Long, Y., Aleven, V. (2011). Students' Understanding of Their Student Model. In G. Biswas, S. Bull, J. Kay & A. Mitrovic (Eds.), Proceedings of the 15th International Conference on Artificial Intelligence in Education, AIED 2011, (pp. 179-186), New York: Springer. doi: 10.1007/978-3-642-21869-9_25

Students' Understanding of Their Student Model

Yanjin Long and Vincent Aleven

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

Abstract. Open Learner Models (OLM) are believed to facilitate students' metacognitive activities in learning. Inspectable student models are a simple but very common form of OLM that grant students opportunities to get feedback on their knowledge and reflect on it. This paper uses individualized surveys and interviews with high school students who have at least three years experience learning with the Cognitive Tutor regarding the inspectable student model in the Tutor. We also interviewed a teacher. We found that: i) students pay close attention to the OLM and report that seeing it change encourages them to learn; ii) there is a significant discrepancy between the students' self-assessment and the system's assessment; iii) students generally rely on the OLM to make judgments of their learning progress without much active reflection. We discuss potential revisions to the student model based on the findings, which aim to enhance students' reflection on and self-assessment of their own learning.

Keywords: Open learner model, student model, self-assessment, Cognitive Tutor

1 Introduction

Recently, many Intelligent Tutoring Systems (ITSs) researchers have studied the potential benefits of an Open Learner Model (OLM), in particular, whether it can help to improve students' metacognitive skills [5]. An OLM is a model accessible to the students that displays details of the student's learning status, such as their knowledge, difficulties, misconceptions, etc. [4]. Bull summarizes four primary OLM types: inspectable, co-operative, editable, and negotiated models [3]. The current work focuses on the first type, inspectable student models, which are the least sophisticated but probably the most common, as we argue below. As Bull and Kay [4] point out, a key purpose of an OLM is to support metacognitive activities such as reflection, planning and self-assessment. The model provides feedback with respect to students' learning and knowledge and it may trigger and facilitate metacognitive activities.

There has been only a limited amount of empirical work that supports the notion that OLMs can facilitate metacognition. In a survey study by Bull regarding college students' attitudes toward potential OLMs [3], most students expressed interest in accessing the models for the purpose of planning their learning and reflecting on it. The OLM was also viewed as a useful navigation aid. However, this survey was conducted before students actually used the tutor. A small number of investigations concentrated on students' field experience with student models. Three such studies suggest that even relatively simple inspectable student models can foster useful reflection by students and can enhance their domain-level learning and motivation. Arroyo *et al.* conducted an experiment to investigate the effects of an OLM that presented simple statistics about the given student's recent domain-level performance, together with metacognitive tips [2]. They found that students in the OLM group achieved greater learning gains and exhibited higher engagement than students who learned without the OLM. By contrast, metacognitive tips alone, without the accompanying simple OLM, were ineffective. A study by Mitrovic and Martin [8] with the SQL tutor investigated the effect of a simple inspectable student model that displayed (in the form of skill bars) students' progress in learning key concepts. They found that this OLM enhanced students' self-assessment and domain-level learning, especially for the less-able students. Finally, a study by Walonoski and Heffernan showed that an inspectable OLM can help reduce behaviors that reflect poor metacognition [10]. They designed an OLM for the purpose of counteracting students' "gaming the system" behaviors. The model plots a graphical trace of student actions with the system, in which gaming behaviors are easily visible. They found that the graphical feedback led to reduced gaming, perhaps due to greater reflection on the part of students, or because the display results in social pressure not to engage in gaming behaviors. However, no significant advantage on learning was found.

Although these studies highlight interesting connections between metacognition and OLMs and some tantalizing evidence about a potential positive influence of OLMs on metacognitive processes, little is known about whether and how OLMs might enhance the accuracy of students' self-assessment of their mastery of *specific* skills and concepts targeted in the instruction. Self-assessment has been recognized as a crucial metacognitive skill in self-regulated learning [11]. Accurate self-assessment can help students be aware of their difficulties and misconceptions, allocate attention to the proper learning topics, and even assist them in making learning plans [7].

We investigate relations between self-assessment and inspectable OLMs in the context of Cognitive Tutor, an ITS developed at Carnegie Mellon University since the early 1980s. This ITS is being used as part of the regular mathematics instruction in many US schools, and therefore provides an opportunity to study relations between self-assessment and OLMs in a real educational context with students who use the tutor over extended periods of time. In the current Cognitive Tutors, a skillometer (Fig.1) serves as an inspectable student model. It displays probabilities of skill mastery for the skills targeted in the current section of the tutor curriculum. Although the skillometer is a simple inspectable OLM, this type is in widespread use, not only in Cognitive Tutors, but also in constraint-based tutors, as mentioned above. The probabilities in the skillometer are calculated using a knowledge-tracing algorithm [6]. The skill bars gradually "grow" as students progress in the tutor and finally turn gold when the skill is fully mastered. The skillometer was added to the Cognitive Tutor to give students a sense of progress, and to help them understand how close they are to finishing a section of the tutor curriculum. An important assumption in Cognitive Tutors is that the skills in the tutor's cognitive model (and displayed in the skillometer) correspond closely to students' psychological reality. This assumption finds support both in Anderson's ACT-R theory [1] and in educational data mining results which show that the particular cognitive models used in tutors accurately account for student performance change over time [9].

Anecdotal reports from Cognitive Tutor classrooms indicate that students tend to pay close attention to their skillometers, perhaps affirming that they indeed serve as useful progress indicators. One might expect that the skillometer would also afford students opportunities to get feedback on the state of their knowledge and reflect on it, such as, for example: "Why have I not mastered this skill yet?" However, little prior work has investigated how students actually use the skillometers and whether this use facilitates students' self-assessment and reflection on their own skill mastery.

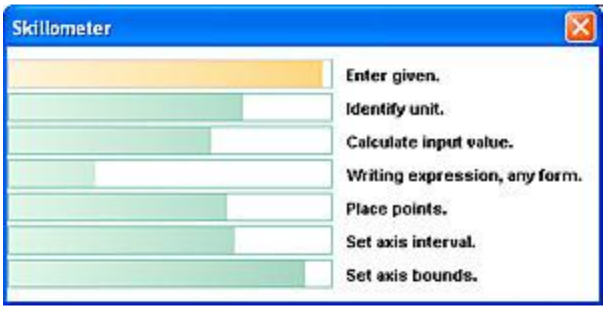

Fig. 1. Screenshot of the Skillometer.

The current study uses data from an individualized survey to find out whether an inspectable model can influence students' self-assessment. Specifically, we compared students' self-assessment against the system's assessment of their skill mastery, as displayed in the skillometer. We also investigated whether students were more likely to reflect on their own skill mastery when they *disagree* with the skillometer, which Bull and Kay suggest may be a key advantage of an inspectable student model [4]. Finally, we conducted interviews with students and a teacher to supplement the findings from the survey with detailed observations and explanations.

2 Survey with Cognitive Tutor Students

The purpose of the survey is to find out i) to what extent students' self-assessment of their skill mastery agrees with the system's student model (which, as mentioned, reflects the probability of mastery of each skill, as inferred from their performance over a range of problems) and ii) the relation between students' disagreement with the student model and their reflective activities.

2.1 Participants, Materials and Procedure

The survey was conducted in a high school in a school district near Pittsburgh. A total of 47 students completed the survey. All the students were enrolled in Cognitive Tutor classes with the same teacher, including Algebra I, Algebra II and Geometry. The age ranged from 15 to 18 years old, and all the students have been in Cognitive Tutor classes for at least three years.

In order to investigate relations between students' assessment of their own skills and the system's assessment, *individualized* survey forms were created, as follows: For each student, a "high skill" and a "low skill" were identified just prior to administering the survey, using automated reports provided by the tutoring software. A high skill had a probability of mastery above 0.6 (according to the tutor's knowledge-tracing algorithm), a low skill a probability lower than 0.4. Individualized survey forms were then put together with three groups of questions, the first two of which varied by the individual student: (1) questions about the high skill (2) questions about the low skill and (3) general questions about the skillometer. For both skills, the participants were asked to rate their overall mastery of the skill on a 7-point Likert scale. They were also asked to self-rate various additional aspects of their mastery and understanding of the skill, such as whether they are good at using this skill, whether they can give an example of a problem in which the skill would be used, and whether they feel they need more practice with the skill. Due to technical problems, we did not have skill levels available for all students at the time we designed the surveys, so we also created a generic version of the survey, which was the same as the individualized version, except that the skills referred to in the first two sections were randomly picked. Only the third sections of these generic surveys were analyzed; the first two parts were added only to make all surveys look equivalent to the participants.

All the surveys were handed out during the students' Cognitive Tutor class time and each took less than 10 minutes to finish. The students were not logged in to the tutor at the moment the surveys were taken, so they could not look at the OLM.

2.2 Results

A total of 47 students participated in the survey, of whom 35 completed an individualized version and 12 completed the generic one.

Agreement between Self-Assessment and System-Assessment. The 35 individualized surveys were analyzed to test whether students' self-assessment of their skills agrees with the system's assessment, as captured in the student model. Specifically, we tested whether the survey scores for the high skill are higher than those for the low skill. As mentioned, students rated their skill mastery on a scale from 1 to 7, where 7 represents greatest level of mastery. For the high skill, the average rating was 4.969 (*SD*= 1.402), and for the low skill, the average rating was 5.156 (*SD*: 1.629); this difference is not statistically significant ($t(30)=-1.329$, $p = 0.194$).

Table 1. Participants' Responses to Other Self-Assessment Questions.

		High Skill	Low Skill
Good at the Skill or Not?	Yes	24	25
	No	8	6
	Not Sure	3	4
Give an Example of the Skill	Yes	8	9
	No	27	26
More Practice on this Skill?	Yes	23	24
	No	7	8
	Not Sure	5	3

Additionally, Table 1 summarizes results from the other three self-assessment questions for both the high and low skills. We see that students' answers to the three questions do not differ much between the high and low skills. For example, 24 and 25 participants rated they were good at using the high and low skills, respectively. The results indicate a discrepancy between the students' perception of their skill mastery and the system's OLM. This discrepancy may be due to inaccurate self-assessment on the part of the students regarding their skill levels. Additionally, it is possible that the descriptions of the skills as they occur in the skillometer are not meaningful or understandable to the students. In the survey, the skills were described using the same short phrases that appear in the skillometer, illustrated in Fig. 1.

The question asking the students to give an example of a mathematics problem that involves the given skill was included mainly to test students' understanding of the skills displayed in the OLM. Two raters independently evaluated the answers. Not surprisingly, given the challenging nature of the question, only 8 (22.9%) participants gave examples for the high skill and 9 (25.7%) for the low skill. The examples given by the 17 students were mostly correct and were in the same format as they were presented in the Cognitive Tutor. We also found that the majority of students (23 for high skill, and 24 for low skill) preferred more practice on the skills. This preference for more practice is quite interesting. Again it is striking that there is no difference between the high skill and low skill questions, which may be evidence that students have difficulty in assessing their own skill.

Relation between Disagreement and Reflection. The results came from the third part of the survey, and all 47 participants' answers were analyzed.

Table 2. Cross-Table of Disagreement and Reflections.

		Disagreement		
		Yes	No	Total
Reflection	Yes	19	13	32
	No	11	4	15
	Total	30	17	47

Table 2 presents results from Question 1 "Do you sometimes disagree with the skillbar?" and Question 4 "Do you reflect on what you have learned in the tutor when you finish each section?" A majority of participants indicated that they sometimes disagreed with the skillometer (30 participants, 63.8%) and reflected on their learning (32, 68.1%). The relationship between students' disagreement and reflection is not statistically significant (*chi*(1)=.862, *p*=.353). Thus, our study finds no strong support for Bull and Kay's hypothesis [4] that disagreement with the OLM leads to reflection.

For Question 3 "Does the skillbar accurately describe what you know and what you don't know in the tutor?", students' answers varied considerably. 23 students (48.9%) answered yes, 15 (31.9%) answered no, 2 (4.3%) answered "sometimes" and 7 (14.9%) indicated "not sure". In response to the question "How often do you look at the skillbar in your tutor?" 28 (59.6%) participants reported they look at the skillometer each time they finish a problem and 9 (19.1%) that they refer to it several times per session. These findings confirm that the majority of the students pay close attention to the skillometer, as we had heard in anecdotal reports from the classroom.

2.3 Discussion

It is notable that there is a significant discrepancy between students' self-assessment and the system's assessment. It is reasonable to assume that the tutor's knowledge-tracing algorithm is accurate and that the skills in the tutor's cognitive model (which are displayed in the skillometer) accurately represent the knowledge components that students are actually learning, given the amount of research and development effort that has been invested in this area [1][6][9].Therefore, the discrepancy between the student's and system's assessment may indicate inaccurate self-assessment abilities of the students. It is possible also that the students have trouble understanding the skill names used in the skillometer, especially outside the Tutor.

In an inspectable student model, the students are simply viewing the model. Even if they sometimes disagree with the model, they cannot express this disagreement or "argue" with the model. Results from the survey suggest the need for negotiation with the students to some extent, since more than 60% students expressed disagreement with the skillometer. One of the goals of the interview portion of our study, therefore, was to hear students' viewpoints with respect to a possible negotiable student model.

3 Interview

Individual interviews were conducted to further investigate students' understanding about the skillometer, as well as to clarify some issues that emerged from the surveys.

3.1 Participants, Materials and Procedure

Five male students from the same teacher's Cognitive Tutor classes volunteered to participate in the interview. The interview was conducted individually in a conference room at the school All interviews were audio recorded with consent from both the students and parents. Each interview took 15 to 20 minutes. The students answered 15 questions regarding their perception and understanding of the skillometer. The 15 questions addressed the following themes: 1) how well do the students understand the skillometer? 2) how often do they trust/disagree with the skillometer? And 3) how much control do they prefer to have in the Tutoring system?

In order to gain a perspective from an instructor, a follow-up interview was conducted through email with the Cognitive Tutor teacher.

3.2 Results and Discussion

All five participants claimed that they paid close attention to the skillometer when they were using the Tutor. They also said that seeing the skill bars change encouraged them to learn in the system. As one student said "It keeps you wanting to go. If it goes down, you get mad. If it goes up, that makes you want to work better."

Understanding of the Skillometer. In general, the participants understand how the skillometer changes in response to their interactions with the tutor, although some misunderstandings exist as well. For example, three participants indicated that the bars would keep going down when they asked for further hint levels, which is not accurate. On the other hand, the teacher stated that some of the skill names were confusing even to her, and the students would ask her for explanations for the skill names from time to time. In future designs of the skillometer, we need to ensure that all skill names can be easily understood, or that other means are used to communicate what the names mean (e.g., examples linked to the skillometer).

Need for Negotiation and Control. All participants stated that they sometimes disagreed with the skillometer, and they would be upset if the system did not allow them to progress to the next section when all skills were mastered. It might be an interesting idea to let them choose their own problems when working with the Cognitive Tutor. However, none of the participants actually prefer to pick their own problems instead of letting the system choose. One of the students said "that could be useful, but I can see … how it could be abused, just like you just choose problems that were easier for you to do." In general, these findings suggest that there is interest in negotiating with the system about the content of the student model. At the same time, the students do not seem to want strong control over their learning process. They trust the system and find it convenient to rely on it. So it is still an open question how much control a negotiable student model should give to students.

Lack of Reflection and Self-Assessment. The students rely heavily on the skillometer to decide what they know and what they still need to learn, in other words, the students do not usually reflect on or try to assess their own mastery of the skills targeted in the tutor. The perspective from the teacher confirms this observation. She wrote "I do not think students have good self assessment of their own skill levels. I feel they are just concerned with getting their bars yellow, but are not too concerned with what the bars mean or say." Also "I do not think that most of my students take time to reflect. Unfortunately, they just want to get it done and move on." These results bring up an essential question. The inspectable student model supports students in telling what they have mastered and what they have yet to master, and thus gives them clues as to what they should still work on. However, such convenience may hinder their thinking and reflection during the learning process, and reinforces a simplified notion of progress as only the changing of the skill bars. Perhaps prompting the students to assess their own skills first, before comparing with the skillometer, can be a better way of facilitating reflection on the part of students.

4 Future Work and Conclusion

In sum, this study confirms that students generally pay close attention to the skillometer. They stated that seeing the skill bars change encourages them to learn. We also find a significant discrepancy between students' self-assessment and the system's, which indicates perhaps that the student model is not fully understandable, but also that there is room for improvement in students' self-assessment abilities.

The long-term goal of the current project is to investigate how a student model can assist students in productive reflection and better self-assessment of skill mastery. Specifically, an interactive negotiable student model that prompts students to reflect may result in more advanced self-assessment abilities, combined with support for comparing with the information in the inspectable student model. It will be interesting to investigate how much control a negotiable student model should give to students in order to achieve the best learning outcome. Another interesting future topic might be showing students indicators of their improvement in the skillometer, analogous to Arroyo et al.'s simple progress indicators [2]. Finally, more in-depth qualitative methods like think-aloud protocols can be used in investigations to find out more information regarding students' understanding of the skillometer and motivation.

Acknowledgments. We would like to thank Gail Kusbit, Matthew Bernacki and Tristan Nixon for their kind help with this work. We would also like to thank the participating teacher and students. This work was funded by an NSF grant to the Pittsburgh Science of Learning Center (NSF Award SBE0354420).

5 References

1. Anderson, J. R. *Rules of the Mind.* Hillsdale, NJ: Erlbaum, (1993)
2. Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Mehranian, H., Fisher, D., Barto, A., Mahadevan, S., Woolf, B. Repairing Disengagement with Non Invasive Interventions. *International Conference on Artificial Intelligence in Education, Marina del Rey, CA, 195-202*, (2007)
3. Bull, S. Supporting Learning with Open Learner Models. *4th Hellenic Conference: Information and Communication Technologies in Education, Athens,* (2004)
4. Bull, S., Kay, J. Student Models that Invite the Learner In: The SMILI Open Learner Modeling Framework, *International Journal of Artificial Intelligence in Education, 17(2) 89-120,* (2007)
5. Bull, S., Kay, J. Metacognition and Open Learner Models. *Proceedings of Third Workshop Meta-Cognition and Self-Regulated Learning in Educational Technologies (ITS '08),* (2008)
6. Corbett, A., McLaughlin, M., Scarpinatto, K. Modeling Student Knowledge: Cognitive Tutor in High School & College, *User Modeling and User-Adapted Interaction, 10, 81-108,* (2000)
7. Mitrovic, A. Investigating students' self-assessment skills. In M. Bauer, P.J. Gmytrasiewicz & J. Vassileva (Eds.) *Proceedings of the 8th International Conference on User Modeling, 247-250. LNAI 2109. Berlin: Springer-Verlag.* (2001)
8. Mitrovic, A. & Martin, B. Evaluating the Effects of Open Student Models on Learning, in P. de Bra, P. Brusilovsky & R. Conejo (eds), *Proceedings of 2nd International Conference on Adaptive Hypermedia and Adaptive Web-based Systems, Springer-Verlag, Berlin Heidelberg, 296-305,* (2002)
9. Ritter, S., Harris, T., Nixon, T., Dickison, D., Murray, C., Towle, B. Reducing the Knowledge Tracing Space. In Barnes, Desmarais, Romero, & Ventura (Eds.). *Proceedings of the 2nd International Conference on Educational Data Mining. pp. 151-160. Cordoba, Spain,* (2009)
10. Walonoski, J.A., Heffernan, N.T. Prevention of Off-Task Gaming Behavior in Intelligent Tutoring Systems. *Proceedings of the 8th International Conference on Intelligent Tutoring Systems, 722-724,* (2006)
11. Winne, P. H., & Hadwin, A. F. Studying as Self-Regulated Learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in Educational Theory and Practice, 279-306,* Hillsdale, NJ: Erlbaum, (1998)