Supporting Students' Self-Regulated Learning with an Open Learner Model in a Linear Equation Tutor

Yanjin Long and Vincent Aleven

Human Computer Interaction Institute, Carnegie Mellon University,
5000 Forbes Avenue, Pittsburgh, PA 15213, USA
{ylong,aleven}@cs.cmu.edu

Abstract. Self-assessment and study choice are two important metacognitive processes involved in Self-Regulated Learning. Yet not much empirical work has been conducted in ITSs to investigate how we can best support these two processes and improve students' learning outcomes. The present work redesigned an Open Learner Model (OLM) with three features aimed at supporting self-assessment (self-assessment prompts, delaying the update of the skill bars and progress information on the problem type level). We also added a problem selection feature. A 2x2 experiment with 62 7th graders using variations of an ITS for linear equation solving found that students who had access to the OLM performed significantly better on the post-test. To the best of our knowledge, the study is the first experimental study that shows an OLM enhances students' learning outcomes with an ITS. It also helps establish that self-assessment has key influence on student learning of problem solving tasks.

Keywords: Self-regulated learning, open learner model, self-assessment, study choice, intelligent tutoring system, classroom evaluation.

1 Introduction

Theories of Self-Regulated Learning (SRL) emphasize that students are active learners [13]. Different metacognitive processes are involved in SRL, such as goal setting, self-assessment, help-seeking, self-monitoring, study choice, etc. Two common metacognitive processes are self-assessment and study choice. Self-assessment refers to students' ability to evaluate how well they are learning/have learned. Study choice means that students make their own decisions with respect to the learning materials they study. More accurate self-assessment can lead to better study choice, which can further result in more efficient and effective learning [13]. Studies conducted with memory tasks and reading comprehension have found some ways to scaffold students' self-assessment and study choice, such as generating delayed key words [5]. Nevertheless, not much such work has been conducted with problem solving tasks, which is an area that Intelligent Tutoring Systems (ITSs) frequently focus on. The mechanism of self-assessing for solving math problems could be significantly different from memory task and reading comprehension.

K. Yacef et al. (Eds.): AIED 2013, LNAI 7926, pp. 219–228, 2013.

ITS researchers have been interested in the potential of Open Learner Models (OLM) to prompt students' reflection and metacognition [3]. Many ITSs have a learner model that intelligently tracks students' learning progress or their skill mastery. An OLM affords students access to part/all of progress information, often in different formats, which may help them reflect on what they know well and not so well. Bull and colleagues [4] found that first year college students were interested in viewing their misconceptions in an OLM, and believed that viewing such information could help them better assess their learning and allocate efforts. Hartley and Mitrovic [6] compared students' learning gains when with or without access to an inspectable OLM, but found no significant effect on the learning gains due to the OLM [6]. In our own prior work, we conducted surveys and interviews with experienced Cognitive Tutor users and found that they inspect the tutor's OLM (the Skillometer) quite frequently but do not actively use it to help them reflect or self-assess [8]. Similar work has also been conducted in the field of adaptive hypermedia. Brusilovsky et al. [2] found that with adaptive navigation support in QuizGuide (an adaptive system provides students self-assessment quizzes), students' participation was increased in the system, as well as their final academic performance. The adaptive navigation support has similar features as the OLMs, as it highlights to the students the important topics and topics that need more practice. Thus, as Bull et al. [3] point out, more empirical studies are needed to investigate how we can design an OLM to effectively facilitate students' metacognition, such as self-assessment and study choice. Moreover, it is also worth investigating to what extent access to an OLM and particular features of OLMs can significantly increase students' learning gains.

There has been limited prior work on study choice within ITS; typically, the ITS is responsible for selecting problems for the students. Mitrovic and Martin [9] found that in an ITS for SQL, lower-performing students learned in a "faded" condition in which they went from system-selected problems to student-selected problems. However, this study did not establish a statistically significant difference with other problem selection methods (fully system-selected or fully student-selected) [9]. The effect of problem selection on students' learning outcomes is still open for further investigation.

In the current work, we redesigned the Skillometer (OLM) of an ITS for linear equation solving so that it facilitates students' self-assessment. Specifically, we designed and implemented three new features for the Skillometer to support a brief self-assessment phase at the end of each tutor problem: self-assessment prompts, delaying the update of the skill bars (so that the updating of the skill bars can function as feedback on students' self-assessment) and showing students' progress on the problem type level in addition to on the skill level (to give students an overview of their progress in the tutor). We also implemented a problem selection feature in the tutor that lets students select their next problem.

We hypothesize that 1) having access to the redesigned OLM can enhance students' learning outcomes and self-assessment accuracy; 2) letting students select their own problems in the tutor could afford them opportunities to apply the results of their self-assessment and improve their learning outcomes even further. We conducted a 2x2 classroom experiment with 62 7^{th} graders with the linear equation tutor to investigate the hypotheses.

2 Methods

2.1 Linear Equation Tutor and the Open Learner Model

We investigate the relationship between OLM, self-assessment and study choice within an ITS for linear equations. This tutor is an example-tracing tutor built using the Cognitive Tutor Authoring Tools [1, 11]. It was first designed and implemented by Maaike Waalkens [11] and has been used in two prior studies with around 150 students from grades 7 and 8. The tutor teaches five types of linear equations of varying difficulty levels (see Table 1). Figure 1 shows the main interface of the tutor: in addition to solving the equations, students need to self-explain each main step. The tutor provides step-by-step guidance for each problem. It also applies knowledge tracing and mastery learning to adaptively select problems for each student, so as to make sure the student reaches mastery on all targeted skills.

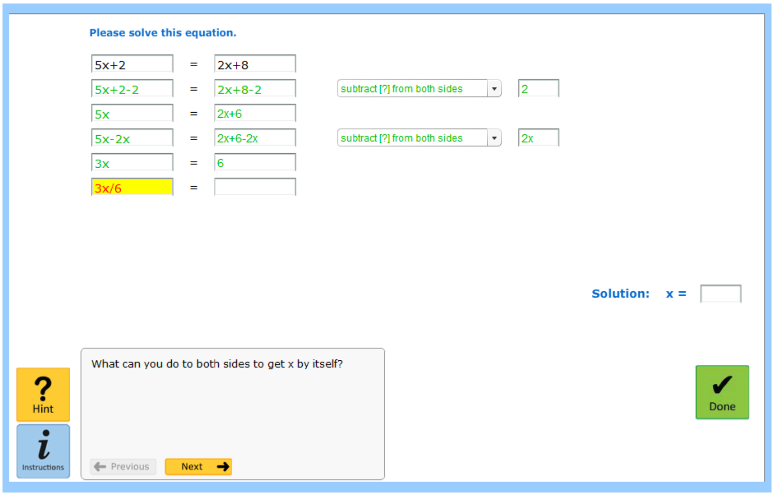

Fig. 1. The interface of the linear equation tutor

Table 1. Five types of equations in the linear equation tutor

Equations	Example	Level
One Step	x+5 = 7	Level 1
Two Steps	2x+1=7	Level 2
Multiple Steps	3x+1=x+5	Level 3
Parentheses	2(x+1)=8	Level 4
Parentheses, more difficult	2(x+1)+1=5	Level 5

As discussed in the introduction, we redesigned the OLM so as to support students' self-assessment and reflection at the end of each problem. We used a user-centered design approach to redesign the OLM. We started with building paper and digital prototypes for the OLM based on literature review. To refine the initial prototypes, we conducted think-aloud sessions with these prototypes in a local middle school with 7 students. Based on the findings from the think-alouds, we finalized the design as shown in Figure 2 and Figure 3. The five types of equations were categorized from level 1 to level 5 based on the skills involved, in order to more systematically reflect students' learning progress in the OLM. We implemented two views of the OLM with three new features: self-assessment prompts, delaying the update of the skill bars and showing progress on the problem type level. We also implemented a problem selection feature in the tutor to let students select their next problem.

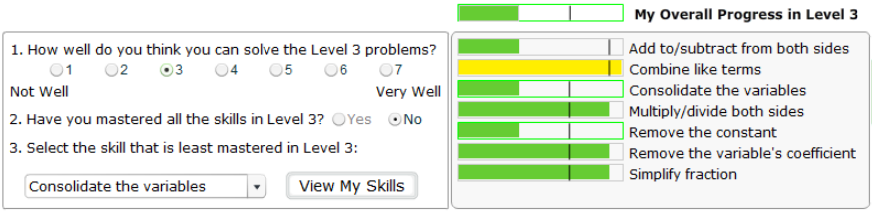

Fig. 2. View-1 of the OLM

Self-Assessment Prompts. View-1 of the OLM is initially hidden on the tutor interface but is revealed after the student finishes the problem. After students complete each problem, three self-assessment prompts are shown one by one (see Figure 2). (The level and skill bars on the right in Figure 2 are not displayed yet at this point in time, so that students answer the self-assessment questions unaided by the skill bars.) Students are asked to rate how well they think they can solve the problems in the current level on a scale from 1 to 7, then answer whether in their own assessment they have mastered the skills in the current level, and finally select the skill that they think is least mastered at this time. After that, the "View My Skills" button appears.

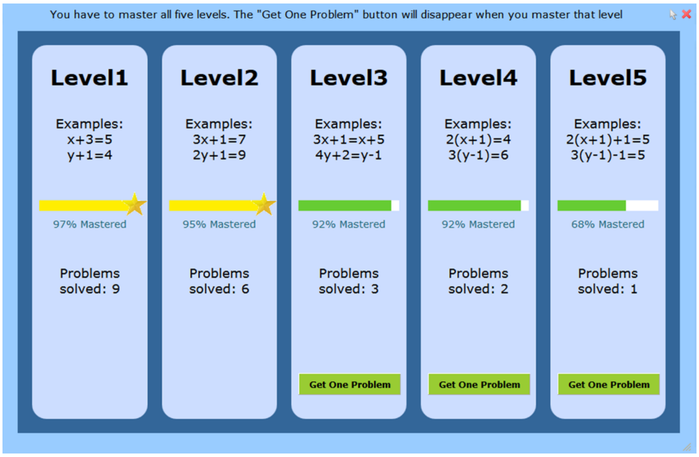

Fig. 3. View-2 of the OLM

Delaying the Update of the Skill Bars. Once students click the "View My Skills" button, the level and skill bars (on the right of Figure 2) are shown and start updating after 1 second (i.e., they move to their new positions, based on the student's performance on the problem they just completed). The updating of the bars serves as feedback on students' responses to the self-assessment prompts. The black vertical lines allow for a before/after comparison.

Showing Progress on the Problem Type Level. Figure 3 shows View-2 of the OLM, which is displayed to students in between problems (when they click the done button after the skill bars have finished updating). View-2 shows a summary of their progress with respect to each level as well as how many problems they have solved at that level.

Selecting the Next Problem. Further, on View-2, students can select the level they want to work on next by clicking the "Get One Problem" button for the preferred level. If a level is fully mastered, the "Get One Problem" button is hidden, so students can only select levels that contain unmastered skills. To complete the tutor they must master all levels.

2.2 Experimental Design, Participants, Procedure and Measurements

We conducted a 2x2 experiment with independent factors OLM (whether or not both views of the OLM are shown to the students) and PS (whether or not students could select their next problem from an unfinished level) with 62 7th grade students from one teacher's three classes at a local public middle school in Pittsburgh. The participants were randomly assigned to one of the four conditions. The OLM+PS condition used the interfaces we introduced in 2.1. The other three conditions used versions of the interfaces that were modified to match the manipulation. Specifically, for the OLM+noPS condition, View-1 of the OLM was unchanged, but View-2 was revised to have only a single "Get One Problem" button, rather than one for each level. Students in this condition were given problems from level 1 to 5 sequentially (they needed to finish level 1 first and then get problems from level 2, and so on). For the noOLM+PS condition, View-1 was not shown to the students. On View-2, all progress information was hidden (i.e., the progress bars and the number of problems completed for each level), but students could freely select their next problem from unmastered levels. Lastly, for the noOLM+noPS condition, View-1 was also not shown. For View-2, the progress information was hidden and there was only one single "Get One Problem" button.

The four conditions followed the same procedure. They all completed a paper pretest on the same day for around 25 minutes, and started to work with the tutor in their computer lab from the next day for five consecutive days. On each day, all students worked on the tutor for one class period of 41 minutes. If a student finished early (in less than 5 periods), they were directed to work in a Geometry unit. After the five days, all conditions again completed an immediate paper post-test on the same day in one class period.

The pre- and post-tests were in similar format and measured students' knowledge of solving linear equations. We created two equivalent test forms and administered them in counterbalanced orders. There were two types of test items on both tests: procedural and conceptual items. Procedural items were the same five types of equations students had practiced in the tutor. Conceptual items were True/False questions measuring the knowledge and understanding of the key concepts involved in equations. We also measured students' self-assessment accuracy for the procedural items on both tests. Students were asked to rate from 1 to 7 regarding how well they think they can solve each equation before they actually solved it. Formula 1 calculates the absolute accuracy of students' self-assessment [10], where "N" represents the number of tasks, "c" stands for students' confidence ratings on their ability to finish the task while "p" represents their actual performance on that task.

$$\text{Absolute Accuracy Index} = \frac{1}{N}\sum_{i=1}^{N}(c_i - p_i)^2 \tag{1}$$

Besides the pre- and post-tests, we analyzed tutor log data to determine if there were differences between the conditions in students' learning behaviors in the tutor.

3 Results

56 students finished all five levels (reached mastery) after 5 class periods. We analyzed the 56 students' pre-test and post-test performance, tutor log data and their self-assessment data. We report the p-values and effect sizes (partial η^2) for the main effects and interactions. An effect size partial η^2 of .01 corresponds to a small effect, .06 to a medium effect, and .14 to a large effect (Cohen's guidelines for effect sizes).

Learning Effects of the Linear Equation Tutor. There were 7 procedural items and 12 conceptual items on both tests. The procedural items were graded from 0 to 1, with partial credit given where appropriate. Cronbach's Alpha for the 7 procedural items on the pre-test is .794, and .669 on the post-test. For the conceptual items, the Cronbach's Alphas are .626 and .672 for pre- and post-test respectively.

Table 2. Means and SDs for the test performance for all four conditions

Conditions	Pre-Test (Procedural)	Post-Test (Procedural)	Pre-Test (Conceptual)	Post-Test (Conceptual)
OLM+PS	.439 (.263)	.711 (.230)	.483 (.215)	.515 (.188)
OLM+noPS	.555 (.347)	.684 (.222)	.472 (.166)	.541 (.230)
noOLM+PS	.358 (.201)	.625 (.237)	.391 (.216)	.357 (.195)
noOLM+noPS	.490 (.204)	.634 (.290)	.436 (.164)	.462 (.202)

A 1-way ANOVA shows that there were no significant differences between the conditions on the pre-test. To examine the learning gains from pre- to post-test, we ran repeated measures ANOVAs (with OLM and PS as independent variables) on procedural items, conceptual items and the sum of the two (the overall test score). The results reveal that the conditions together improved significantly from pre- to post-test on the test as a whole ($F(1, 52) = 13.927$, $p = .000$, $\eta^2 = .211$) and on the procedural items separately ($F(1. 52) = 35.239$, $p = .000$, $\eta^2 = .404$), both with effect sizes considered to be very large. No significant improvement on conceptual items was found.

Effects of Open Learner Model (OLM). We also ran ANOVAs (with OLM and PS as independent variables) for the post-test results. There was a significant main effect of OLM on the overall test scores ($F(3, 52) = 4.903$, $p = .031$, $\eta^2 = .078$), as well as on the conceptual items ($F(3, 52) = 5.212$, $p = .026$, $\eta^2 = .082$). No significant main effect was found for the procedural items. In short, the two OLM conditions performed better on the post-test than the two groups who did not have access to the OLM. We then looked at process measures from the tutor log data to determine whether having access to the OLM significantly influenced students' behaviors while learning with the tutor. The process measures shown in Table 3 are commonly used in Cognitive Tutor studies [7]. As shown in Table 3, the two OLM conditions made fewer incorrect attempts, requested fewer hints and had a lower average assistance score ((hints + incorrect attempts) / total steps). ANOVAs (with OLM and PS as independent variables) show that there was a marginally significant main effect of OLM on incorrect attempts ($F(3, 52) = 3.608$, $p = .062$, $\eta^2 = .059$), and a significant main effect of OLM on average assistance score ($F(3, 52) = 3.292$, $p = .009$, $\eta^2 = .116$). There was no significant main effect of OLM on the number of hints.

Table 3. Means and SDs of process measures for all four conditions

	OLM+PS	OLM+noPS	noOLM+PS	noOLM+noPS
Total number of problems	32.80 (9.15)	36.93 (11.50)	34.23 (6.51)	39.31 (9.30)
Incorrect attempts per step	.248 (.180)	.261 (.164)	.337 (.256)	.364 (.182)
Hints per step	.157 (.138)	.190 (.178)	.221 (.197)	.268 (.433)
Average assistance score	.260 (.178)	.268 (.166)	.321 (.123)	.532 (.368)

Effects of Problem Selection (PS). ANOVAs (with OLM and PS as independent variables) found no significant main effect of PS on the overall post-test score or on the two categories of post-test items separately. For log data, the students in the PS conditions made fewer incorrect attempts, requested fewer hints, had a lower average assistance score, and needed fewer problems to reach mastery in the tutor. The effect of PS was marginally significant on the average assistance score ($F(3, 52) = 3.292$, $p = .075$, $\eta^2 = .056$), but was not significant for the other dependent measures mentioned above.

Effects of the Interaction between OLM and PS. We did not find any significant interactions between OLM and PS on the post-test results. From the log data, we found an interaction that was on the borderline of significance for the average assistance score (ANOVA, $F(3, 52) = 2.804$, $p = .100$, $\eta^2 = .049$). Specifically, when students did *not* have access to the OLM, control over problem selection led to a lower assistance score, whereas with access to the OLM, their assistance score was the same regardless of whether they had control over problem selection.

Self-Assessment (SA) Accuracy. We also evaluated students' self-assessment accuracy. Figure 4 shows the frequencies of each self-assessment score (on the left) as well as how students' actual test performance relates to their self-assessment score (on the right). For both pre- and post-tests, the actual test scores increase as the self-assessment scores increase. We also compared students' self-assessment scores on the pre- and post-tests. A repeated measures ANOVA reveals that students' self-assessment scores increased significantly from pre- to post-test ($F(1, 52) = 13.078$, $p = .001$, $\eta^2 = .201$; pre-test Mean = 4.706, post-test Mean = 5.270). No significant differences were found between the conditions.

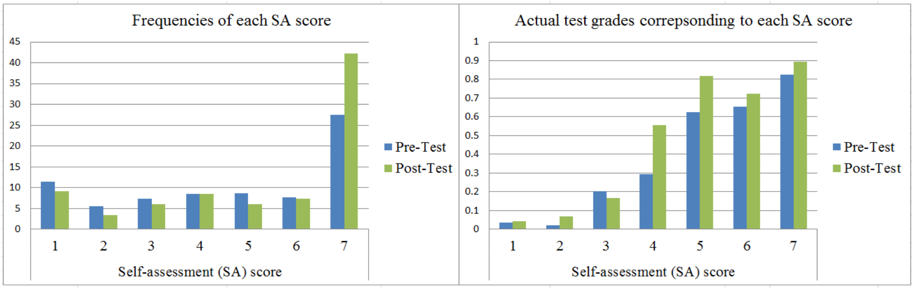

Fig. 4. The frequencies of different SA scores and the distribution of the test performance

Table 4 shows students' absolute accuracy of self-assessment (the lower the index, the better students' self-assessment). An absolute accuracy index of .14 means that a student answers a question correctly and s/he is 62.6% confident (according to Schraw [10], 50% confident is considered to be moderately accurate). Therefore, as shown in Table 4, the students had moderate to high accuracy of self-assessment. No significant differences were found among the conditions.

Table 4. The absolute self-assessment (SA) accuracy for different conditions

	OLM+PS	OLM+noPS	noOLM+PS	noOLM+noPS
Pre-test SA accuracy	.186 (.163)	.146 (.145)	.147 (.128)	.143 (.146)
Post-test SA accuracy	.127 (.115)	.127 (.088)	.166 (.077)	.106 (.084)

4 Discussion, Conclusion and Future Work

We conducted a controlled classroom experiment to investigate the effectiveness of having access to an OLM and having problem selection in an ITS, an area where not much empirical work has been conducted. Firstly, the pre- and post-test results reveal that students' knowledge of solving linear equations improved significantly, with large effect sizes on both the procedural problems and whole test, affirming the effectiveness of the tutor. Secondly, having access to an OLM resulted in better performance on the post-test. OLMs are a common feature in ITS. Although much effort has been put into the design and evaluation of the OLMs, and it has often been theorized that OLMs enhance the effectiveness of ITSs, we know of no prior experimental studies that had demonstrated an OLM significantly enhances student learning compared to a noOLM condition. The advantage of our OLM conditions suggests that the reflective self-assessment activities scaffolded by the OLM can significantly enhance students' learning outcomes, similar to the paper-based support in White and Frederiksen [12]. Specifically, students were prompted to reflect and self-assess on their learning status after each problem, with the display and updating of the OLM functioning as implicit feedback on their self-assessment. In this way, students might have been reminded of the errors and difficulties they had while solving each problem, as well as how they had corrected/resolved them. Such reflective process could enhance their understanding and help them learn from their errors. In addition, being exposed to their progress could also keep the students alerted the whole time. They would be more careful and motivated to stay focused on the learning. As revealed by the log data, the students with the OLM needed significantly less assistance from the system and made marginally significantly fewer incorrect attempts.

Thirdly, we did not find any significant main effect of PS on post-test results. In the log data, we only found that the students in the PS conditions had a marginally significant lower assistance score, suggesting that having control over problem selection leads to a somewhat smoother experience when solving problems. We also found the interaction between OLM and PS was on the borderline of significance for the average assistance score. When students had to select their own problems, they might be spurred to be more careful and active in their learning process, as evidenced by the lower assistance score. However, the fact that no significant results were found on the post-test suggests that more studies are still needed to investigate whether and how problem selection can enhance students' learning outcome in ITSs.

In regard to self-assessment, we found that students' self-assessment scores (confidence ratings) increased significantly from pre- to post-test, with a large effect size. Another interesting finding is that the participating students generally had moderate to high accuracy of self-assessment on the procedural problems, which is different from what have been observed in lots of prior work focusing on memory tasks and reading comprehension [5]. One explanation could be that the superficial features of equations correspond well with their difficulty levels, i.e. equations with more terms (or with parentheses) normally are more difficult. Consequently, it might be easier for the students to make accurate self-assessment on these questions. However, the mechanisms of self-assessment for different learning tasks need to be clarified in future research. Regardless, the increased self-assessment, especially given that it was accurate, should be viewed as positive result in its own right; arguably, learning is not truly robust if not accompanied by accurate self-assessment.

In sum, the present study shows that having an OLM while learning with a tutor leads to better learning outcomes, while the effects of having control over problem selection still need further investigation. Our findings help establish that reflective self-assessment is beneficial for students learning with math problem solving tasks in ITSs. To the best of our knowledge, our study is the first controlled experiment that supports the theoretical claim that OLMs can enhance students' learning outcomes. The future design of effective OLMs should consider incorporating features that can facilitate students' self-assessment to better support metacognition and Self-Regulated Learning.

Acknowledgements. We thank Jonathan Sewall, Borg Lojasiewicz, Octav Popescu, Brett Leber, Gail Kusbit and Emily Zacchero for their kind help with this work. We would also like to thank the participating teacher and students. This work is funded by an NSF grant to the Pittsburgh Science of Learning Center (NSF Award SBE0354420).

References

1. Aleven, V., et al.: A New Paradigm for Intelligent Tutoring Systems: Example-Tracing Tutors. International Journal of Artificial Intelligence in Education 19, 105–154 (2009)
2. Brusilovsky, P., Sosnovsky, S., Shcherbinina, O.: QuizGuide: Increasing the Educational Value of Individualized Self-Assessment Quizzes with Adaptive Navigation Support. In: Nall, J., Robson, R. (eds.) Proceedings of World Conference on E-Learning, E-Learn 2004, pp. 1806–1813. AACE (2004)
3. Bull, S., Dimitrova, V., McCalla, G.: Open Learner Models: Research Questions (Special Issue of IJAIED Part 1). International Journal of Artificial Intelligence in Education 17(2), 83–87 (2007)
4. Bull, S., Jackson, T., Lancaster, M.: Students' Interest in Their Misconceptions in First Year Electrical Circuits and Mathematics Courses. International Journal of Electrical Engineering Education 47(3), 307–318 (2010)
5. Dunlosky, J., Lipko, A.: Metacomprehension: A Brief History and How to Improve Its Accuracy. Current Directions in Psychological Science 16, 228–232 (2007)
6. Hartley, D., Mitrovic, A.: Supporting Learning by Opening the Student Model. In: Cerri, S.A., Gouardères, G., Paraguaçu, F. (eds.) ITS 2002. LNCS, vol. 2363, pp. 453–462. Springer, Heidelberg (2002)
7. Koedinger, K.R., Baker, R.S.J.d., Cunningham, K., Skogsholm, A., Leber, B., Stamper, J.: A Data Repository for the EDM Community: The PSLC DataShop. In: Romero, C., Ventura, S., Pechenizkiy, M., Baker, R.S.J.d. (eds.) Handbook of Educational Data Mining. CRC Press, Boca Raton (2010)
8. Long, Y., Aleven, V.: Students' Understanding of Their Student Model. In: Biswas, G., Bull, S., Kay, J., Mitrovic, A. (eds.) AIED 2011. LNCS, vol. 6738, pp. 179–186. Springer, Heidelberg (2011)
9. Mitrovic, A., Martin, B.: Scaffolding and Fading Problem Selection in SQL-Tutor. In: Hoppe, U., Verdejo, F., Kay, J. (eds.) Proceedings of the 11th International Conference on Artificial Intelligence in Education, pp. 479–481. Springer, Berlin (2003)
10. Schraw, G.: A Conceptual Analysis of Five Measures of Metacognitive Monitoring. Metacognition and Learning 4(1), 33–45 (2009)
11. Waalkens, M., Aleven, V., Taatgen, N.: Does Supporting Multiple Student Strategies Lead to Greater Learning and Motivation? Investigating a Source of Complexity in the Architecture of Intelligent Tutoring Systems. Computers & Education 60, 159–171 (2013)
12. White, B.C., Fredrickson, J.: Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students. Cognition and Instruction 16, 39–66 (1998)
13. Zimmerman, B.J.: Investigating Self-Regulation and Motivation: Historical Background, Methodological Developments, and Future Prospects. American Educational Research Journal 45(1), 166–183 (2008)