The Role of Feedback in Preparation for Future Learning: A Case Study in Learning by Teaching Environments

Jason Tan and Gautam Biswas

Dept. of EECS & ISIS, Box 1824 Sta B,
Vanderbilt University
Nashville, TN 37235. USA
Tel.: (615)-343-6204
{jason.tan, gautam.biswas}@vanderbilt.edu
http://www.teachableagents.org

Abstract. Past research on the timing and content of feedback on student learning in computer-based learning environments has shown that directed or corrective feedback helps with immediate learning, whereas guided and metacognitive feedback help in gaining deep understanding of the domain and developing the ability to transfer this knowledge. Feedback becomes important in discovery learning environments, where novice students are often overwhelmed by the cognitive load associated with learning and organizing new knowledge while at the same time monitoring their own learning progress. We focus on feedback mechanisms in Betty's Brain, a teachable agent system in the domain of river ecosystems. Our goal is to help improve students' abilities to monitor their agent, Betty's knowledge, and, in the process their own learning and understanding. Our studies demonstrate the effectiveness of guided metacognitive feedback in preparing students for future learning.

1 Introduction

For the last five years we have been designing, implementing, and evaluating computer-based learning environments called Teachable Agents (TAs) that are based on the *learning by teaching* paradigm. In the TA paradigm, students learn science and math by teaching a computer agent through well-structured visual representations that help them organize their knowledge and thinking. The TA, using artificial intelligence techniques, reasons with the facts and relations it is taught to answer questions and solve problems. The TA can also illustrate its reasoning using graphics and animation, and this provides valuable feedback to the students. One of our TAs, called Betty's Brain, has been successfully used to teach about river ecosystems in 5th grade science classrooms [5]. This paper discusses the effectiveness of feedback mechanisms that we have built into the Betty's Brain system to help students with their assessment and learning tasks.

The cognitive science and education research literature supports the idea that teaching others is a powerful way to learn [4]. The literature on tutoring has shown that tutors benefit as much from tutoring as their tutees [7]. Beyond preparing to

M. Ikeda, K. Ashley, and T.-W. Chan (Eds.): ITS 2006, LNCS 4053, pp. 370 – 381, 2006.

teach, the act of teaching taps into three aspects of learning interactions – *structuring, taking responsibility, and reflecting* [3]. Whereas preparation to teach involves significant amounts of learning and organizing [3], our studies have found that for novices, a great deal of learning can occur through assessment and reflection during teaching [5, 13]. The feedback students receive by observing their TA's performance helps them discover what to prepare, and how to structure what they are learning to ensure the agent can understand and apply what it has been taught. However, 5th grade students, who are both domain novices and novices in teaching practices, often do not possess the skills to monitor their own knowledge. As a result, they often fail to analyze relevant pointers to errors and omissions in their knowledge. We focus on explicit feedback mechanisms that help improve students' abilities to monitor their agent's, and as a result, their own learning in a way that improves their understanding and problem solving capabilities. Feedback in the Betty's Brain environment comes from a Mentor agent who answers student's queries, and Betty who demonstrates the use of metacognitive learning strategies while being taught by the student.

This paper focuses on the impact of directed versus guided feedback on student learning in TA environments. The results of an experimental study run in a 5th grade science classroom are discussed in terms of the students' immediate learning abilities and their preparation for future learning [12].

2 Betty's Brain

Our TA, Betty, shown in Fig. 1, is taught using a concept map representation [10]. Students teach her about entities, such as fish and algae, and their relations, (e.g., fish consume dissolved oxygen, algae replenish it) as they pertain to river ecosystems. Once taught, Betty uses qualitative reasoning methods to reason through chains of links [5], which helps her answer questions, such as "*if macroinvertebrates increase what happens to bacteria?*" Learning by teaching is implemented as three primary components: (i) *teach* Betty using a concept map, (ii) *query* Betty with your own questions to see how much she has understood, and (iii) *quiz* Betty with provided tests to see how well she does on questions you may not have considered. These activities are usually embedded within a larger narrative (e.g., teach Betty so she can pass a test to join a science club) [5].

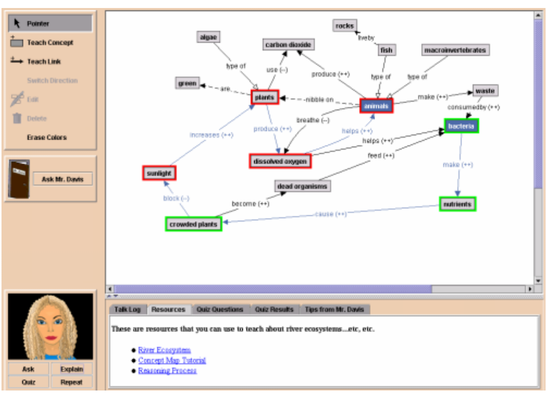

Fig. 1. Betty's Brain – Interface

When asked, Betty explains her answers using text and animation. Students reflect on Betty's answers and revise their own knowledge as they make changes to the concept maps. Details of the Betty's Brain system are summarized in [5]. Our work has demonstrated that one of the primary benefits of learning by teaching a TA is the need to structure knowledge in a compact and communicable format so that the student-teacher may develop important explanatory structures for the domain. The fact that TAs have independent performance and can show their reasoning based on how they have been taught also helps students (and teachers) assess their teaching. This should provide metacognitive and self-assessment opportunities for students that can lead to superior learning and transfer.

To help novice students with their learning and teaching tasks, we built additional resources into the environment: (i) domain resources organized as searchable hypertext so that students can look up information as they teach Betty, (ii) a concept map tutorial that provides students information on causal structures and how to reason with these structures, and (iii) a Mentor agent, Mr. Davis, who provides on-demand feedback about learning, teaching, and domain knowledge ("Ask Mr. Davis"). The Mentor also provides feedback immediately after Betty takes a quiz.

An experimental study conducted in a 5^{th} grade science classroom has demonstrated that learning-by-teaching with metacognitive support for self-regulated learning helps students develop better learning strategies, and better prepares them for future learning on related topics, even when this learning happens outside of the support provided by the TA environment [5]. Students were divided into groups to work on three versions of the system: (i) Intelligent Tutoring System (ITS), where the Mentor asked students to create a concept map that would correctly answer a set of test questions, (ii) Learning by Teaching (LBT), where students taught Betty to help her pass a test so she could become a member of the school Science club, and (iii) Self-Regulated Learning (SRL), where students taught a Betty persona that incorporated metacognitive learning strategies [11, 15]. All three groups had access to identical resources on river ecosystems, the same quiz questions, and similar access to the query feature and the Mentor agent. The differences in performance for the three groups in the main study were not significant (we studied the quality of the concept maps students generated and the quiz scores). However, in a *preparation for future learning* task, where all students had to construct a concept map to answer questions about the land-based nitrogen cycle (a topic they had not studied before), the SRL group created maps with more concepts and links than the ITS and LBT groups. The effects of teaching self-regulation strategies had an impact on the students' abilities to learn a new domain [5]. These results were very encouraging, and prompted us to study metacognitive feedback in a more systematic way.

3 Previous Studies on Feedback

A number of researchers have studied the effects of timing and content of feedback provided by computer-based tutors and pedagogical agents on student learning and problem solving performance. With their Lisp Tutor, Corbett and Anderson [8] found that students who received immediate feedback (the tutor intervened as soon as students made errors and forced them to correct the error before they could move on) went through the tutoring lessons faster than students who had to ask the tutor for feedback. However, the latter group made lesser errors in post-test debugging and code generation tasks. This study demonstrated that immediate error feedback helped with immediate learning, but there were indications that providing students with more control (on-demand feedback) led to better retention and deeper understanding.

Aleven and Koedinger [1] performed studies on students' help-seeking behavior with a Cognitive Tutor for Geometry that provided on-demand help at multiple levels of detail, starting with general strategies relevant to a problem solving step to specific hints, which explicitly outlined the correct solution to the step. In initial studies, the researchers found that most students did not know when to ask for help and when prompted, they quickly clicked down to the most detailed hint. Students with higher pretest scores made fewer errors, asked for less hints, and did better on post tests. However, it was not clear that feedback helped students improve their overall learning. In later work [2], they incorporated self-explanation, where the students were required to explain their problem solving steps. In addition to error feedback, the system provided self-explanation hints centered on general strategies related to the current problem. Students showed deeper learning when the tutor required them to explain their steps. Students in the explanation condition spent more time on the system than students who were not required to provide self-explanations, but they needed fewer problems to achieve predetermined mastery levels for skills.

Moreno [9] used feedback as a mechanism for decreasing the cognitive load of novice students using software agents in discovery-based multimedia environments. Her study included: (i) a guided learning environment, where the agent provided explanatory feedback and (ii) a directed learning environment, where the agent provided corrective feedback. Her guided discovery hypothesis centered on the belief that learning occurs when learners actively construct a coherent knowledge representation by meaningful interactions with resource materials, converting the information extracted into representations, and integrating new information into existing representations. Typically, discovery learning results in high cognitive load for students with low prior knowledge, making it hard for them to learn. Her studies showed that the guided feedback group found the instructional material easier to follow, made significantly fewer errors on post-test questions, and was much better than the directed feedback group at transfer tasks that involved novel situations.

These studies support our early findings with the Betty's Brain system. On-demand, guided feedback is very likely to lead to better learning and transfer in discovery learning environments, where students are involved in the construction of knowledge. However, the complexity of the constructive discovery and problem solving tasks may overwhelm students who have low prior knowledge of the domain. The resulting frustration may distract the students and cause them to abandon their learning tasks. In the remainder of this paper, we discuss three versions of the Betty's Brain system designed to compare the differences between directed and guided feedback and study the effects of the content of guided feedback on students' understanding of domain knowledge and their preparation for future learning.

4 Different Forms of Feedback in Betty's Brain

For the three versions of Betty's Brain, we started with the previous LBT and SRL versions of the system. The LBT system is designed to provide directed or corrective feedback. Betty provides answers to queries and explanations of how she derived her answers using text, speech, and animation when asked by the student [5]. The Mentor provides feedback after Betty takes a quiz by overlaying the part of the expert map used to answer the quiz questions on the concept map the student created to teach Betty. For every incorrect answer, the Mentor looks to see if the concepts in the quiz question appear in the student's map. If they do not, the Mentor suggests that the student study these concepts in the resources. Otherwise, the Mentor looks for the first (i) missing expert concept, (ii) missing expert link, and (iii) incorrect expert links, in that order, to generate the appropriate feedback content. Like the Cognitive Tutors, the Mentor is designed to provide hints that range from general (e.g., "read about algae and dissolved oxygen") to specific ("you are missing a link between algae and dissolved oxygen in your concept map").

Our SRL system feedback is designed to teach students a set of comprehensive skills that includes setting goals for learning new materials and applying them to problem solving tasks, deliberating about strategies to enable this learning, monitoring one's learning progress, and then revising one's knowledge, beliefs, and strategies as new materials and strategies are learnt. Betty's persona in the SRL version incorporates self-regulation [5] and metacognitive strategies [13]. For example, when the student is building the concept map, she occasionally responds by demonstrating reasoning through chains of events. She may query the user, and sometimes remark (right or wrong) that the answer she is deriving does not seem to make sense. The idea of these spontaneous prompts is to get the student to reflect on what they are teaching, and perhaps, like a good teacher, check on their tutee's learning progress. At times, Betty may directly suggest to the students that they need to query her to ensure that she can reason correctly with the current concept map. At other times, Betty refuses to take a quiz, because she feels that she has not been taught enough, or that the student has not given her sufficient practice by asking queries.

In addition to comparing directed versus guided feedback, we were also interested in determining how the content of the feedback provided by the TA would affect the students' learning behaviors. Feedback content can be categorized as either cognitive or affective. Cognitive feedback is based on beliefs, thoughts, and rational arguments

Table 1. Example patterns of behavior andBetty's SRL-C and SRL-A responses

Pattern	Cognitive Response	Affective Response
If after four questions, Betty has not been queried on an unlinked concept	Excuse me. You taught me a concept, but didn't teach me any relationships between it and other concepts. Please teach me more, and ask me questions to make sure I understand	Hey, I'm confused and I don't understand what you taught me. Please teach me more, and ask me some questions.
If quiz and causal query but no update	Hey, you haven't taught me anything new since my last quiz. My score will surely be the same. Teach me something, and ask me some questions to make sure I understand, before you send me to take another quiz	Hey! You're making me do really hard things and I don't like it.
If no resource access and no improvement on previous quiz score	Excuse me. I like what you are teaching me, but it may not help me pass the quiz. I would like to be better prepared when I take it again. Could you check the resources and teach me about what you find there? Thanks.	Excuse me, but that quiz is very difficult. I really don't want to take it now. Can we do something else?

that are related to a problem-solving task while affective feedback is based on feelings or emotions that arise when performing a task [14]. Cognitive feedback is directed to helping the student develop better skills whereas affective feedback more likely invokes empathy and feelings for the agent. Both kinds of feedback may promote better learning among students, even if they are for different reasons.

To study the differences between affective and cognitive feedback, we created two versions of the SRL system that take their cues from the same set of patterns, but provide either cognitive or affective feedback. Several examples of the two kinds of feedback are illustrated in Table 1. The SRL-Cognitive (SRL-C) feedback is more content-directed, and students are provided with hints that help them apply metacognitive strategies to improve their learning, monitoring, and debugging tasks. The SRL-Affective (SRL-A) feedback is triggered by the same patterns as the SRL-C system, but Betty's responses are emotional and based on her feelings. The next section presents the experimental study we conducted to compare the effectiveness of the different kinds of feedback on student learning.

5 Experimental Study and Results

This new study, conducted with two 5th grade classrooms, was designed to compare the effects of the different types of feedback. 37 students from the two classrooms were divided into the three groups (SRL-C, SRL-A, and LBT) using a stratified sampling method based on standard achievement scores in mathematics and language. The students worked with Betty's Brain for seven 45-minute sessions. Their goal was to successfully teach Betty about river ecosystems and get her to pass three quizzes (answer all questions correctly). Approximately 10 weeks later, the students were given the transfer test, where they again taught Betty about a new domain: the land-based nitrogen cycle. The students worked for three sessions and this permitted us to determine which group was better prepared to learn in situations where scaffolds and feedback from their previous environments were removed. We measured student learning along two dimensions: (1) Students' performance in creating concept maps and (2) students' learning behaviors (in the form of quiz attempts, queries, and resource accesses). We believed that students previously in the SRL-C condition would demonstrate the best performance for future learning, and students in LBT condition with directed feedback would perform better than the students in the SRL-A condition, who received no useful feedback.

Experimental Results

Students' activities on the TA systems in the main and transfer study were recorded in log files, along with Betty's and Mr. Davis' feedback. The students' concept maps were also saved at the end of each session. In evaluating the students' concept maps we considered both "expert" and "relevant" concepts and links. Concepts and links that appeared in the expert map were labeled as "expert." However, other concepts and links that corresponded to a correct understanding about the domain (though they were not required to answer quiz questions) were coded as "relevant." The set of expert and relevant concepts and links are called "valid."

Table 2 summarizes the average number of expert and valid concepts and links for each condition at the end of the final session. Overall, the LBT group had more expert links in their maps than the two SRL groups. However, the SRL-C group had more valid links than the LBT and SRL-A groups. We believe that the directed feedback focused on the quiz questions directed the LBT students only to the missing expert concepts and links in their maps. The metacognitive feedback for the SRL-C group was more general, and focused students on acquiring knowledge from the resources and organizing it into their concept maps. We believe that learning such behavior would promote better abilities to learn on one's own.

Table 2. Means and standard deviations (sd) of the number of concepts and links in the final concept map for the main study. (Differences were not statistically significant).

Main study	SRL Cognitive Mean (sd)	SRL Affective Mean (sd)	LBT Mean (sd)
Expert concepts	9.23 (1.7)	9.15 (1.8)	8.27 (2.2)
Valid concepts	21.69 (11.0)	18.00 (9.9)	14.27 (6.2)
Expert links	5.38 (3.0)	5.62 (3.9)	7.18 (4.7)
Valid links	18.15 (10.5)	14.92 (12.6)	13.36 (8.0)

To study this, we performed a similar analysis on the transfer test concept maps. All students used identical barebones systems with no feedback. Analysis of the expert concepts and links produced no statistically significant differences. However, the SRL-C group had more valid concepts and links than the other two groups, and the differences in the number of valid concepts was statistically significant (see Table 3). This demonstrates the effectiveness of metacognitive, guided feedback in preparation for future learning.

Table 3. Means and standard deviations (sd) in students' final concept map by group for the transfer test concepts and links. Significant differences are indicated.

Transfer test	SRL Cognitive Mean (sd)	SRL Affective Mean (sd)	LBT Mean (sd)
Expert concepts	6.54 (1.9)	5.31 (2.7)	6.09 (4.1)
Valid concepts	14.69[ab] (5.5)	10.23 (4.9)	10.27 (5.6)
Expert links	1.923 (1.4)	2.25 (2.7)	3.80 (5.0)
Valid links	10.85 (7.6)	8.5 (6.5)	9.3 (6.9)

[a] Significantly greater than SRL-A, $p < .05$; [b] Significantly greater than LBT, $p < .05$

In addition to evaluating students' performance, we also monitored their behaviors in the main study and the transfer test. Analyzing the student log files revealed differences between the three groups that can be attributed to the differences in feedback they received. We focus on quiz attempts, queries asked, and resource accesses as they demonstrate the students' abilities to monitor their learning and seek new information. We computed the average counts of each behavior across all sessions in both the main study and the transfer test.

Correlations were computed to see if the behavior patterns could be linked to improvements in the students' concept maps on a session by session basis. The number of new valid links added in each session for each student was compared against the number of times the specific actions were performed for that session. These pairs were aggregated for each group, and a Pearson's r test was performed to determine the significance of the association between the two sets of variables.

Figure 2a shows that the SRL-A group made a much larger number of quiz attempts when compared to the SRL-C group in all of the main study sessions. Asking Betty to take a quiz allowed the students to monitor Betty's progress and their own teaching and learning. However, making Betty repeatedly take quizzes showed that the student resorted to a trial-and-error pattern. To avoid this behavior, we created an SRL behavior pattern (see Table 1) where Betty refuses consecutive quiz attempts if the concept map is not updated and the student has not queried Betty between the quiz attempts. The number of quiz attempts refused shows an even larger difference between the two SRL groups (Figure 2d and Table 4). Although Betty refuses to take a quiz in exactly the same situations for both SRL groups, the SRL-C group seemed to avoid the trial-and-error strategy and focus on more systematic and effective learning methods. Interestingly, when the SRL patterns and feedback were removed in the transfer test, the SRL-C group continued to use the quiz feature more effectively than the SRL-A group (see Figure 3a). The difference in quiz attempts, combined with the fact that the SRL-C students were more successful in adding valid links to their concept maps in the transfer test, shows that they effectively used the quiz feature even when the scaffolding and feedback were removed.

Table 4. Means, standard deviations, and statistical significance for the main study counts

Main study	SRL Cognitive Mean (sd)	SRL Affective Mean (sd)
Quiz attempts refused (session 2)	1.77 (1.4)[a]	3.54 (2.2)
Quiz attempts refused (session 3)	1.92 (1.8)[a]	4.15 (3.3)
Quiz attempts refused (session 4)	4.31 (4.8)	10.54 (10.0)
Quiz attempts refused (session 7)	4.08 (3.9)[a]	9.33 (6.8)

[a] Significantly less than SRL-A, $p < .05$

Main study	SRL Cognitive Mean (sd)	SRL Affective Mean (sd)	LBT Mean (sd)
Queries (session 6)	7.00 (6.0)	12.92 (9.1)[ab]	3.73 (3.8)
Queries (session 7)	6.62 (6.7)	16.17 (13.7)[ab]	4.00 (4.6)

[a] Significantly greater than SRL-C, $p < .05$; [b] Significantly greater than LBT, $p < .05$

Main study	SRL Cognitive Mean (sd)	SRL Affective Mean (sd)	LBT Mean (sd)
Resource accesses (session 5)	4.92 (4.6)	1.92 (2.2)[ab]	5.00 (3.2)
Resource accesses (session 6)	5.08 (3.3)	2.00 (2.6)[a]	3.55 (2.7)

[a] Significantly less than SRL-C, $p < .05$; [b] Significantly less than LBT, $p < .05$

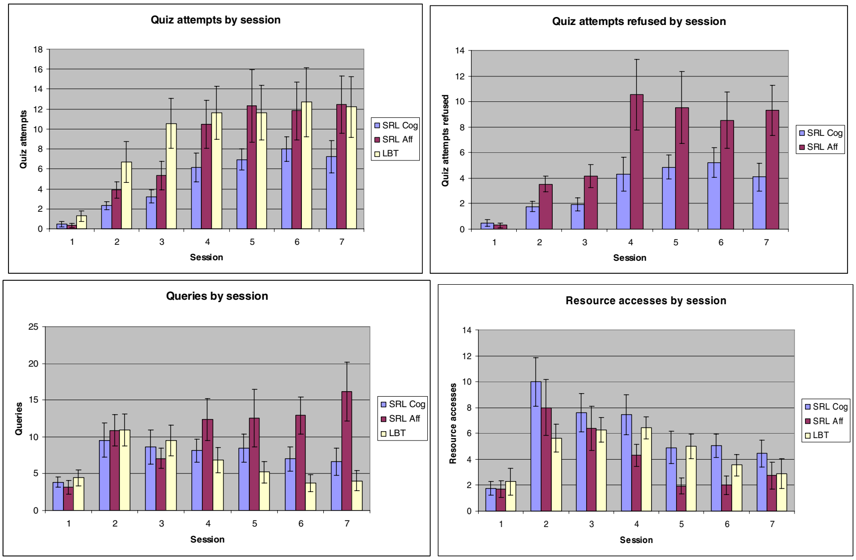

Fig. 2. Number of quiz attempts (a – top left), queries (b – bottom left), and resource accesses (c – bottom right) by session for the main study. Quiz attempts refused per session (d – top right) for the two SRL groups (the LBT system does not have SRL patterns).

In support of this claim, the correlation analysis showed that only the SRL-C group demonstrates a significant correlation between quiz attempts per session and number of added valid links for that same session. This correlation is negative, indicating that the students who were successful in adding valid links did not have to repeatedly use the quiz feature (Pearson r=-0.247, n=91, p=0.009). This behavior persisted in the transfer test.

Figure 2b shows that the SRL-A group used the query feature excessively in the later sessions. In particular, sessions 6 and 7 show a statistically significant difference between the use of the query feature between groups (see Table 4). Since the query feature is an important mechanism for monitoring Betty's knowledge, and, therefore, the student's teaching performance, the SRL Betty prompts the student to query her from time to time. This occurs most often after Betty refuses to take a quiz because the student has not checked to see if she has learnt what she was taught. Whereas the SRL-C group seemed to realize the role of the query feature in monitoring Betty's and their own knowledge, it is clear that the SRL-A group were generating queries just to get Betty to take the quiz. This group had many more quiz attempts refused, despite their efforts to use the query feature to get Betty to take the quiz. In the transfer test, there is no significant difference in the number of queries asked by the groups (see Figure 3b). This further supports the fact that the SRL-A group did not understand the role of queries. A correlation analysis for the main study showed that both the SRL-C and LBT groups exhibited a positive correlation between queries asked per session and number of added valid links for that same session (Pearson R for SRL-C: r=0.196, n=91, p=0.031; for LBT: r=0.226, n=77, p=0.024). This did not hold for the SRL-A group.

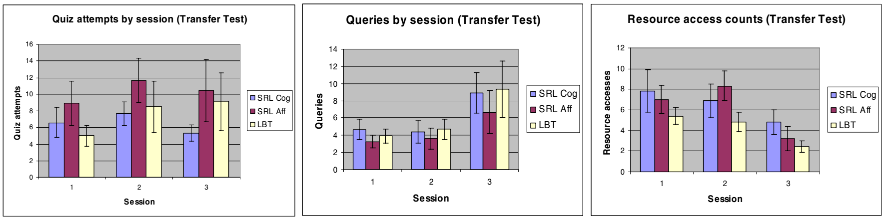

Fig. 3. Number of quiz attempts (a – left), queries (b – middle), and resource accesses (c – right) for all three groups across all sessions for the transfer test

Finally, examining students' resource accesses across the three groups reveals an interesting pattern in the students' behavior. The key to successfully teaching Betty hinges upon extracting information from the resources and transferring it into the concept map structure. The correlation analysis for the resource accesses per session and number of added valid links for that same session showed a positive correlation for all groups, although the data is statistically significant only for the SRL-A and SRL-C groups (Pearson R for SRL-C: $r=0.319$, $n=91$, $p=0.001$; for SRL-A: $r=0.260$, $n=88$, $p=0.007$; for LBT: $r=0.099$, $n=73$, $p=0.202$). This supports the obvious claim that the amount of resource use is directly related to the quality of the student's concept map. However, Figure 2c shows that the SRL-A group's resource accesses showed a significant drop from session 4 on (see Table 4). This suggests that Betty's affective explanations may have been more of a hindrance than a help for the SRL-A students, and their enthusiasm for teaching Betty may have dropped because of their lack of success. It is interesting to see the LBT group with no metacognitive prompting used the resources more often than the SRL-A group, despite the fact that they received corrective feedback from the Mentor.

6 Conclusions

The results from this study demonstrate the performance and behavioral differences in learning that can be associated with the different types of feedback provided by our TA system. Although directed or corrective feedback may allow the student to quickly achieve immediate goals set by the learning environment, like earlier work [9], we have demonstrated that guided metacognitive feedback better prepares the student for learning even when the student is removed from the learning environment. This was illustrated in the transfer study, when the students from the three different conditions were asked to learn a new domain in an environment with no scaffolding and very little feedback.

We have also demonstrated that the *presence* of guided feedback based on metacognitive cues is not enough. Students have to be explicitly taught the metacognitive strategies, and be given enough opportunities to practice them like the main study. The differences between the SRL-C and SRL-A groups indicate that the *type* of feedback received has a significant effect on learning outcomes. Students receiving cognitive content feedback were better able to learn from the teachable agent's metacognitive behaviors, but it is clear that the SRL-A students did not learn even though feedback was given in the same situations for both groups.

There are other issues in the learning by teaching framework with computer-based Teachable Agents that may contribute to student learning. One issue is the social aspect of the teaching process that very likely contributes positively to the student's motivation, and therefore, enhances his or her ability to learn. A second issue has to do with the notion of shared responsibility. In the TA environment, the agent knows only what the student has taught her. On the other hand, the student typically knows little of how to reason with causal structures, but learns by observing the agent answer questions and explain her answers. This results in a significant decrease in the student's cognitive load during initial learning. Last, there is the issue of the student monitoring their agent's knowledge, as opposed to their own (though, in reality it is their own knowledge). This again may result in a reduction of cognitive load, since the student is not problem solving and debugging their problem steps at the same time. (They would have to do this if they were generating a self-explanation for a problem they had solved). In future work, we would like to bring together all of these issues in redesigning our learning environments with guided, metacognitive feedback to provide better mechanisms that enable learning with deep understanding.

Acknowledgements

This work has been supported by a NSF ROLE Award #0231771. The assistance provided by the Teachable Agents group, especially Daniel Schwartz, Kadira Belynne, Bilikiss Adebiyi, Krittaya Leelawong, Marc Santoro, John Bransford and Nancy Vye are gratefully acknowledged.

References

1. Aleven V. & Koedinger, K. R., Investigations into Help Seeking and Learning with a Cognitive Tutor. *Working Notes AIED Workshop on Help Provision and Help Seeking In Interactive Learning Environments*. San Antonio, Texas, May 2001.
2. Aleven, V., McLaren, B., Roll, I., & Koedinger, K., Toward tutoring help seeking: Applying cognitive modeling to meta-cognitive skills. *Proceedings of Seventh International Conference on Intelligent Tutoring Systems*, pp. 227-239, 2004.
3. Artzt, A. F., & Armour-Thomas, E. Cognitive Model for Examining Teachers' Instructional Practice in Mathematics: A Guide for Facilitating Teacher Reflection. *Educational Studies in Mathematics*, 1999: 40(3), pp. 211-335.
4. Bargh, J.A. and Y. Schul, On the cognitive benefits of teaching. *Journal of Educational Psychology*, 1980. 72(5): pp. 593-604.
5. Biswas, G., D. Schwartz, K. Leelawong, N. Vye, and TAG-V, "Learning by Teaching: A New Agent Paradigm for Educational Software," *Applied Artificial Intelligence*, special issue on Educational Agents, vol. 19, no. 3, pp. 363-392, March 2005.
6. Bransford, J. D., & Schwartz, D. L. Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of Research in Education*, vol. 24: 1999, pp. 61-101. Washington DC: American Educational Research Association.

7. Chi, M.T.H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G., Learning from Human Tutoring. *Cognitive Science*, 2001. 25(4): pp. 471-533.
8. Corbett, A.T. and Anderson, J.R., Locus of Feedback Control in Computer-Based Tutoring: Impact on Learning Rate, Achievement and Attitudes, *Proc. CHI Letters*, vol. 3, no. 1, pp. 245-352, 2001.
9. Moreno, R., Decreasing Cognitive Load for Novice Students: Effects of Explanatory versus Corrective Feedback in Discovery-Based Multimedia. *Instructional Science*, vol. 32, pp. 99-113, 2004.
10. Novak, J.D., Concept Mapping as a tool for improving science teaching and learning, in *Improving Teaching and Learning in Science and Mathematics*, D.F. Treagust, R. Duit, and B.J. Fraser, Editors. 1996, Teachers College Press: London. pp. 32-43.
11. Pintrich, P. R. and E. V. DeGroot, Motivational and self-regulated learning components of classroom academic performance, *Journal of Educational Psychology*, 1990, vol. 82: pp. 33-40.
12. Schwartz, D. L. and Martin, T. Inventing to prepare for learning: The hidden efficiency of original student production in statistics instruction. *Cognition & Instruction*, 2004: 22, pp. 129-184.
13. Schwartz, D.L., K.B. Pilner, G. Biswas, K. Leelawong, and J.P. Davis, Animation of Thought: Interactivity in the Teachable Agent Paradigm. *Learning with Animation: Research and Implications for Design*, Lowe, Cambridge Univ. Press, to appear.
14. Verplanken, B., Hofstee, G. & Janssen, H. Accessibility of affective versus cognitive components of attitudes. *Eur. Journal of Social Psychology*, vol. 28, pp. 23-35, 1998.
15. Zimmerman, B. J. A Social Cognitive View of Self-Regulated Academic Learning. *Journal of Educational Psychology*, vol. 81(3), pp. 329-339, 1989.