To Tutor or Not to Tutor: That is the Question

Leena RAZZAQ and Neil HEFFERNAN
{leenar, nth}@wpi.edu
Worcester Polytechnic Institute

Abstract. Intelligent tutoring systems often rely on interactive tutored problem solving to help students learn math, which requires students to work through problems step-by-step while the system provides help and feedback. This approach has been shown to be effective in improving student performance in numerous studies. However, tutored problem solving may not be the most effective approach for all students. In a previous study, we found that tutored problem solving was more effective than less interactive approaches, such as simply presenting a worked out solution, for students who were not proficient in math. More proficient students benefited more from seeing solutions rather than going through all of the steps. However, our previous study controlled for the number of problems done and since tutored problem solving takes significantly more time than other approaches, it suffered from a "time on task" confound. We wanted to determine whether tutored problem solving was worth the extra time it took or if students would benefit from practice on more problems in the same amount of time. This study compares tutored problem solving to presenting solutions while controlling for time. We found that more proficient students clearly benefit more from seeing solutions than from tutored problem solving when we control for time, while less proficient students benefit slightly more from tutored problem solving.

Keywords. Intelligent tutoring system, Tutored problem solving, Scaffolding, Feedback, Time on task.

Introduction

Many intelligent tutoring systems (ITS) rely on tutored problem solving (TPS), which requires students to work through problems step-by-step while the ITS provides hints and feedback, to improve student learning. Numerous studies have shown TPS to be effective in helping students to learn [1], [8], [10], [6]. However, TPS may not be the best approach for some students. In a previous study [5], we reported that the more time-consuming interactive TPS was indeed more helpful to less proficient students when compared to simply showing them a solution to the problem. On the other hand, it was not as helpful to more proficient students who benefited more from seeing solutions. We hypothesize that in the classroom setting, TPS was superior for less proficient students due to the fact that the higher interactivity level required from TPS better engages students' focus. This theory would suggest that students who were better able to learn from reading a solution had greater focus. In addition, the more proficient students may have more prior knowledge that prepares them to learn from reading text [7].

Our previous study controlled for the number of problems done and since TPS takes significantly more time than seeing a solution, it suffered from a "time on task" confound. We wanted to determine whether TPS was worth the extra time it took or if students would benefit from practice on more problems in the same amount of time. This becomes more pertinent as teachers are expected to cover more and more material to address all of the topics covered in standardized tests and instructional time becomes more precious.

Our hypothesis is that less proficient students in a classroom setting will benefit more from interactive tutored problem solving than from reading solutions and doing more problems. We expect that more proficient students will benefit more from reading solutions and doing more problems than less proficient students. We used the ASSISTment System, described in the next section, to test our hypothesis.

1. Brief Overview of ASSISTments

The ASSISTment[1] System [6] tutors students while providing a detailed assessment of their abilities to teachers. The ASSISTment System can identify the difficulties individual students are having, as well as difficulties the class as a whole is having, so teachers can use this detailed data to tailor their instruction to focus on the areas identified by the system. Unlike other assessment systems, the ASSISTment system also provides students with tutoring assistance while the assessment information is being collected.

An ASSISTment is the basic unit of our tutoring application. It consists of a single main question which students are asked to answer. Unlike Cognitive Tutors [1], the ASSISTment System allows students to attempt to solve a problem without going through all of the steps first. However if they get the problem wrong or ask for help, students are provided with tutored problem solving in the form of "scaffolding questions" which break down problems into steps when they need assistance. Each scaffolding question targets a single skill needed to solve the problem. Students must answer each scaffolding question in order to proceed and can ask for hints for extra help. Additionally, students receive feedback on each response telling them whether it is right or wrong and occasionally providing *buggy messages* (messages that address specific errors or misconceptions).

The structure and the supporting interface used to build ASSISTments are simple enough so that users with little or no computer science and cognitive psychology background can use it easily. Content authors can easily enter question text, hints and buggy messages by clicking on the appropriate field and typing; formatting tools are also provided for easily bolding, italicizing, etc. Images and animations can also be uploaded in any of these fields.

Most ASSISTments are based on math problems taken from the Massachusetts Comprehensive Assessment System (MCAS) tests, which is the Massachusetts state test that students take every year from grades 3 – 10. The system is primarily used by middle and high school teachers throughout Massachusetts who are preparing students for the MCAS tests.

[1] The term ASSISTment was coined by Kenneth Koedinger and blends Assisting and Assessment.

Currently, there are over 3000 students and 50 teachers using the ASSISTment System as part of their regular math classes. We have had over 30 teachers use the system to create tutoring content.

2. Experiment Design

A repeated measures design was used where each student participated in two conditions: TPS and Solutions. We designed two problem sets: 1) slope, intercept and linear equations and 2) symbolization. Students had been introduced to these topics in their regular math class before this study took place. Figure 1 shows a problem that appeared in the symbolization problem set, with the TPS approach shown on the left and the Solutions approach shown on the right. We counterbalanced to control for order effects, with each group receiving treatments in a different order. Table 1 shows the experiment design.

Table 1. Experiment design.

Session	Symbolization First Group		Slope and Intercept First Group	
	Class A	**Class B**	**Class C**	**Class D**
Session 1	Symbolization *Solutions*	Symbolization *TPS*	Slope & Intercept *TPS*	Slope & Intercept *Solutions*
Session 2	Slope & Intercept *TPS*	Slope & Intercept *Solutions*	Symbolization *Solutions*	Symbolization *TPS*

We did not use mastery learning during the study, but we believe that ASSISTments protected against a mastery learning bias since it would not matter which condition a student was in if they got problems correct: they would not see either TPS or Solutions but would proceed to the next question.

Four classes of 8^{th} grade students participated in the study, which took place over two days, resulting in a total of 88 students. On the first day, students worked for 20 minutes on one of the problem sets using one of the strategies. On the second day, students worked for 20 minutes on the second problem set using the second strategy. After 20 minutes of working on the problem set, all of the students were asked to stop working. Then they were given the post-test and asked to finish all of the problems on the post-test. Students could work on the post-test until the end of class time, approximately 20 minutes. The pre-test problems were the same as the post-test problems, although students received no feedback on the pre-test whether they answered them correctly or not.

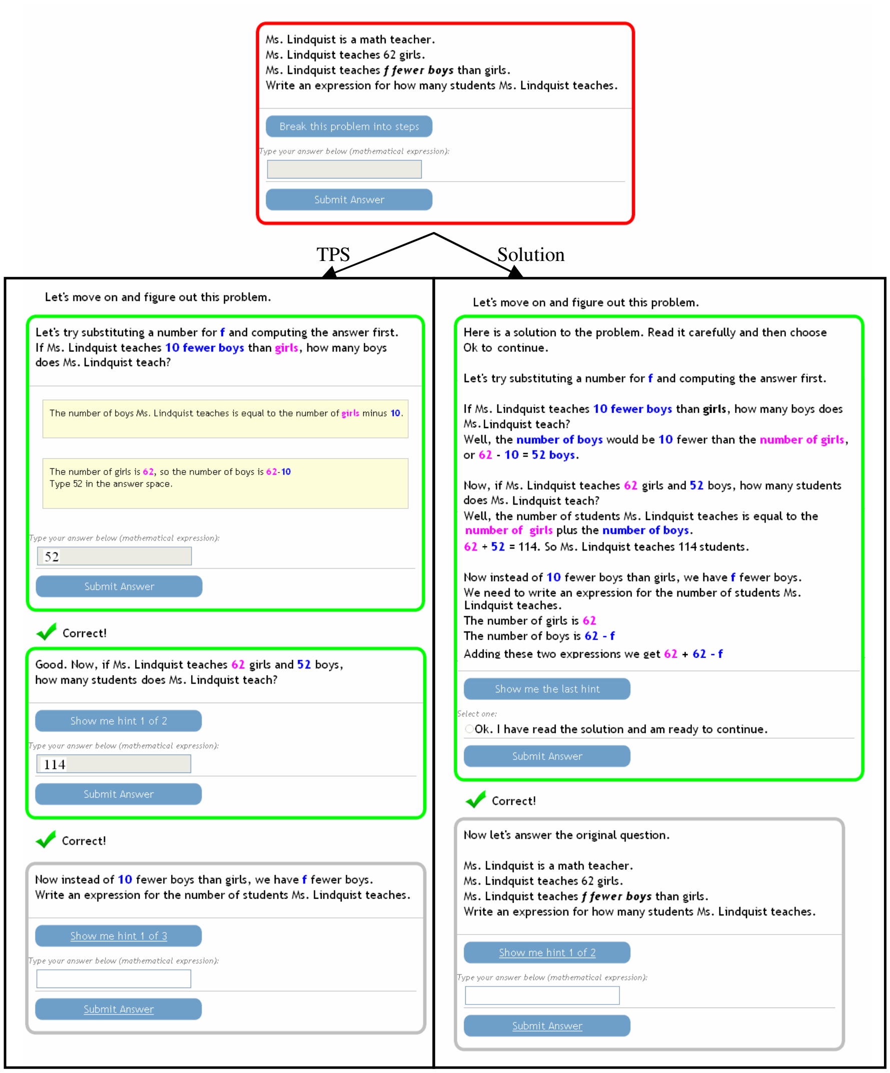


Figure 1. A symbolization problem shows the TPS approach on the left and the Solutions approach on the right.

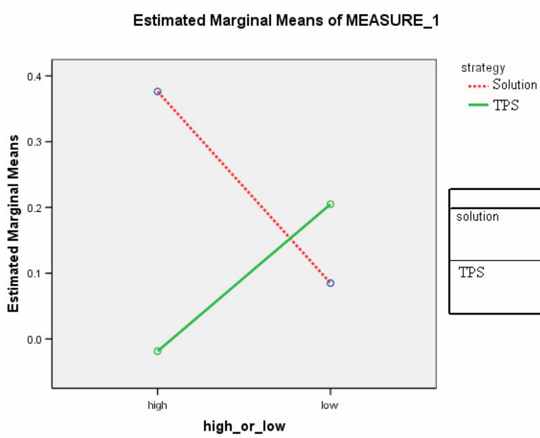
3. Results

A total of 88 eighth grade students participated in the study with 82 students completing both problem sets. We calculated a gain score for the students by subtracting their post-test scores by their pre-test scores. The slope and intercept problem set had three pre- and post-test problems and the symbolization problem set had four pre- and post-test problems. For this reason, we calculated a z-score for each student's gain score on each problem set (to compare gain scores from distributions with different means). Thus, the transformed scores have a mean of zero and a standard deviation of one. Overall, we found that there was significant learning in both problem sets ($p < 0.01$).

Students did significantly more problems with Solutions than with TPS in both problem sets. For instance, in the symbolization problem set, students using Solutions did an average of 16.57 problems and students using TPS did an average of 11.59 problems ($t(82) = 16.66$, $p < 0.001$).

We used students' performance on a practice Massachusetts Comprehensive Assessment System (MCAS) math test for 8th grade to categorize them as high proficiency or low proficiency. The practice MCAS test was given to the students as preparation for the MCAS test they will take at the end of the school year. The average score on the practice MCAS for the students who participated in this study was 56% correct. Therefore we placed students who scored greater than 56% on the practice test in the high math proficiency group and students who scored 56% or less in the low math proficiency group.

Our hypothesis was that highly proficient students would benefit more from Solutions and practice on more problems and that students with low proficiency would benefit more from TPS even though it was more time-consuming. In the following analysis, we excluded students who got all of the pre-test problems correct in one or both of the problem sets assuming that they had already learned the material (resulting in N=54). Treating the problem sets as a repeated measure, we see an interesting interaction between proficiency and condition ($p = 0.17$). Highly proficient students clearly did better when they were shown Solutions than with the TPS. Less proficient students learned slightly more from TPS than from seeing Solutions. Figure 2 shows the results of this analysis.



Descriptive Statistics

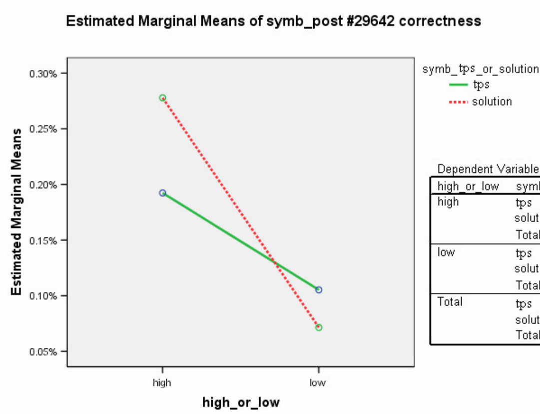
	high_or_low	Mean	Std. Deviation	N
solution	high	.375004	.9199809	28
	low	.086262	.8608142	26
	Total	.235980	.8954903	54
TPS	high	.016579	.7162554	28
	low	.167183	1.4878675	26
	Total	.089092	1.1451383	54

Figure 2. Highly proficient students appeared to do much better with Solutions, and students with low proficiency did slightly better with TPS.

We decided to look more closely at how students performed on individual problems. One problem in the symbolization pre- and post-test appeared to be more difficult than the others and no student got it correct on the pre-test. The problem reads as follows:

Linda is training for a marathon, which is a race that is ***26*** *miles long.*
Her average training time for the 26 miles is 208 minutes,
but the day of the marathon she was ***x*** *minutes* ***faster*** *than her average time.*
What was Linda's running speed for the marathon in miles per minute?
Answer: 26/(208 – x)

The interaction between proficiency and condition was significant, $F(1, 73) = .460$, $p = 0.006$. Highly proficient students learned more from reading Solutions while less proficient students learned more from TPS when working on this difficult problem.



Descriptive Statistics

Dependent Variable: symb_post #29642 correctness

high_or_low	symb_tps_or_solution	Mean	Std. Deviation	N
high	tps	.1923	.40192	26
	solution	.2778	.46089	18
	Total	.2273	.42392	44
low	tps	.1053	.31530	19
	solution	.0714	.26726	14
	Total	.0909	.29194	33
Total	tps	.1556	.36653	45
	solution	.1875	.39656	32
	Total	.1688	.37706	77

Figure 3. The interaction between high and low proficiency and condition was significant on difficult problems.

4. Contributions and Conclusions

We believe this study will aid the intelligent tutoring community in addressing the "assistance dilemma" coined by Koedinger and Aleven [4]. The assistance dilemma seeks to answer the question of how tutoring systems should balance *giving* and *withholding* information to optimize learning. Giving information can benefit students in that it is less time-consuming and students will make fewer errors [3]. However, students may find it hard to stay focused and engaged. Withholding information can help students to stay focused and engaged while helping them to generate the information on their own [2]. However, it is more time-consuming and students can make more errors that are difficult to recover from.

In this work, interactive TPS represents withholding information in an attempt to encourage students to construct knowledge themselves. Students must respond to questions and solve each step in order to proceed. They can get help and feedback on each step to help them solve the problem. On the other hand, presenting Solutions represents giving information where students are given all of the information needed to solve the problem (including the answer). Students do not have to produce any response to the solution although they are asked to read and understand how to do the problem before moving on.

In a previous study [5], we found evidence that choosing between giving or withholding information from students may depend on a student's knowledge level. However, we did have a time-on-task confound in that study because we controlled for number of problems and not for time spent. This study attempted to answer the question of when TPS is worth the extra time that it takes and who benefits most from it.

Of course, we do not claim that this study will definitively answer the assistance dilemma, but we believe it may take us a small step closer. We found that a student's math proficiency determined whether we should withhold information by presenting TPS or give information by presenting a solution to the problem. Students with high proficiency benefited from seeing Solutions whether the problem was of high or low difficulty. Students with low proficiency benefited slightly more from TPS when problems were of low difficulty, but this difference was small so we think either approach would be beneficial. When problems had high difficulty and students had low proficiency, we recommend TPS. We present our decision criteria for deciding when to switch between these two approaches in Table 2.

Table 2. Decision criteria for switching between TPS and Solutions.

	High Knowledge	**Low Knowledge**
High Difficulty	Solutions	TPS
Low Difficulty	Solutions	Solutions or TPS

The advantage of having the decision criteria shown in Table 2 is that we can attempt to optimize learning in an intelligent tutoring system by presenting the most effective and efficient approach to students determined by their knowledge level and the problem's difficulty. Students who have high proficiency would not have to waste time going through long problems step-by-step, causing them to become frustrated or bored. Students who have low proficiency may need to spend the extra time and help focusing that TPS provides.

5. Future Work

For future work, we would like to determine whether motivation plays a role in whether a tutoring system should give or withhold information. For instance, could a student who has low knowledge but high motivation benefit more from seeing a Solution rather than TPS? Do students with high knowledge and low motivation need the extra focus of TPS or would it make them even less motivated?

This study took place over a short period of time. We would also like to know how well both groups retain the information learned by giving them a delayed post-test.

Our study was limited in that we do not know if our results will generalize to other domain areas or student populations. However, we believe this study and others like it will help us to address the assistance dilemma by further characterizing the criteria under which we should switch between information giving and information withholding in general.

Acknowledgements

We would like to thank all of the people associated with creating the ASSISTment system listed at www.ASSISTment.org including investigators Kenneth Koedinger and Brian Junker at Carnegie Mellon University. We would also like to acknowledge funding from the US Department of Education, the National Science Foundation, the Office of Naval Research and the Spencer Foundation. All of the opinions expressed in this paper are those solely of the authors and not those of our funding organizations.

References

[1] Anderson, J. R., Corbett, A., Koedinger, K., & Pelletier, R. Cognitive Tutors: Lessons Learned. *Journal of the Learning Sciences* **4(2)** (1995), 167-207.

[2] Chi, M. T. H., Siler, S., Jeong, H., Yamauchi, T., & Hausmann, R. G. Learning from tutoring. *Cognitive Science*, **25** (2001), 471-533.

[3] Klahr, D. & Nigam, M. The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, **15(10)** (2004), 661–667.

[4] Koedinger, K. R., & Aleven, V. Exploring the assistance dilemma in experiments with Cognitive Tutors. *Educational Psychology Review,* **19** (2007), 239-264.

[5] Razzaq, L., Heffernan, N. T., Lindeman, R. W. (2007) What level of tutor interaction is best? In Luckin & Koedinger (Eds) Proceedings of the 13th Conference on Artificial Intelligence in Education. IOS Press. pp. 222-229.

[6] Razzaq, L., Heffernan, N. T., Koedinger, K. R., Feng, M., Nuzzo-Jones, G., Junker, B. Macasek, M. A., Rasmussen, K. P., Turner, T. E., & Walonoski, J. A. (2007). Blending Assessment and Instructional Assistance. In Nadia Nedjah, Luiza deMacedo Mourelle, Mario Neto Borges and Nival Nunesde Almeida (Eds). Intelligent Educational Machines within the Intelligent Systems Engineering Book Series. 23-49 Springer Berlin / Heidelberg.

[7] Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. Cognition and Instruction, 16(4), 475–522.

[8] VanLehn, K., Lynch, C., Schulze, K. Shapiro, J. A., Shelby, R., Taylor, L., Treacy, D.,Weinstein, A., & Wintersgill, M. The Andes physics tutoring system: Lessons Learned. *International Journal of Artificial Intelligence and Education,* **15 (3)** (2005), 1- 47.

[9] VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rose, C. P. When is reading just as effective as one-on-one interactive human tutoring? In Proceedings of the 27th Annual Meeting of the Cognitive Science Society (pp. 2259-2264). Mahwah, NJ: Erlbaum. (2005) 2259-2264.

[10] Warnakulasooriya, R. & Pritchard, D. E. (2005) Learning and problem-solving transfer between physics problems using web-based homework tutor. *Conference Proceedings: EdMedia -World Conference on Educational Multimedia, Hypermedia & Telecommunications.* Vol. 2005, 2976–2983.