A Formative Classroom Evaluation of a Tutorial Dialogue System that Supports Self-Explanation

Vincent ALEVEN
Octav POPESCU
Amy OGAN
Kenneth R. KOEDINGER
Human Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Ave, Pittsburgh, PA 15213, USA
aleven@cs.cmu.edu, octav@cmu.edu, aeo@andrew.cmu.edu, koedinger@cmu.edu

Abstract. Previous research has shown that self-explanation can be supported effectively in an intelligent tutoring system by simple means, namely, menu selection (Aleven & Koedinger, 2002). We are exploring whether self-explanation can be supported even more effectively by means of a tutorial dialogue system. When students explain in their own words, they may pay more attention to the crucial features of the problem and thus acquire more transferable knowledge. To test this hypothesis, we are developing the Geometry Explanation Tutor, a tutor which supports students as they solve problems and engages them in a restricted form of dialogue to help them state general explanations of their problem-solving steps. We conduced a formative evaluation study in a local junior high school, comparing two tutor versions, one in which students explain in their own words, one in which they explain by means of a simple menu. We found that there was little difference between the learning results of the two conditions. We also found that the pre-test scores in both conditions were unexpectedly high. Thus, the hypothesized advantages for explaining in one's own words do not seem to materialize for students who start out with considerable knowledge of the subject matter, yet may still materialize for students more typical of the target population.

Introduction

A number of cognitive science studies have shown that self-explanation is an effective metacognitive strategy that helps students learn with greater understanding(Aleven & Koedinger, 2002; Bielaczyc, Pirolli, & Brown, 1995; Chi, 2000; Renkl, 1999). Self-explanation can be supported effectively in intelligent tutoring systems, using fairly simple means such as templates or menus (Aleven & Koedinger, 2002; Conati & VanLehn, 2000; Trafton & Trickett, 2001). For example, Aleven and Koedinger showed that students working with a Cognitive Tutor learn more effectively when they explain their problem-solving steps simply by selecting from a menu the name of the problem-solving principle that justifies each step, as compared to solving problems without explaining.

However, self-explanation may be supported even more effectively by means of a tutorial dialogue system that allows students to state explanations in their own words and helps them in a dialogue to improve explanations that are deemed inadequate. When students explain in their own words, they are likely to pay more attention to the crucial features that must be present in a problem in order for a certain operator or rationale to apply. They may thus be less likely to acquire “shallow knowledge” by means of implicit learning processes and may be more likely to acquire knowledge at the right level of generality. Further, when students explain problem-solving principles in their own words, they may reveal more of the current state of their knowledge to the tutor. The tutor may thus be in a better position to help them remedy misconceptions and construct new knowledge to fill in gaps. Finally, an often-cited advantage of tutorial dialogue systems may be that they cause a “generation effect” by forcing students to recall information from memory rather than recognize items in menus (see e.g., Anderson, 1999, p. 194ff).

Such potential advantages must of course be weighed against potential disadvantages. The downside of having students explain in their own words may be that it takes students more time to produce adequate explanations, even if the system feedback is very helpful. Inevitably, some amount of effort and attention will be spent to work out “details” at the linguistic level, rather than thinking about the domain of interest. Just typing the explanation also takes time.

The question, then, is whether students get more out of solving fewer problems that they explain in their own words, or whether they learn more deeply when they solve more problems, but explain them more superficially, by means of menus or templates. In light of other recent results that show that “less can be more” in learning with instructional technology (e.g., Aleven & Koedinger, 2002), we pursue the hypothesis that the first alternative is better, that is, that students learn more deeply when they explain in their own words, supported by system dialogue. In order to test this hypothesis, we are developing a tutorial dialogue system called the Geometry Explanation Tutor (Aleven, Popescu, & Koedinger, 2001; 2002). This system provides support as students solve geometry problems (e.g., find unknown angle measures in a diagram) and helps them through a restricted form of dialogue to state general explanations for their problem-solving steps.

In a previous classroom experiment, we compared explanation by means of dialogue (as supported by the Geometry Explanation Tutor) versus explanation using a menu. Although the students who explained by means of dialogue had better learning results than those who explained by means of a menu (Aleven, Koedinger, & Popescu, in press), it is hard to come up with a coherent interpretation of the detailed pattern of results. Thus, we consider these results to be inconclusive. The fact that this study did not definitively show an advantage of dialogue over menus does not necessarily invalidate the main hypothesis presented above. It may well be that the dialogue system was not yet functioning well enough or that the types of dialogues that it supports were not well-suited for the given student population.

Therefore, we conducted a second study with the same design, comparing a system version in which students explain in their own words during dialogues with the system with a version one in which they explain by selecting from a menu. There were a number of differences between the studies: First, the system’s natural language understanding and dialogue management compoments had been improved considerably the second time around. Second, the student population was different. The first study was conducted in an inner-city school, the second in a suburban school. In the first study, one out of three class periods involved consisted of honors students, in the second study, all students were honors students. Finally, the second study was of shorter duration than the first.

In this paper, we describe the Geometry Explanation Tutor and illustrate its capabilities. We present the results from the second classroom study and discuss their implications.

1.	**Student:**	the bottom angles are the same
2.	**Tutor:**	You are getting there, but you need to be more precise. Are base angles always congruent?
3.	**Student:**	the bottom angles are congruent
4.	**Tutor:**	Base angles in what type of geometric figure are congruent?
5.	**Student:**	the bottom angles in an isosceles triangle are congruent
6.	**Tutor:**	[accepts as correct and complete explanation by means of implicit feedback]

1.	**Student:**	three angles in an isosceles triangle sum to 180 degrees
2.	**Tutor:**	Your explanation is close to being perfect. Except for the following: your explanation should not be specifically about isosceles triangles. Better delete that from your explanation.
3.	**Student:**	the sum of the three angles in all triangles is 180 degrees
4.	**Tutor:**	[accepts as correct and complete explanation by means of implicit feedback]

Figure 1: Two dialogues that students had with the Geometry Explanation Tutor, focused on the Isosceles Triangle Theorem and the Triangle Sum theorem.

The Geometry Explanation Tutor

The Geometry *Explanation* Tutor is based on the Geometry *Cognitive* Tutor, which is an integrated part of a full-year geometry course. The course and sofware together have been shown to be better than typical classroom instruction (Koedinger, Corbett, Ritter, & Shapiro, 2000). The Geometry Explanation Tutor focuses on one unit of the tutor curriculum, the Angles unit, which deals with the properties of angles in various types of diagram configurations. The problems that the tutor presents to students typically involve a diagram and given angle measures. The student is then asked to find certain unknown angle measures. The tutor provides assistance for problem solving in the form of context-sensitive hints and feedback, on solutions as well as intermediate steps. In addition, it supports self-explanation by means of a restricted form of dialogue.

Example student-tutor dialogues

The Geometry Explanation Tutor requires that students provide general explanations of their problem-solving steps in their own words and engages them in a restricted form of dialogue, illustrated in Figures 1 and 2, to help them improve their explanations. All three dialogues shown in these figures are actual dialogues of students working with the Geometry Explanation Tutor, collected during the evaluation study described in this paper. The dialogues are not representative of the complete set that was collected during the experiment; they show the tutor at its best. They were selected because they illustrateimportant capabilities of the system. When the students focus on the right geometry rule, the system helps them to state it at the right level of generality. For example, when students state overly general rules, as they tend to do, the tutor helps them to add the missing conditions, as is illustrated in the first example in Figure 1. When students state an overly specific rule, as in the second example in Figure 1, the tutor helps them make it more general. The tutor will also make it clear when the student focuses on the wrong rule or seems to focus on irrelevant concepts, as illustrated in Figure 2.

The dialogues presented in Figures 1 and 2 illustrate the hypothesized advantages and disadvantages of dialogue over menus, discussed in the previous section. Let us contrast these dialogues with a system in which students can explain their reasoning by means of a menu from which they could select the name of a problem-solving principle that justifies

1.	**Student:**	in an isosceles triangle, base angles are the same
2.	**Tutor:**	You gave a correct statement of a geometry rule, but that rule is not needed here.Please focus on a different rule.
3.	**Student:**	interior angles are congruent
4.	**Tutor:**	You may be thinking of the wrong geometry rule. Please focus on a different rule.
5.	**Student:**	when two parallel lines are cut by a transversal, the interior angles are congruent
6.	**Tutor:**	Your explanation looks good. However, you do not need to talk about interior angles here. So it is better to delete or change this part.
7.	**Student:**	when two parallel lines are cut by a transversal, the exterior angles are congruent
8.	**Tutor:**	Yes, you are on the right track. One thing, though: mentioning exterior angles here is not correct. So you should probably change that part of your explanation.
9.	**Student:**	when two parallel lines are cut by a transversal, the corresponding angles are congruent
10.	**Tutor:**	[accepts as correct and complete explanation by means of implicit feedback]

Figure 2: Example dialogue between a student and the Geometry Explanation Tutor

the step (i.e., geometry definitions and theorems). Such a system was used as the control condition in the experiment reported in this paper. In the two dialogues in Figure 1, the students focused on the right geometry rule right from the start. If the same students had explained the same steps by means of a menu, it is likely that they would have selected the right rule from the menu at the first try, leaving the tutor with the impression that they could explain their reasoning adequately. On the other hand, during the two dialogues shown it becomes clear that they do *not* have a solid understanding of the geometry rules in question or at least, are not able to state them adequately. In the course of these dialogues, the system focuses the students' attention on certain features of the problem that at first the students did not seem to recognize as being crucial to their explanation. It seems quite plausible that by pointing out specific elements that were missing from the students' explanation, the system helped the student solidify their understanding. More generally, it seemes plausible that the exemplified kinds of dialogues have a positive effect on students' learning outcomes, as compared to explaining by means of a menu.

On the other hand, keyboard-to-keyboard dialogue is likely to take more time than explaining by means of a menu. This hypothesized disadvantage of dialogues is perhaps illustrated in the dialogue shown in Figure 2. In constrast to what happened in the previous two example dialogues, the student does not quickly focus on the right geometry rule, but once she does, the explanation is complete and accurate at the first try. The analogous interaction with a menu-driven interface (i.e.,selecting the same sequence of geometry rules from a menu: Isosceles Triangle Theorem, Alternate Interior Angles Theorem, Alternate Exterior Angles Theorem, Corresponding Angles Theorem) would probably have taken less time. Bu then again, the menu-driven interaction misses some of the attractive properties of natural language dialogue: In the dialogue in Figure 2 the student mentions an important feature of the problem, namely, that it involves parallel lines intersected by a transversal, that probably would have been less salient in a simple menu-driven interaction.

Architecture of the Geometry Explanation Tutor

The system's architecture has been described in more detail elsewhere (Aleven, Popescu, & Koedinger, 2001), so here we provide only a brief outline. The system is built on top of an existing Cognitive Tutor for geometry problem solving, the Geometry Cognitive Tutor. As all Cognitive Tutors do (Anderson, Corbett, Koedinger, & Pelletier, 1995), the Geometry Explanation Tutor uses a cognitive model, in the form of production rules, to interpret the students' problem-solving steps (i.e., their calculations of unknown angle measures in diagrams). To evaluate and respond to student explanations, it uses a knowledge-based natural language understanding (NLU) component, described in more detail in Popescu, Aleven, and Koedinger (in press), combined with a simple dialogue management algorithm. Each student input is assumed to be an attempt at stating an explanation and is processed in three steps:

1. Create semantic representation—First, the system parses the student's explanation, using the LCFLEX left-corner chart parser (Rosé & Lavie, 1999) and simultaneously creates a representation of the semantic content of the explanation, implemented in the Loom term description logic system (MacGregor, 1991).

2. Classify the semantic representation—Next, the system (specifically, Loom's classifier) classifies the semantic representation. For this purpose, the system has an "Explanation Hierarchy," a fine-grained set of approximately 200 explanation categories, also defined in Loom. The categories represent common ways in which students state incomplete or incorrect explanations. The set of categories was developed in a data-driven manner, by analysis of several corpora of student explanations. Attached to each category is a sequence of (canned) messages that constitute appropriate feedback when an explanation classifies under that category.

3. Generate feedback—Finally, the system decides what feedback to give based on (a) the set of categories under which the explanation was classified and (b) the set of geometry rules (definitions and theorems) that can be used to justify the current step—the system determines which geometry rules have that property by running its cognitive model for geometry problem solving. It produces feedback as follows:

First, if the set of explanation categories indicates that the student's explanation contains any errors of commission, that is, if the explanation mentions any geometric concepts that are not relevant to the correct geometry rules, the tutor points out the concept that need not or should not be mentioned (see for example, Figure 1, second dialogue, step 2, or Figure 2, steps 6 and 8). If the explanation contains no content related to any of the correct rules, the tutor points out that the student is focusing on the wrong rule entirely (e.g, Figure 2, steps 2 and 4).

Second, if the set of categories under which the student explanation is classified contains a category representing a complete and correct explanation of a correct geometry rule (i.e., a geometry rule that can be used to justify the current problem-solving step), the tutor accepts the explanation by means of implicit feedback, as illustrated in the last step of each of the three example dialogues shown in Figures 1 and 2.

Finally, if the explanation is an incomplete statement of a correct geometry rule, the system provides feedback hinting at or indicating what is missing. It does so by selecting, from the set of categories under which the explanation was classified, the one that is closest to a category that represents a complete and correct statement of one of the correct geometry rules. It then presents the first of the sequence of feedback messages attached to that category (e.g., see Figure 1, first dialogue, step 2). If the student does not improve the explanation in subsequent attempts, that is, if the next attempt at stating an explanation yields the same set of explanation categories as the previous, the system selects the next message in the sequence, thus providing more specific feedback. For example, in the first dialogue in Figure 1, steps 1 and 3, the student changed the explanation from "the bottom angles are *the same*" to "the bottom angles are *congruent*". Since the tutor interprets those two sentences as having the same meaning, it follows up after step 3 with the next more specific feedback message attached to the relevant category, CONGRUENT-BASE-ANGLES.

As mentioned, the system had been improved quite considerably before the evaluation study reported in this paper, as compared to the previous study (Aleven, Koedinger, & Popescu, in press). The speed with which the system's NLU component processes student explanations had been increased greatly, mainly by reducing ambiguity in the grammar, the NLU component's reference mechanism had been revamped, the coverage of the grammar, lexicon, and ontology had been extended, the system's strategy for providing feedback had been extended so that it responds to commission errors and gives higher priority to commission errors than to omission errors, the Explanation Hierarchy had been expanded, and a significant number of feedback messages had been added to categories in the Explanation Hierarchy.

A Classroom Evaluation Study

In the spring of 2003, we conducted an evaluation study in a suburban junior high school. The purpose of the study was to test the hypothesis that students learn with deeper understanding when they explain their problem-solving steps in their own words, as compared to explaining by means of a menu. A secondary goal was to gather data on how well the system is functioning. As mentioned, this experiment constitutes the second elaborate classroom evaluation of the Geometry Explanation Tutor. The study took place within the context of a course based on the Integrated Mathematics curriculum, which includes concepts from both algebra and geometry. A total of 71 students participated in the study, three class periods, all taught by the same teacher. All students were honors students, meaning that they were the most gifted and diligent students within the given age group and school. At the start of the experiment, the students were assigned to two conditions, a "Dialogue" condition and a "Menu" condition. Two class periods were assigned in their entirety to one of the conditions. The students in the third period were assigned randomly to one of these conditions. Prior to the experiment, the teacher and students covered the textbook chapter on proof, which involves many of the geometry theorems that are covered in the tutor's Angles unit. The first activity of the experiment was an in-class, paper-and-pencil pre-test. During the same session, the first author demonstrated the Geometry Explanation Tutor in front of the class, using a data projector. All students then worked on the tutor for four 40-minute sessions. In the sixth and final session, all participants took a paper-and-pencil post-test.

Two tutor versions were used. The students in the Dialogue condition used the Geometry Explanation Tutor described above, in which they explain their reasoning steps in a (restricted kind of) dialogue with the tutor. The students in the Menu condition used a tutor version in which they could explain their steps by giving the name of a geometry definition or theorem. They could either type the name or select it from an on-line Glossary of geometry knowledge. The Glossary listed the relevant geometry theorems and definitions and provided further information about each rule upon the student's request. The Glossary was available freely to the students in both conditions, but only for the students in the Menu condition did it function as a menu from which to select rule names. The tutor versions were the same in all other respects.

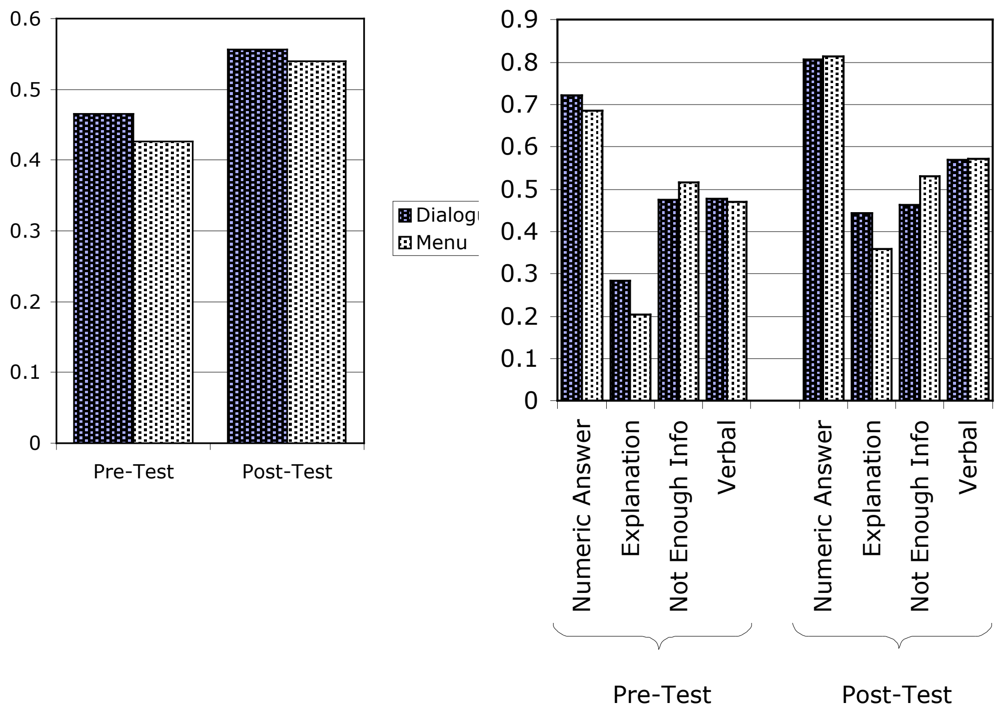

Figure 3: Overall (left) and detailed (right) pre- and post test results

The pre-test and post-test included regular "Numeric Answer" and "Explanation" items which were of the same type as the problems that students had encoutered during their work with the tutor. They also included two types of transfer items in order to measure any improvements in students' understanding: In some test items, the students were asked to judge whether there was enough information to find a particular unknown quantity. Items that involved a quantity whose value could not be uniquely determined we call "Not Enough Info" items. Items that dealt with a quantity whose value could be determined, given the available information, were grouped with the Numeric Answer items. We also included "Verbal" items, in which the students were given a general statement about geometry (e.g., "the sum of supplementary angles is 90 degrees") and were asked to say whether the statement is correct and to make any necessary corrections.

Of the 71 students, 62 completed the pre-test and post-test. Of those, 46 students worked on the tutor for at least 80 minutes, not considering idle time. We report the results of those 46 students. The cutoff point of 80 minutes may seem somewhat arbitrary, but none of the results noted below is very sensitive to the threshold used. The 46 students included 21 students in the Dialogue condition and 25 students in the Menu condition.

Results

The pre-test and post-test scores did not differ much between the conditions (see Figure 3). A repeated-measures ANOVA of students' test-scores with independent factors condition, test-time, and item type revealed a main effect of test time ($F(1,44) = 15.7$, $p < .0005$) but no significant interaction between condition and test-time. Thus, the students in both conditions improved and seemed to have learned about equally much from their respective instruction.

The data about the student-tutor interactions, shown in Table 1 and Table 2, provide some insight into how the experience of working with the tutor was different for

Table 1: On-line measures related to the interactions that students had with the Geometry Explanation Tutor

Condition	Time (mins)	Nr. of Answer Steps	Nr. of Expl Steps	Answer Time (mins)	Expl Time (mins)
Dialogue	116±15	34±9.8	33±9.9	30±12	75±14
Menu	103±12	57±7.1	56±7.2	48±8.3	36±9.7

Table 2: More measures related to the interactions that students had with the Geometry Explanation Tutor

Condition	Answer %Correct	Expl %Correct	Attempts/ Answer	Attempts/ Expl
Dialogue	65±13	29±9.3	1.59±.28	3.60±1.5
Menu	64±8.7	59±11	1.58±.24	1.76±.40

the students in each of the two conditions. Overall, the students in the Dialogue condition completed considerably fewer problem-solving and explanation steps than their counterparts in the Menu condition (see Table 1). Further, the students in the Dialogue condition spent a much greater proportion of their time explaining their problem-solving steps. During their work on the tutor, the students in both conditions did equally well on problem-solving steps (referred to as "Answer steps" in the two tables), that is, were equally good at determining the measures of angles in the given problems. The students in both conditions got an equal number of answers right without making any errors and also needed the same number of attempts needed to find the right answer. With respect to the explanation steps, however, there were considerable differences between the conditions. The students in the Dialogue condition got fewer explanations right at the first attempt than the students in the Menu condition and generally needed more attempts to arrrive at an explanation that the tutor deemed correct.

Discussion

The on-line data are consistent with the expectation that explaining in one's own words takes more time and is harder than explaining by means of a menu. To some degree, the difference may reflect imperfections in the system's dialogue management strategy, but we suspect that natural language self-explanation inherently is more difficult and time-consuming than explaining by means of a menu. Clearly, if self-explanation in students' own words is to result in deeper learning than explaining by means of a menu, in the same amount of time, students will have to get more out of each problem.

All in all, the results presented above provide no support for the hypothesis that students learn better when they explain their problem-solving steps in their own words, supported by dialogue, than when they explain by means of menus. If anything, they show a slight advantage for the Menu condition, whose pre-test to post-test improvement was .90 standard deviations, compared to .63 in the Dialogue condition, although as mentioned we found no statistically significant difference between the conditions. The results of the experiment therefore provide a further glimpse of evidence that supporting self-explanation by means of a simple menu is quite effective in a domain like mathematics, where one can easily identify a valid set of "reasons" in advance. As mentioned, in a previous experiment it was shown that the simple menu condition is in itself better than a Cognitive Tutor that supports problem-solving without explanation (Aleven & Koedinger, 2002).

While we find no evidence to confirm the main hypothesis of our research, we do not think the results from the current experiment disconfirm that hypothesis either. In light of the high pre-test scores, it seems that the students in this experiment were more advanced than the typical target population of the Geometry Explanation Tutor. Wo do not know whether they acquired their geometry knowledge in the context of earlier mathematics courses or during the classroom instruction in the weeks leading up to the experiment reported here. Regardless, it is possible that the students' relatively high level of geometry knowledge at the outset of the experiment obscured any differential effect that the two tutor versions may have had. If explaining in one's own words is more effective early on in learning, one would expect to see a greater advantage of the Dialogue condition with students who are not as well-prepared. The opposite effect however cannot be ruled out.

A second reason why we do not yet accept that the experiment disconfirmed the given hypothesis is the state of development of the system. Overall, based on our informal observations, the current experiment was a better experience for the students involved than the previous (Aleven, Koedinger, & Popescu, in press). Differences in student population, teachers, or system versions probably all contributed. The students in the second study seemed more able and better prepared learners and generally seemed more motivated to learn. The teacher involved in the second study seemed to enjoy a very good rapport with the students. Finally, the system used in the second experiment had been improved quite considerably. In particular, there was a very noticeable difference in the speed with which it processed student explanations and it also seemed to be more accurate. The teacher clearly shared our impression that the current classroom evaluation was a positive experience for the students involved.

Nonetheless, there seems to be room for improvement in the system's performance. We are in the process of analysing the data about student-tutor interactions extensively in order to get a sense for how well the system performed and how it can be improved. Our analysis of the accuracy of the system's NLU component (Popescu, Aleven, & Koedinger, in press) indicates that the system is beginning to approximate the accuracy of human raters, but that there is room for improvement. We are analyzing the data in a number of additional ways. First, we are assessing the quality of the system's feedback and analyzing the causes of low-rated feedback. As we did in a previous study (Aleven, Popescu, & Koedinger, 2002), we are looking at the relation between the feedback quality and the amount of progress that the students make in the next attempt at improving their explanations. We are studying the relation between the length of the dialogues that students' had with the system and their learning results. Finally, we are identifying the causes of lengthy dialogues, focusing on the 10% or 15% longest dialogues that we collected during the study. The results of these analyses should provide a wealth of interesting data about the particular strengths and weaknesses of knowledge-based natural language processing combined with a simple dialogue management strategy. They will point the way towards specific improvements that can be made.

Condusion

In a classroom study involving the Geometry Explanation Tutor, we found that there was little difference in the learning results of students who explained their problem-solving steps in a natural language dialogue with the system, compared to those who explained by means of a menu. Thus, the experiment does not confirm the hypothesis that explaining in one's own words represents a more favorable trade-off between time spent and learning gains.

However, the students who participated in the experiment seemed better-prepared and more able than the typical student in our target population. Further, our analyses of the system's performance indicate that there still is some room for improvement in the system, although considerable progress has been made. Therefore, we still think that the main research hypothesis is plausible. While the results of the current experiment indicate that the hypothesized advantages of explaining in one's own words are not pronounced for learners with relatively high knowledge of the subject matter, they do not rule out a greater effect for more average students.

Acknowledgements

The research reported in this paper is supported by NSF grants 9720359 and 0113864. We thank Jay Raspat of North Hills Junior High School for his inspired collaboration.

References

Aleven, V., & Koedinger, K. R. (2002). An Effective Meta-cognitive Strategy: Learning by Doing and Explaining with a Computer-Based Cognitive Tutor. *Cognitive Science,* 26(2), 147-179.

Aleven, V., Koedinger, K. R., Popescu, O. (in press). A Tutorial Dialog System to Support Self-Explanation: Evaluation and Open Questions. In *Proceedings of AIED 2003, The 11th International Conference on Artificial Intelligence in Education.*

Aleven V., Popescu, O. , & Koedinger, K. R. (2002). More on pilot-testing a tutorial dialog system that supports self-explanation: Good feedback on explanations helps, how we can generate more of it? In C. P. Rose & V. Aleven (Eds.), Working notes of the ITS 2002 Workshop on Empirical Methods for Tutorial Dialogue Systems. Available via http://www-2.cs.cmu.edu/~aleven/ITS2002DialogueWS/

Aleven V., Popescu, O., & Koedinger, K. R. (2001). Towards Tutorial Dialog to Support Self-Explanation: Adding Natural Language Understanding to a Cognitive Tutor. In J. D. Moore, C. L. Redfield, & W. L. Johnson (Eds.), *Proceedings of AI-ED 2001* (pp. 246-255). Amsterdam, IOS Press.

Anderson, J. R. (1999). *Cognitive Psychology and Its Implications* (5th ed.). New York: Worth Publishing.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive Tutors: Lessons Learned. *The Journal of the Learning Sciences,* 4, 167-207.

Bielaczyc, K., Pirolli, P. L., & Brown, A. L. (1995). Training in Self-Explanation and Self-Regulation Strategies: Investigating the Effects of Knowledge Acquisition Activities on Problem Solving. *Cognition and Instruction,* 13, 221-252.

Chi, M. T. H. (2000). Self-Explaining Expository Texts: The Dual Processes of Generating Inferences and Repairing Mental Models. In R. Glaser (Ed.), *Advances in Instructional Psychology,* (pp. 161-237). Mahwah, NJ: Erlbaum.

Conati C. & VanLehn K. (2000). Toward Computer-Based Support of Meta-Cognitive Skills: a Computational Framework to Coach Self-Explanation. *Intern. Journal of Artificial Intelligence in Education,* 11, 398-415.

Koedinger, K. R., Corbett, A. T., Ritter, S., & Shapiro, L. (2000). Carnegie Learning's Cognitive Tutor: Summary Research Results. White paper. Available from Carnegie Learning Inc., 1200 Penn Avenue, Suite 150, Pittsburgh, PA 15222, info@carnegielearning.com, http://www.carnegielearning.com.

MacGregor, R, (1991). The Evolving Technology of Classification-Based Knowledge Representation Systems. In J. Sowa (ed.), *Principles of Semantic Networks: Explorations in the Representation of Knowledge.* San Mateo, CA: Morgan Kaufmann.

Popescu, O., Aleven V., & Koedinger, K. R. (submitted). A Knowledge-Based Approach to Understanding Students' Explanations. Submitted to the AIED 2003 Workshop on Tutorial Dialogue Systems: With a View Towards the Classroom.

Renkl, A. (1999). Learning mathematics from worked-out examples: Analyzing and fostering self-explanations. *European Journal of Psychology in Education,* 16(4), 477-488.

Rosé, C. P. & Lavie, A. (1999). LCFlex: An Efficient Robust Left-Corner Parser. User's Guide, Carnegie Mellon University.

Trafton, J.G., & Trickett, S.B. (2001). Note-Taking for Self-Explanation and Problem Solving. *Human-Computer Interaction* 16, 1-38.