Worked Examples and Tutored Problem Solving: Redundant or Synergistic Forms of Support?

Ron J. C. M. Salden (rons@cs.cmu.edu)
Human-Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213, USA

Vincent A. W. M. M. Aleven (aleven@cs.cmu.edu)
Human-Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213, USA

Alexander Renkl (renkl@psychologie.uni-freiburg.de)
Educational Psychology, Department of Psychology
University of Freiburg
Engelbergerstr. 41, D-79085 Freiburg, Germany

Rolf Schwonke (rolf.schwonke@psychologie.uni-freiburg.de)
Educational Psychology, Department of Psychology
University of Freiburg
Engelbergerstr. 41, D-79085 Freiburg, Germany

Abstract

The current research investigates a combination of two instructional approaches, tutored problem solving and worked-examples. Tutored problem solving with automated tutors has proven to be an effective instructional method. Worked-out examples have been shown to be an effective complement to *untutored* problem solving, but it is largely unknown whether they are an effective complement to *tutored* problem solving. Further, while computer-based learning environments offer the possibility of adaptively transitioning from examples to problems while tailoring to an individual learner, the effectiveness of such machine-adapted example fading is largely unstudied. To address these research questions, one lab and one classroom experiment were conducted. Both studies compared a standard Cognitive Tutor with two example-enhanced Cognitive Tutors, in which the fading of worked-out examples occurred either fixed or adaptively. Results indicate that the adaptive fading of worked-out examples leads to higher transfer performance on delayed post-tests than the other two methods.

Keywords: Cognitive Tutor, worked-out examples, adaptive fading

Introduction

Learning and cognitive skill acquisition can be supported effectively in a number of different ways. One very successful approach is the use of "tutored problem solving" by intelligent tutoring systems (Beal, Walles, Arroyo, & Woolf, 2007; Koedinger & Aleven, 2007; Koedinger, Anderson, Hadley, & Mark, 1997; Mitrovic, 2003; Razzaq et al., 2005; VanLehn et al., 2005). These systems provide individualized support for learning by doing (i.e., solving problems) by selecting appropriate problems to-be-solved, by providing feedback and problem-solving hints, and by on-line assessment of the student's learning progress. Cognitive Tutors are one particular form of intelligent tutoring systems, grounded in cognitive theory; they individualize instruction by selecting problems based on a model of the students' knowledge state that is constantly being updated (Corbett & Anderson, 1995).

Although Cognitive Tutors have many advantages, they are not without limitations. As is the case with most tutoring systems, their main focus is on correct answers during problem solving and not on gaining an understanding of how the domain principles apply in problem solving (cf. VanLehn et al., 2005).

One instructional idea to further improve the focus on principles in Cognitive Tutors, and thereby their effectiveness, is to reduce problem-solving demands by providing worked-out solutions (Renkl, 2005; Renkl & Atkinson, 2007) when the primary instructional goal is to gain understanding (cf. Sweller, van Merriënboer, & Paas, 1998). Thereby, more of the learners' limited processing capacity (i.e., working memory capacity) can be devoted to understanding the domain principles and their application in problem solving, especially when worked-out examples are combined with self-explanation prompts (cf. Renkl, 2005).

However, as learners progress through training, worked-out examples might not be as effective in later stages of the training as postulated by the expertise reversal effect (Kalyuga, Ayres, Chandler, & Sweller, 2003). The empirical results actually indicate that problem solving is more favorable in later stages, whereas worked examples are more favorable in earlier stages. This means that the effectiveness of learning methods reverses in the course of skill development. Consequently, several instructional implications should be taken into account in order to obtain effective worked-out examples. Initially, worked-out steps should be presented together with self-explanation prompts. When the learner demonstrates understanding, the worked-out steps should gradually be 'faded' from worked-out examples (solution is presented to the learner) to problems (learner must find the solution) (Atkinson, Renkl, & Merrill, 2003; Renkl, Atkinson, & Große, 2004; Renkl, Atkinson, Maier, & Staley, 2002; for a detailed theoretical rationale see Renkl & Atkinson, 2003; Renkl & Atkinson, 2007).

An important issue that has remained unaddressed until recently is whether "tutored problem solving" and worked-out examples are redundant or synergistic forms of support. On one hand, it might be that the guidance that Cognitive Tutors give to learners is so effective that embedding worked-out examples within a tutored problem solving setting would not improve learning. This kind of tutored problem solving represents a far tougher control condition than those that have been investigated in previous studies of the value of worked examples. On the other hand, it is conceivable that the two forms of instruction are synergistic: it is plausible that early cognitive skill acquisition is better supported by examples than by tutored problem solving, because examples prevent potential pitfalls such as a performance orientation, combined with the use of shallow strategies or general heuristics instead of efforts to apply domain principles in problem solving (cf. also VanLehn et al., 2005).

This issue was addressed in two recent studies by Schwonke et al. (2007), which found that tutored problem solving combined with examples that are gradually faded has beneficial learning effects. In this approach, examples are added to tutored problem solving, and are faded gradually, according to a "fixed" fading scheme that is the same for all learners. Students self-explain the example steps, as well as problem steps, with feedback from the tutor, by identifying the geometry theorem that justifies the worked-out step (see Figure 1). The results indicated that tutored problem solving combined with example fading leads to more transfer than tutored problem solving alone. Furthermore, the combination was less time consuming with no loss in terms of performance.

As suggested by Schwonke et al. (2007) the fading of examples could be even more beneficial for learning if the rate at which the worked-out steps are faded would be adapted to the students' individual learning progress. While studying and self-explaining worked-out solution steps prepares the learner to deal with subsequent problem-solving demands in a principle-based way, a learner who has not yet gained a basic understanding of a principle and of the way in which it is applied to solve problems should not be exposed to the corresponding problem-solving demands. Once the student shows a basic understanding of a principle and its application, should s/he go one step further and apply this knowledge to solve problem steps. An adaptive fading procedure will make it more likely that the student will be able to solve a faded step correctly. Such an adaptive fading method can be based on the quality of students' menu-based answers to self-explanation prompts, as a measure of their understanding of the underlying problem-solving principles.

In order to investigate whether tutored problem solving and worked-out examples are synergistic when examples are adaptively faded, three experimental conditions were compared: 1) a problem solving condition that uses the standard Cognitive Tutor; 2) an example-enhanced Cognitive Tutor that fades worked-out steps in a fixed manner; and 3) an example-enhanced Cognitive Tutor that fades worked-out steps adaptively for each individual learner. The main hypothesis states that an adaptive fading procedure, combined with tutored problem solving, will lead to better learning and higher transfer than a pure tutored problem solving procedure and a fixed non-adaptive procedure for fading examples (also combined with tutored problem solving). The three experimental conditions were addressed in both a lab setting (in Freiburg) and a classroom setting (in Pittsburgh). The main reason why the same manipulation was used in both settings lies in the question how robust effects found in a lab setting transfer to a real-life environment. By linking both lab and classroom settings a stronger and more ecologically valid investigation of the experimental methods is created and possible effects will have higher implications.

Experiment 1: Freiburg Lab Study

For this study 57 students (19 in 9th grade; 38 in 10th grade) were recruited from a German "Realschule", which is equivalent to an American high school. The participants (age M = 15.63, SD = .84) were randomly distributed across the three experimental conditions.

The experiment focused on a unit in the tutor curriculum that deals with the geometric properties of angles, covering four theorems: angle addition, separate complementary angles, vertical angles, and linear pair. Every aspect (interface, hints, and glossary) of the Cognitive Tutor was translated into German. In order to be able to implement a consistent fading procedure we created new Geometry problems covering the theorems of the selected unit. The problems were sequenced from simpler to more complex; with one-step problems presented first, followed by two step problems, and eventually by three step problems. Furthermore, in contrast to the problem solving condition, not all problems were pure problem solving in the fixed fading condition. See Table 1 for the sequencing of the problems (P1 to P11) and which theorem steps (T1 to T4) were worked-out (W) or faded (S for solving) for the fixed fading condition.

Table 1: The sequence of problems and the fading of worked-out steps in the "Fixed Fading" condition

	Problem solving				Examples			
	T1	T2	T3	T4	T1	T2	T 3	T 4
P1	S				W			
P2		S				W		
P3			S				W	
P4				S				W
P5	S		S		W		W	
P6		S		S		W		W
P7	S	S	S		W	W	S	
P8		S	S	S		S	S	W
P9	S	S		S	W	S		S
P10	S		S	S	S		S	S
P11	S	S	S		S	S	S	

For the adaptive fading condition the presentation of worked-out steps was the same as the fixed fading condition up until the three-step problems (problems 7 to 11). Once students got to those problems any value step could be presented as either pure problem solving or as worked-out, depending on the student's performance explaining worked-out steps in earlier problems that involve the same geometry theorem (see Figure 1). Thus, fading was based on the correctness of the menu-based explanations of the worked-out step, as described above and illustrated in Figure 1.

Specifically, the fading decisions were based on the tutor's estimates of each individual student's ability to produce valid self-explanations on the reason steps, for each of the four theorems. The tutor maintains these probability estimates (separately for each of the four theorems) using a Bayesian knowledge-tracing algorithm (Corbett & Anderson, 1995). The estimates are updated each time the student explains a step involving the giving geometry skill; the direction of the update depends on whether the explanation was correct or not. The knowledge-tracing algorithm is a well-established method for student modeling in intelligent tutoring systems. In prior research, Cognitive Mastery Learning built on top of Bayesian Knowledge Tracing has been shown to significantly improve student learning (Corbett & Anderson, 1995). Further, the estimates of skill mastery based on the Bayesian knowledge tracing algorithm have been shown to accurately predict students' posttest scores (Corbett & Anderson, 1995).

In the current project, in order to achieve effective fading of the worked-out steps, the estimates of an individual student's mastery of each the geometry theorems were compared against two thresholds: a high threshold and a low threshold, set at .7 and .5, respectively. The high threshold represents an estimate of the level of understanding that the students needs to attain at which a worked-out step is faded. However, even if a student attains this level of understanding s/he may fall below that level by making errors on subsequent steps of that specific theorem. Once the estimate of skill mastery falls below the low threshold the Tutor will present the student with a worked-out step

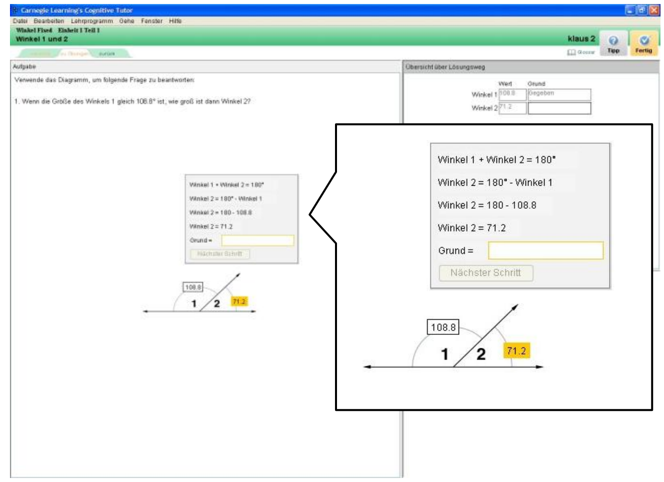

Figure 1: The circled work area shows the worked-out steps and the self-explanation to be done.

again until s/he reaches the high threshold again. In this manner, the Adaptive Fading method adapts to each individual student's level of understanding.

An example for the linear pair ("Lineares Paar") theorem is given in Figure 1 and while the value step (of "Winkel 2") is worked-out the student still has to find that theorem to explain the worked-out step. To fill in this explanation step (called "Grund") they can do either type the name of the theorem, or select the theorem from the tutor's online glossary of geometry knowledge. Figure 1 shows the "Glossar" hyperlink in the upper right corner which will open the glossary in which students can browse relevant theorems and definitions; each is described and illustrated with a simple example.

The experiment consisted of two lab sessions. Since the German students were unfamiliar with the Cognitive Tutor they received paper instructions before using the Geometry Tutor during the first lab session. Then they received a built-in untutored pre-test, the actual Cognitive Tutor training, a built-in untutored post-test plus a paper post-test. The built-in pre- and post-test consisted of the same three tutor problems during which the students did not receive feedback prompts nor could they request hints (by pressing the "Tipp" button, see Figure 1). In contrast, during the Cognitive Tutor training students did receive corrective feedback after each step they performed; for both value steps and self-explanation steps. Furthermore, during the Tutor training the students could request hints on each of these steps and often each hint had several hint levels with the final level being the bottom-out hint (i.e., containing the answer).The paper post-test consisted of three different tasks with the first task being word problems from different domains with different structures. In another task, participants had to decide whether a given problem was solvable and if so provide the principles. In a third task they had to generate real world examples for the to-be-learned principles and to illustrate that example in form of a drawing. During the second session, which occurred one week later, a delayed post-test on paper was administered which contained the same tasks as the immediate post-test. The students received 20 euro for their participation in the study.

Results

A planned contrast comparing the adaptive fading condition with the problem solving + fixed fading conditions revealed higher transfer performance for the adaptive fading condition on the regular post-test ($F(1, 54) = 5.05$, $p < .05$, $\eta^2 = .09$). This effect was replicated on delayed post-test ($F(1, 54) = 4.42$, $p < .05$, $\eta^2 = .08$). There were no differences in time spent on either of the post-tests (Fs < 1).

Experiment 2: *In Vivo* Study

The study took place at a vocational school which already uses the Cognitive Tutor for their Geometry curriculum. The participants consisted of three 9th grade classes with 51 students led by one teacher. For the assignment to the conditions the students were compiled top down with the best students on top. The first top three students were then randomly assigned to one of the experimental conditions, followed by the second three students on the list, and so on.

Since the students were already acquainted with the Cognitive Tutor we gave instructions to the teacher about the differences between the standard Cognitive Tutor and the two example-enhanced Cognitive Tutors. Overall, the materials and procedure were very similar to the German lab study with a few differences. First of all, mastery learning was on during this study, leading students to receive remedial problems for the theorems/skills they have not fully mastered yet. Secondly, since the school where the study took place uses the Geometry Cognitive Tutor as part of their regular geometry instruction, the study covered more material and had a longer duration.

The study comprised all five sections in the tutor curriculum that deal with the geometric properties of angles, including the unit that was used in the Freiburg study. New problems were developed for all units, as our fading procedure required problems that involve particular skill combinations. Over a period of three weeks, the students worked with the Cognitive Tutor for two hours per week, each according to the condition s/he was assigned to.

Furthermore, online pre- and post-tests were administered to the participants. These tests were created with the Cognitive Tutors Authoring Tools (CTAT) and presented students with problems covering the same Angles theorems as they learned in the Cognitive Tutor. The pre-test and immediate post-test contained the same ten transfer problems of which eight problems were transfer problems (problem solving items) in which the students needed to indicate whether the step was solvable, and if so then find out the value, its corresponding theorem and to which angles it applied to. The remaining two problems were transfer problems (composer items) were students were presented with an image and given values for a few angles. From this information the students needed to extract what angles they could find within just one step. More specifically, with each angle they found another angle might be found within one step.

In addition to the immediate post-test a delayed post-test was administered three weeks after the students finished working on the Cognitive Tutor. This post-test contained six transfer problems of which four were problem solving items and two were composer items. Since the Angles Unit is part of their regular curriculum no participant fee was given.

Results

It should be noted that a considerable amount of attrition occurred throughout the study which explain the varying degrees of freedom in the analyses. Of the 51 students only 20 completed all three CTAT tests. Furthermore, only 28 students completed both the pre-test and the immediate post-test. For the analysis on the delayed post-test scores the criterion was set to include those students who not only completed the delayed post-test but also completed at least one other test ($N = 35$).

Using students who completed both pre-test and regular post-test ($N = 28$) it was shown that significant learning occurred throughout all conditions from pre-test ($M = 15.46$, $SD = 14.01$) to post-test ($M = 22.93$, $SD = 16.64$; $t(27) = 2.27$, $p < .05$, $d = .87$; cf. Cohen, 1988). The planned contrast of adaptive fading condition versus the problem solving + fixed fading conditions revealed no differences in performance either on the pre-test or on the regular post-test (Fs < 1). Furthermore, while the planned contrast did not show an effect of the adaptive fading condition ($M = 12.80$, $SD = 5.61$) over the other two conditions ($M = 8.73$, $SD = 4.97$) on the delayed post-test ($F = 2.38$, $p = .11$), it did indicate a tendency in the expected direction ($t(32) = 2.10$, $p < .05$, $d = .74$). When excluding the fixed fading condition, the adaptive fading condition ($M = 11.64$, $SD = 6.58$) did attain higher transfer performance on the delayed post-test than the problem solving condition ($M = 6.93$, $SD = 5.21$; $t(20) = 2.15$, $p < .05$, $d = .96$). Interestingly no differences in Cognitive Tutor time ($F < 1$) were found. Lastly, the results show that the overall number of worked-out example steps between the fixed fading and adaptive fading conditions is fairly close to each other (F < 1).

Discussion

Two studies were conducted comparing “standard” tutored problem solving with a Cognitive Tutor versus two conditions in which tutored problem solving was enriched with worked-out examples. The worked-out examples were faded in either a fixed or in an adaptive manner. These manipulations were tested both in a lab study and in an actual classroom setting as part of a regular high school curriculum. The results of the lab study show that adaptively fading worked-out examples leads to higher transfer performance on both regular post-test *and* delayed post-tests. While this effect was not fully replicated in the classroom study, a significant benefit in transfer performance for the adaptive fading condition over the problem solving condition was revealed on the delayed post-test.

A likely explanation for the lesser effect in the classroom study can be found in the larger amount of noise that inherently exists within a real life environment, as compared to the laboratory. Also, the classroom study took place over a longer period of time with a fairly high amount of attrition of students. More specifically, a considerable number of students missed one (some even missed two) of the three online tests that were given. Yet despite the general difficulty of replicating lab results in the classroom, the current study still shows a benefit of the adaptive fading condition.

Another explanation could be the use of the Cognitive Tutor's mastery learning criterion which led students in the classroom study to receive remedial problems for the theorems they had not mastered fully yet (on an individual basis, after they completed the problem sequence described in Table 1). These remedial problems represent additional learning opportunities for students. (The mastery learning mechanism was turned on in all three conditions in the *in vivo* study.) It could be that the mastery learning mechanism caused the students' knowledge level to be more equal than the students in the lab experiment who did not receive any remedial problems. If the mastery learning criterion made students exit with similar exit knowledge of the theorems then similar performance can be expected on the immediate post-test. The results indicate that a possible equalizing effect of mastery due to mastery learning did wear off over time, since the adaptive fading condition attained higher delayed post-test performance than the tutored problem solving condition. In other words, even with mastery learning on, a benefit of worked examples is seen.

The current findings confirm and extend the findings of Schwonke et al. (2007) which indicated that tutored problem solving, combined with fixed fading of worked-out steps, leads to better transfer performance, as well as to more efficient learning. A tentative explanation might be that working with examples increases students' procedural *and* conceptual knowledge compared to tutored problem solving without examples.

It is interesting to view the current findings in light of the Assistance Dilemma issue that was recently raised by Koedinger and Aleven (2007). The Assistance Dilemma states that the balance between giving assistance to students and withholding it (while letting students generate information by themselves, possibly with feedback) exerts a major influence on students' learning. The choice between worked examples and problems is a key manifestation of this dilemma: how should a tutor effectively switch from a "high assistance" form of instruction (i.e., worked examples) to a "low assistance" form of instruction (i.e., problem solving) in a manner that is adaptive to individual students' needs? The present results show that an adaptive example fading method, in which the rate of fading is based on the quality of students' self-explanations, is a promising way to make this determination, in a manner adaptive to students' individual learning trajectories.

In short, the results of both studies indicate that the implementation of an adaptive fading procedure of worked-out examples within a Cognitive Tutor can be useful in both lab and actual classroom environments. Tutored problem solving and worked examples, adaptively faded, are *synergistic*, not redundant, forms of support.

Acknowledgements

This work was supported in part by the Pittsburgh Science of Learning Center which is funded by the National Science Foundation award number SBE-0354420.

References

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Combining fading with prompting fosters learning. *Journal of Educational Psychology, 95*, 774-783.

Beal, C. R., Walles, R., Arroyo, I., & Woolf, B. P. (2007). Online tutoring for math achievement: A controlled evaluation. *Journal of Interactive Online Learning, 6*, 43-55

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4*, 253-278.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*, 23-32.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8*, 30-43.

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

Mitrovic, A. (2003). An Intelligent SQL Tutor on the Web, *International Journal of Artificial Intelligence in Education, 13*, 171-195.

Razzaq, L., Feng, M., Nuzzo-Jones, G., Heffernan, N. T., Koedinger, K. R., Junker, B., Ritter, S., Knight, A., Aniszczyk, C., Choksey, S., Livak, T., Mercado, E., Turner, T. E., Upalekar. R, Walonoski, J. A., Macasek. M. A., Rasmussen, K. P. (2005). The Assistment Project: Blending Assessment and Assisting. In C.K. Looi, G. McCalla, B. Bredeweg, & J. Breuker (Eds.) *Proceedings of the 12th International Conference on Artificial Intelligence In Education*, (pp. 555-562). Amsterdam: ISO Press.

Renkl, A. (2005). The worked-out-example principle in multimedia learning. In R. Mayer (Ed.), *Cambridge*

Handbook of Multimedia Learning (pp. 229-246). Cambridge, UK: Cambridge University Press.

Renkl, A. & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skills acquisition: A cognitive load perspective. *Educational Psychologist, 38,* 15-22.

Renkl, A., & Atkinson, R. K. (2007). An example order for cognitive skill acquisition. In F. E. Ritter, J. Nerb, E. Lehtinen, T. O'Shea (Eds.), *In order to learn: How the sequence of topics influences learning* (pp. 95-105). New York , NY : Oxford University Press.

Renkl, A., Atkinson, R. K., & Große, C. S. (2004) How fading worked solution steps works - a cognitive load perspective. *Instructional Science, 32*, 59-82.

Renkl, A., Atkinson, R. K., Maier, U. H., & Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *Journal of Experimental Education, 70*, 293-315.

Schwonke, R., Wittwer, J., Aleven, V., Salden, R., Krieg, C., & Renkl, A. (2007). Can tutored problem solving benefit from faded worked-out examples? In S. Vosniadou, D. Kayser, & A. Protopapas (Eds.), *Proceedings of EuroCogSci 07. The European Cognitive Science Conference 2007* (pp. 59-64). New York, NY: Erlbaum.

Sweller, J., Merriënboer, J. J. G. van, & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Van Lehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, A., Taylor, D., Weinstein, A., & Wintersgill, M. (2005). The Andes physics tutoring project: Five years of evaluations. *International Journal of Artificial Intelligence in Education, 15*, 1-47.