often students set goals, changed their plan, and marked steps complete.
2. *Plan Adherence*: the proportion of time students' spent performing activities that were specified in their plans.
3. *Content Evaluation Proficiency*: content evaluation is an SRL process in which learners evaluate the utility of information. During each day that learners used *SimSelf*, they were expected to model a particular aspect of thermoregulation. By tracking the resource pages students viewed, we determined whether they were viewing potentially *relevant* material. Content evaluation proficiency is the percentage of a student's reading time spent viewing potentially relevant material for their assigned goal.
4. *Content Evaluation Monitoring*: students' ability to assess their own content evaluation skills. When students accessed the monitoring tool, they were asked the question *"have I been reading things that are related to my current goal?"* A student's content evaluation monitoring score is the difference between the proportion of times they responded with "yes" and the proportion of time they spent reading relevant material. A score closer to 0 is interpreted as more accurate monitoring.

4.4 Procedure

Study duration was 10 50-minute class periods. During the first two periods of the study, one author led classroom lessons introducing students to the modeling languages and presenting an overview of SRL. During the third period, students were introduced to *SimSelf* and its features, and they were allowed to practice on the system. During the fourth period, students completed the pretest.

During period 5 they worked on the thermoregulation structure map. During period 6, one author led a classroom lesson on thermoregulation. He explained how the body detects cold temperatures and how it responds by triggering shivering and vasoconstriction responses. Students then spent three class periods working on the three sections of the behavior map (as described previously). During periods 8 and 9, students started with the correct map from the previous period. During period 10, students completed the post-test.

During the classroom intervention, we observed that very few students completed the structure map on its allocated day. Therefore, if a student completed a behavior unit early, they were allowed to continue working on their structure map. Because students worked for different amounts of time on the structure map, the focus of the data analyses in this paper is on students' behaviors while building the thermoregulation behavior map.

5 Results

Table 1 summarizes students' pre- and post-test scores. Overall, students exhibited moderate gains on science content (d = 1.02), suggesting that the intervention helped them learn to recognize and reason with concepts and relations describing thermoregulation. Students did not show learning gains on causal reasoning questions. This may be explained by the decision to administer the pre-test after the lesson on reasoning with structure and behavior maps.

Measure	**Max.**	**Pretest**	**Posttest**	t	p	**Cohen's d**
Causal reasoning	7	4.20 (1.65)	4.16 (1.67)	0.123	0.903	0.025
Science content	17	3.32 (1.26)	5.68 (2.96)	4.264	0.001	1.024

Table 1. Means (and standard deviations) of assessment test scores.

To address our first research question, we created a heat map representation of students' planning activities over the course of their time in each unit (Figure 2). The results show that students primarily worked on their plans during the first 10–20% of their time on the system. Further analysis indicated that most planning activities after 20% of their time involved marking existing plan steps complete. In terms of our first research question, this suggests that students did keep track of their plans as they worked on *SimSelf* but did not revise them.

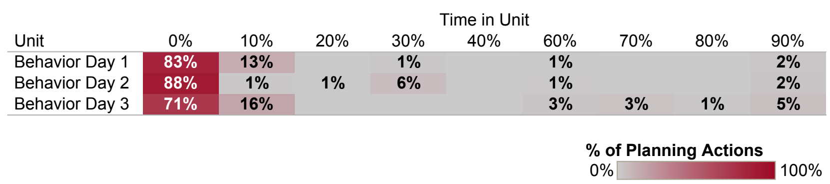

Fig. 2. Proportions of students' planning behaviors over time

Table 2 shows the means (and standard deviations) of the percentage of time students spent on activities listed in their plans (Planned Activities) and reading material specified in their plans (Planned Reading). Results show that students spent a majority of their time performing planned activities. Conversely, students spent only a small percentage of their reading time on planned reading. In terms of our second research question, one possible explanation for these behaviors is that students were better at planning high-level strategies (*e.g.*, iterating between reading and map building) than they were at planning specific details of their approaches (*e.g.*, the pages they would read).

	Day 1	**Day 2**	**Day 3**
Planned Activities	55.0% (30.5%)	66.7% (30.5%)	65.0% (30.2%)
Planned Reading	11.9% (27.3%)	9.3% (19.9%)	14.9% (25.0%)

Table 2. Proportion of time performing activities in their plans

Although students exhibited low proportions of planned reading, they may have been better at dynamically evaluating the material they were reading. Table 3 shows the means (and standard deviations) of students' content evaluation proficiency and monitoring scores. The results show that students spent an overwhelming majority of their time (> 89%) reading relevant pages on all three behavior map days. These numbers are higher than the percentage of pages in the resources that were relevant (row 3) on each of the three days, suggesting students' proficiency was better than chance.

	Day 1	**Day 2**	**Day 3**
Proficiency	91.4% (15.1%)	89.8% (15.9%)	93.8% (12.3%)
Monitoring	7.5% (25.3%)	6.5% (36.7%)	11.8% (34.3%)
% Rel. Pages	73.3%	60.0%	73.3%

Table 3. Means (and standard deviations) of content evaluation scores. The bottom row displays the percentage of relevant pages for each day.

In terms of our third research question regarding students' ability to monitor their own content evaluation, the results show that students, on average, slightly underestimated their proficiency. Students' self-assessments of their content evaluation were, on average, 7.5%, 6.5%, and 11.8% less than their actual proficiency scores. This suggests that students were mostly accurate in their self-judgment abilities. If this pattern is indicative of their ability to monitor other aspects of their learning behaviors, it suggests that this population of 5^{th} grade students were quite capable of self-assessing this aspect of their learning behavior.

Our final set of analyses investigated our fourth research question regarding whether the planning and monitoring assessments correlate with learning gains and normalized best map scores. We calculated bi-variate correlations between these measures and students' planning and content evaluation scores, each averaged over the three days. We found significant correlations between best map scores and: (1) average planned activity scores ($r = 0.472$, $p = 0.017$); and (2) average planned reading scores ($r = 0.434$, $p = 0.030$). However, the analysis failed to identify a relationship between map scores and content evaluation proficiency or monitoring. In other words, students who spent a larger proportion of their time engaging in planned activities and planned reading tended to construct more accurate behavior maps, but the same was not true of students who spent more time reading relevant material.

Moreover, students' science learning gains were also moderately (but not significantly) correlated with planned activity scores ($r = 0.340$, $p = 0.096$) and planned reading scores ($r = 0.329$, $p = 0.108$). However, no relationship was found between learning and content evaluation proficiency or monitoring. One possible explanation is that the more engaged students were more willing to exert the effort necessary to create a meaningful plan, stay with it, and carefully analyze their modeling performance. As a result, they achieved greater success and may have learned more of the material.

6 Discussion and Conclusions

The study showed mixed results: students used some aspects of the planning and monitoring interfaces productively, but they did not take advantage of other aspects that would have demonstrated higher SRL proficiency. Reflection on the results reported in the last section provides us with some useful lessons that will motivate the next iteration of SimSelf. Our first lesson pertains to the planning interface. Some students were unsure of how to best use the planner. They were reluctant to mark steps on their plans complete because they expected to repeat these steps in the near future. As a result, we believe it will be more useful to move the focus from *exact plan steps* to *conditional knowledge* that underlies successful planning. We are currently re-designing the planning interface so students can specify the conditions under which they may invoke specific SRL processes during learning. In this interface, students will be responsible for specifying when, for example, it is appropriate to employ content evaluation. We believe that this will help students link generic SRL processes (which they seem to understand) to specific learning tasks (for which they currently do not show proficiency). It will also provide a better framework for feedback that helps students operationalize SRL processes using appropriate conditional knowledge, contextualized by their concrete learning activity.

A second lesson relates to the apparent lack of a relationship between content evaluation and performance or learning in this study. A reasonable expectation is that students who can identify relevant information are more likely to learn and succeed in *SimSelf*, but we could not establish this connection with our results. We realize that identifying important content is not, by itself, sufficient for effective understanding, which also requires the ability to interpret the content under study. In future work, we plan to adopt the approach of [9] and provide students with opportunities to practice related skills and strategies important for success in interpreting the information in *SimSelf*.

The third lesson relates to the inability to identify a relationship between content evaluation monitoring and performance or learning. This may be due to shortcomings in the monitoring tool. Monitoring during problem solving is valuable only when it can be used to identify and correct problems in one's current approach. However, the monitoring tool did not include supports that helped students revise their approaches when faced with challenging circumstances. In future work, we will augment the *SimSelf* monitoring interface and use pedagogical agents to suggest alternate strategies for solving problems.

In summary, we provided students with tools that could be used for regulating learning, but we did not support their learning how to use these tools. In the next version of the system, we will incorporate a focus on both procedural and conditional knowledge in relation to how to use these tools effectively to support learning. We will also include scaffolds designed to help students understand how to use the result of their monitoring processes to adapt their problem solving approaches as they continue working towards completing their tasks. Several researchers have shown that students struggle to succeed in complex OELEs when support and scaffolding is not adaptive to students (*e.g.*, [4, 7]). We are in the process of developing these levels of support, which along with the planned changes to the *SimSelf* interfaces, will bring us closer to our goal of helping middle school students develop SRL processes.

7 Acknowledgments

This work has been supported by the Institute of Education Sciences CASL Grant #R305A120186. We would also like to acknowledge the time and expertise contributed by our collaborators at North Carolina State University led by Prof. Roger Azevedo.

References

1. Azevedo, R., Johnson, A., Chauncey, A., Burkett, C.: Self-regulated learning with metatutor: Advancing the science of learning with metacognitive tools. In: Khine, M., Saleh, I. (eds.) New Science of Learning. pp. 225–247. Springer (2010)
2. Goel, A., Rugaber, S., Vattam, S.: Structure, behavior & function of complex systems: The sbf modeling language. AI for Engineering Design, Analysis and Manufacturing 23, 23–35 (2009)
3. Kinnebrew, J., Segedy, J., Biswas, G.: Analyzing the temporal evolution of students behaviors in open-ended learning environments. Metacognition and learning pp. 1–29 (2014)
4. Land, S.: Cognitive requirements for learning with open-ended learning environments. Educational Technology Research and Development 48(3), 61–78 (2000)
5. Leelawong, K., Biswas, G.: Designing learning by teaching agents: The Betty's Brain system. International Journal of Artificial Intelligence in Education 18(3), 181–208 (2008)
6. Perry, N., Winne, P.: Learning from learning kits: gStudy traces of students self-regulated engagements with computerized content. Educational Psychology Review 18(3), 211–228 (2006)
7. Roll, I., McLaren, B., Koedinger, K.: Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. Learning and Instruction 21(2), 267–280 (2011)
8. Schraw, G., Crippen, K., Hartley, K.: Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. Research in Science Education 36(1), 111–139 (2006)
9. Segedy, J., Biswas, G., Blackstock, E., Jenkins, A.: Guided skill practice as an adaptive scaffolding strategy in open-ended learning environments. In: Proceedings of the 16th International Conference on Artificial Intelligence in Education. pp. 532–541. Springer (2013)
10. Winne, P., Hadwin, A.: The weave of motivation and self-regulated learning. In: Schunk, D., Zimmerman, B. (eds.) Motivation and self-regulated learning: Theory, research, and applications, pp. 297–314. Taylor & Francis, NY (2008)
11. Zhang, L., VanLehn, K., Girard, S., Burleson, W., Chavez-Echeagaray, M.E., Gonzalez-Sanchez, J., Hidalgo-Pontet, Y.: Evaluation of a meta-tutor for constructing models of dynamic systems. Computers & Education 75, 196–217 (2014)
12. Zimmerman, B.: Self-regulating academic learning and achievement: The emergence of a social cognitive perspective. Educational Psychology Review 2(2), 173–201 (1990)