Research Reports

Institute of Psychology

University of Freiburg

Germany

Research Report No. 140

From Example Study to Problem Solving: Smooth Transitions Help Learning

Alexander Renkl, Robert K. Atkinson & Uwe H. Maier

February 2000

Correspondence

Prof. Dr. A. Renkl, Department of Educational Psychology, University of Freiburg, Belfortstr. 16, D-79085 Freiburg, Germany.

Email: renkl@psychologie.uni-freiburg.de

Abstract

Research has shown that it is effective to combine example study and problem solving in the initial acquisition of cognitive skills. Present methods for combining these learning modes are, however, static and do not support a transition from example study in early stages of skill acquisition to later problem solving. Against this background, we propose a successive integration of problem-solving elements into example study until the learners solve problems on their own (i.e., complete example ➔ increasingly more incomplete examples ➔ problem to-be-solved). We tested the effectiveness of such a fading procedure against the traditional method of employing example-problem pairs. In a field experiment and in a more controlled lab experiment, we found that the fading procedure fosters learning, at least when near transfer performance is considered. Moreover, this effect is mediated by a lower number of errors under the fading condition as compared to the example-problem condition.

FROM EXAMPLE STUDY TO PROBLEM SOLVING: SMOOTH TRANSITIONS HELP LEARNING

Worked-out examples consist of a problem formulation, solution steps and the final solution itself. Research has shown that learning from such examples is of major importance for the initial acquisition of cognitive skills in well-structured domains such as mathematics, physics, and programming (for an overview see VanLehn, 1996). In addition, this learning mode is preferred by novices, and they are right: It is quite an effective way of learning. Studies performed by Sweller and his colleagues (e.g., Sweller & Cooper, 1985; for an overview see Sweller, van Merrienboër, & Paas, 1998) showed that learning from worked-out examples can be more effective than pure learning by problem solving.

Although worked-out examples have significant advantages, their employment as a learning methodology does not, of course, guarantee effective learning. First, the extent to which learners profit from the study of examples depends heavily on how well they explain the solutions of the examples to themselves (Chi, in press; Chi, Bassok, Lewis, Reimann, & Glaser, 1989). Second, it is important how the learning materials (examples and problems) are structured (cf. Atkinson, Derry, Renkl, & Wortham, under review). The second aspect is the focus of this study. More specifically, this study investigates one possible approach to integrating elements of problem solving into example study. We propose that these learning modes can be combined by successively introducing more and more elements of problem solving in example study until learners are solving the problems on their own. This rationale can also be used as a way to structure the transition from studying examples in initial skill acquisition to problem solving in later phases of the learning process.

In the next section, the literature with respect to the issue of combining example study and problem solving is discussed. Then we outline open questions and give preliminary answers that were tested in two studies, first in a field experiment and then in a more controlled lab experiment.

How to Combine Example Study and Problem Solving? – State of the Art

Empirical evidence has shown that pure example study is not as effective as learning from examples in which elements of problem solving are integrated. There are two traditional ways in which this can be accomplished: (1) Making the solutions of examples incomplete and (2) employing example-problem pairs.

Incomplete examples

Some researchers argue that incomplete examples, which the learners have to complete, effectively support the acquisition of cognitive skills (van Merrienboër, 1990; van Merrienboër & de Crook, 1992; Paas, 1992; Stark, 1999). Stark (1999) conducted a controlled experiment designed to examine the extent to which the insertion of “blanks” into the solution of examples—which, in a certain sense, forced the learners to determine the next solution step on their own—fostered learning. In his study, half of the participants studied incomplete examples (experimental group), while the other half learned from complete examples (control group). In the experimental group, portions of the example solutions presented to the participants were replaced by "question marks." The learners were then asked to identity what solution step was missing. After doing that, or at least making the attempt, the complete solution step was presented so that learners received feedback on the correctness of their anticipation. When compared to studying complete examples, Stark found that incomplete examples fostered the quality of self-explanations and, as a consequence, the transfer of learned solution methods. The results of Stark's study contrast with observations by Paas (1992), who did not find any difference in performance among participants presented with either incomplete or complete examples. However, the main purpose of Paas' study was not to investigate the effects of complete versus incomplete examples. Taken together, the results of Stark (1999) show that making examples incomplete (at least) can support learning.

Example-problem pairs

Sweller and his colleagues (e.g., Sweller & Cooper, 1985; Mwangi & Sweller, 1998) have conducted several classic studies documenting the effectiveness of learning from worked-out examples. However, in these studies the authors did not compare *pure* learning from examples with *pure* problem solving. Instead, the example condition usually consisted of examples followed by isomorphic problems to-be-solved (example-problem pairs). Thus, the studies of Sweller and colleagues mainly showed that combined learning from examples and problems is more effective than pure learning by solving problems.

Studies on learning from worked-out examples performed by other researchers have focussed on pure learning from examples (e.g., Renkl, 1997). Explicit comparisons between pure example learning and learning from example-problem pairs are, however, rare. One such study was performed by Trafton and Reiser (1993), in which the authors designed two treatments, alternating and blocked: Participants in the alternating condition were exposed to six example-problem pairs, where each example was followed directly by a isomorphic problem, while participants in the blocked condition were exposed to the entire set of six examples, followed by the entire set of six practice problems. The authors found that, as predicted, participants in the alternating-example condition took less time and produced more accurate solutions on the transfer posttest than their counterparts in the blocked-example condition. Based on these findings, the authors asserted that "the most efficient way to present material to acquire a skill is to present an example, then a similar problem to solve immediately following" (Trafton & Reiser, 1993, p. 1022).

In a recent study, Stark, Gruber, Renkl, and Mandl (in press) examined whether there might be another more effective variation of the traditional method of pairing examples with practice problems. Based on a study of learning diagnostic strategies in medicine in which it was found that it is more effective to learn from a "cognitive model" (which can also be regarded as a kind of worded-out example) after an initial problem solving experience (Gräsel & Mandl, 1993), the authors argued that presenting practice problems first followed by isomorphic examples (problem-example pairs) should be an effective mode of instruction. Specifically, the authors proposed that initial problem solving difficulties should motivate the learners to process the examples that followed more deeply and, in particular, more focussed with respect to the specific difficulties the individual learners have in solving such problems. In a comparison between pure example learning and learning from problem-example pairs (domain: calculation of compound and real interest), it was found that the combined learning method (i.e., problem-example pairs) substantially fostered active example processing and, as a result, learning outcomes.

Taken together, combining practice problems and examples is obviously more effective than exposing learners to either of the two pure learning conditions, that is, either to sets of practice problems or sets of examples.

Open Questions and Answers to be Validated

Although there can be little doubt on the effectiveness of a combined learning method, two questions still remain open: (2) Are there more effective ways of combining example study and problem solving than presenting incomplete examples or pairs of examples and problems? (2) What is a sensible rationale for designing the transition from learning from examples in initial stages of cognitive skill acquisition to problem solving in later stages?

Instructional models such as *Cognitive Apprenticeship* (Collins, Brown & Newman, 1989) propose a smooth transition from modeling to scaffolded problem solving to independent problem solving in which instructional support fades during the transition. The use of incomplete examples, at least as realized in previous studies, has not incorporated such a dynamic fading component. To date, studies incorporating the "pairs arrangement" have also not used a fading component. In fact, these studies typically contain abrupt transitions from examples, as a type of models, to independent problem solving. Against this background, it is sensible to combine problem solving and example study in the following way. First, a complete example is presented (model). Second, an example is given in which one single solution step is omitted (scaffolded problem solving). Then, the number of blanks is increased step-by step until just the problem formulation is left, that is, a problem to-be-solved (independent problem solving). In this way, a smooth transition from modeling (complete example) over scaffolded problem solving (incomplete example) to independent problem solving is implemented.

This rationale also provides one possible answer to the second open question that was outlined above: The transition from example study in early phases of skill acquisition to later problem solving, which is—without doubt—the activity that should be learned eventually.

Experiment 1: Field Experiment

As a first test of our assumptions we conducted a small-scale field experiment in which we tested whether a smooth transition from example study to problem solving (gradual insertion of blanks into the solutions of examples) is more effective than learning by example-problem pairs as they are used in many studies on learning from examples.

Methods

Sample and Design. Two ninth-grade classrooms from a German *Hauptschule* (lowest track of the German three track system) participated in this quasi-experiment. In both classrooms, the same teacher (third author) conducted a physics lesson on electricity based on four examples/problems. In one classroom (n = 20) a fading procedure was used and in the other classroom (n = 15) traditional example-problem pairs were employed. Each example/problem involved three solution steps. Across both conditions half of the steps were worked-out whereas the other half was to be generated. Thus, learners in both conditions were required to solve the same number of solution steps.

Learning environment. In the experimental phase, the third author (a professional teacher) conducted a 45 minute lesson in each classroom. Both groups worked on four examples/problems in which the cost for running a variety of electric devices for a certain time had to be determined (e.g., "A aluminum factory has a big melting furnace which is run with 1000 V. A power of 20 A has to flow through the furnace in order to melt aluminum. What does the factory have to pay per month when the furnace always runs and the kWh costs DM 0.22?"). Although the examples/problems were printed on work sheets, the problem formulation of each example/problem was read aloud by one of the students from the class. Following the reading of the problem formulation, the students were permitted ask clarifying questions (of course, no questions on the solution) before working individually on the example or problem. At the end of each incomplete example or problem, the complete solution was presented on an overhead transparency and, if necessary, the students corrected or supplemented their solutions. Then the teacher proceeded to the next example/problem.

In the fading classroom, the teacher presented the instruction in the following order: (1) a complete example, (2) an example with the last solution step left out, (3) an example with the last two steps omitted, and (4) a problem where all three steps were missing ("backward rationale"). In the example-problem group in contrast, a complete example was presented twice, each time it was followed by a corresponding problem.

Procedure. The overall procedure was identical in both classrooms. Basic knowledge of the concepts and rules of electricity was introduced in the context of regular instruction followed by a pretest that tapped into prior knowledge with respect to the ability to apply the abstract rules to domain problems. Two days later, the school lessons in which the experimental variation took place were conducted. Finally, after additional two days, the students worked on a posttest.

Instruments. The pretest consisted of four problems from the physics domain of electricity that were structurally equivalent to the problems in the posttest (e.g., "The electronic motor of an electronic locomotive is supplied by a voltage of 0.6 kV. In the average, a current of 18 A flows through the motor. What does an eight-hour trip from Stuttgart to Hamburg cost when you assume that the *German Railway* pays DM 0.12 per kWh?"). For the correct solution of an item, a maximum of three points could be gained. For partly correct solutions partial credit was dispensed. The score was divided by the theoretical maximum score (12) so that it represent the percentage of points in relation to perfect performance. The pretest had a sufficient reliability (Cronbach's Alpha: .87).

The posttests consisted of six problems. Four near-transfer problems had the same underlying structure (solution rational) as the examples and problems employed in the learning phase but different surface features (cover story, numbers). Two problems were classified as far transfer because both the underlying structure and the surface features differed (e.g., "Tanja pays for her frig 40 DM per year. One kWh costs DM 0.22. What power does the frig have if you assume that it runs all the time?"). For the correct solution on a posttest problem, which always included three solution steps, three points were dispensed. Partial credit was given for partly correct solutions (1 or 2 points). The scores for both scales were finally divided by the theoretical maximum score (12 or 6 respectively) so that they represented the percentage of points in relation to perfect performance. We obtained sufficient reliabilities (Cronbach's Alphas) for both posttest scales: .85 for near transfer and .60 for far transfer.

Results

Table 1 shows the means and the standard deviations of the two experimental groups on the pretest and the posttest scores. Both groups showed almost identical pretest performance ($t(33) = 0.01$; $p > .10$). Hence, there was no *a priori* difference between groups with respect to prior knowledge.

Table 1

Group means (standard deviations in brackets) of pretest and posttest scores

	Fading	Example-problem pairs
Pretest	24.06 (28.12)	23.96 (29.02)
Posttest: near transfer	79.38 (27.42)	62.22 (24.82)
Posttest: far transfer	36.25 (37.29)	21.11 (27.61)

With respect to treatment effects we descriptively obtained higher means in the fading group for both near and far transfer. Comparisons between the experimental conditions by means of an ANCOVA (controlling for prior knowledge) yielded a significant difference for near transfer performance ($F(1,32) = 4.44$; $p < .05$). But, the group difference in far transfer performance failed to reach the level of significance ($F(1,32) = 2.28$; $p > .10$). Thus, the fading procedure clearly fostered near transfer performance. We can not, however, claim that this is also true for far transfer performance.

Discussion

We obtained a positive effect of our fading procedure with respect to near transfer performance. The far transfer performance was also superior in the fading group, but not at the level of statistical significance. Before theorizing on possible reasons for potential differential effects of the fading procedure on near and far transfer, we should wait and see whether the respective finding can be replicated.

A replication is necessary because a field study such as the present one always has some factors that might diminish the internal validity of the findings. For example, the teacher that conducted the instruction in both classrooms was not "blind" with respect to the experimental expectations. Furthermore, the present investigation was "merely" a quasi-experiment (no random assignment of participants to the experimental conditions). Hence, the conditions in both classrooms might not have been totally identical except for the independent variable (fading vs. example-problem pairs). Finally, no data on possible processes that mediate the effects of the fading procedure on the learning outcomes were recorded. These issues were addressed in Experiment 2.

Experiment 2: Lab Experiment

In order to conceptually replicate the results of the preceding field experiment under more controlled conditions, we ran a lab experiment. We also tested for one possible mediating mechanism that may explain the effect found in Experiment 1.

As outlined above, there are quite abrupt changes with respect to the demands placed on the learners in the example-problem conditions. After a first example, the learners have to solve a whole problem totally on their own. In the fading procedure, the first problem solving demand is to generate just a single step, and the demands are only gradually increased. Against this background, we expect that the learners will make fewer errors during learning in the fading condition. In the example-problem condition, in contrast, we expect a relatively high number of errors during learning that may prevent rapid learning progress. This assumption was tested in Experiment 2.

Methods

Sample and Design. The participants of this study were 54 students of psychology (Mississippi State University). They were randomly assigned to the fading or to the example-problem condition, respectively (n = 27 in each group). As with our field experiment, the number of unsolved solution steps was held constant across both conditions.

Learning environment. A computer-based learning program was employed that had been originally developed by Renkl (1997), modified by Stark (1999), and finally adapted to the present needs by the second author. It presented worked-out examples and problems from the domain of probability calculation (e.g., "Jonathan has recently bought a new camera. Independently of each other he frequently makes two errors when he takes a picture. He manages to blur the image in 40% of his photos (p=2/5) and he forgets to activate the flash in 10% of the photos (p=1/10) so that the pictures end up too dark. If you randomly choose one of Jonathan's developed pictures, what is the probability that it will be flawless?"). The examples/problems were displayed in a step-by-step procedure. On the first page of an example/problem, the problem givens were displayed. The learners could read them and then go to the next page where a first solution step was presented or the learners were required to determine a solution step on their own (or at least to attempt it). After inspecting or determining this solution step, the participants proceeded to the following page where the next solution step was added or required, and so on. When the whole solution of a problem was presented or required, the next page contained the first page of a new example/problem until the lesson was completed. In the case of omitted solution steps, the learners had to type in a solution attempt, otherwise they could not proceed. Hence, the correctness of the problem solving attempts could be determined. Note that the correct step was always displayed when the learners went to the next page so that there was feedback on the correctness of the learners' problem solving attempts.

On the whole, there were two sets of four probability tasks. Each set consisted of four tasks with the same underlying structure (solution rationale) but different surface features (cover stories, numbers). In the fading group, the first task was a completely worked-out example. In the second task, the first solution step was omitted. In the third task the first two steps were omitted ("forward rationale" of omitting solution steps). The fourth task was essentially a problem-solving task (all three steps were missing). In the example-problem group, two such pairs were presented.

Procedure. The participants worked in group sessions lasting about 90 minutes. During this time they worked individually in front of a computer. First, a pretest on prior knowledge in probability calculation was presented. In order to provide or re-activate basic knowledge that allowed the participants to understand the worked-out examples, an instructional text on basic principles of probability calculation was given to the participants. After reading this instructional text, the participants were to study the worked-out examples and problems provided by the computer program. In this phase, the experimental variation took place (fading vs. example-problem pairs). The time spent for learning was recorded. Finally, the participants worked on a posttest.

Instruments. A *pretest* was employed in order to assess prior knowledge. It consisted of nine relatively simple problems involving probability calculation (e.g., "When rolling a 6-sided die what is the probability that '2' or '4' will appear?"). For each correct solution, one point was dispensed (no partial credit). The overall score was divided by the theoretical maximum score (9) so that it represents the percentage of points in relation to perfect performance. We obtained a sufficient reliability of .73 (Cronbach's Alpha).

The learning outcomes were assessed by a *posttest* that included thirteen problems. Besides one very simple warm-up problem, which was ignored for further analysis, we employed six near transfer items and six far transfer items. As compared to the examples / problems studied during the learning phase, the near transfer problems had the same underlying structure (solution rationale) but different surface features (cover story, numbers; e.g., "While preparing a batch of rolls at the local bakery, the baker’s assistant forgot to add salt to 30% of the rolls and, independent of this event, he burned 40% of the rolls. If the head baker arrives to examine the quality of his assistant’s work by randomly testing a roll, what is the probability that it is edible; that is, that it has the right amount of salt and is not burned?"). Far transfer problems differed with respect to both structure and surface features (e.g., "When driving to work, Mrs. Fast has to pass the same traffic light twice—once in the morning and once in the evening. It is green in 70% of the cases. What is the probability that she can pass through a green light in the morning but has to stop in the evening?").

For the totally correct solution on a posttest problem, which always included three solution steps, three points were dispensed. Partial credit was provided for partially correct solutions (1 or 2 points). The scores for both scales were finally divided by the theoretical maximum score (18) so that they represent the percentage of points in relation to perfect performance. We obtained sufficient reliabilities (Cronbach's Alphas) for both posttest scales: .91 for near transfer and .75 for far transfer.

Results

Table 2 shows the means and the standard deviations of the two experimental groups for the pretest (prior knowledge), the time spent for studying the examples and problems (learning time), the proportion of correct solutions steps generated during learning, and posttest performance with regard to near transfer and to far transfer. The small difference between the pretest scores in favor of the example-problem group was not statistically significant ($t(52) = -0.49$; $p > .10$). Hence, the groups were *a priori* comparable with respect to prior knowledge. In addition, the learning time did not significantly differ between groups ($t(52) = 0.28$; $p > .10$). Thus, possible group differences with respect to learning could not be simply attributed to time-on-task.

Table 2
Group means (standard deviations in brackets) of the pretest, the learning time (min.), the correctness of solution steps during learning (in %), and the posttest.

	Fading	Example-problem pairs
Pretest	55.56 (23.67)	58.85 (25.93)
Learning time	31.15 (10.83)	30.37 (9.41)
Correctness of solution steps	66.42 (31.61)	51.81 (33.13)
Posttest: near transfer	53.91 (32.24)	43.83 (35.35)
Posttest: far transfer	38.68 (25.25)	43.42 (24.60)

With respect to treatment effects, we descriptively obtained substantially higher means in the fading group for the proportion of correct solution steps and for near transfer. We used an ANCOVA (controlling for prior knowledge) to make comparisons between the experimental conditions that yielded a significant difference for near transfer performance ($F(1,51) = 4.58$; $p < .05$), but not for far transfer ($F < 1$). A third ANOCA revealed that there was also a significant difference between groups with respect to the proportion of correct solution steps ($F(1,51) = 7.62$; $p < .05$).

In order to test the mediation hypothesis that fading fosters learning outcomes (at least near transfer) because less errors occur during learning, an additional ANOCA for near transfer performance was performed in which the proportion of correct solution steps was included as covariate in addition to prior knowledge. The mediation hypothesis would have been confirmed if the group effect (more or less totally) disappeared in this case (cf. Baron & Kenny, 1986). This proved to be true. The *F*-value for the group effect was not only smaller than 1, but was a negligible size of 0.23.

Discussion

In the present lab experiment, we conceptually replicated the effectiveness of our fading procedure for near transfer. Both studies also yielded consistent results with respect to far transfer: No significant effect was found. We obtained these converging results even though the present study and our first investigation differed with respect to the type of learners ("low-track" students vs. university students), the learning domain (physics/electricity vs. mathematics/probability calculation), the learning setting (school lesson vs. computer-based learning in the lab), and the kind of fading out worked-out solution steps ("backward" vs. "forward"). We interpreted the stability of the findings despite these very different context conditions as an indicator that our fading procedure has a reliable effect.

Something that we did not expect in advance is that the effect of fading is restricted to near transfer. This differential effectiveness of the fading procedure may have something to do with the mediating mechanism that was identified in this study (amount of errors during learning). The analyses showed that the effect on near transfer is more or less totally mediated by the amount of errors committed during learning. Although we did not directly assess self-explanations, this result suggests that the fading procedure did not enhance learning outcomes via fostering self-explanation quality. This also helps to explain the differential effectiveness of fading. For far transfer performance (e.g., Renkl, 1997; see also Atkinson et al., under review), it is of special importance that the learners explain to themselves the rationale of solution steps in an active way so that they become aware of how domain principles can be applied in a domain and how certain goals can be achieved by certain operators. In other words, reflection about the more general aspects of specific problem solutions is necessary for far transfer. However, this process was obviously not elicited by the fading procedure. "Error-avoiding" instructional procedures such as Direct Instruction or drill-and-practice tutorials are known to effectively foster "low-level" level learning (near transfer). As our fading procedure is a method of avoiding errors during learning, it is understandable why it fosters "merely" near transfer performance.

General Discussion

In the present study, the effectiveness of our fading rationale for designing the transition from example study to problem solving has been affirmed in an highly ecologically valid field experiment as well as in a well-controlled lab study. Thus, we have provided strong evidence that a fading procedure actually fosters near transfer. Nevertheless, there are at least three important questions left that should be addressed in further research:

(1) The results indicate that the effects of fading are more or less totally mediated by the low amount of errors during learning and not by the way in which the examples were processed (self-explanations). In order to obtain more direct evidence for this interpretation, self-explanations should be assessed in a subsequent study on fading in example-based learning. In such a study, the mediation effect involving the amount of errors should be replicated and it should be tested whether there are, as expected, no differences with respect to self-explanations.

(2) In the effort to successively optimize learning from worked-out examples, another issue related to self-explanations should be addressed. If it is true, as argued above, that the quality of self-explanations is especially important for far transfer, it should be tested whether a combination of fading and self-explanation training—such as the one developed and evaluated by

Renkl, Stark, Gruber and Mandl (1998)—can facilitate *both* near and far transfer learning.

(3) We employed two ways of fading out worked-out solution steps, a backward and forward procedure. As the context conditions in our two studies varied substantially, we could not compare the relative effectiveness of these two procedures. In addition, it may well be that other procedures are even more effective. For example, one could first omit the solution step that is the easiest one for the learners to determine, then the second easiest one and so on. Systematic experimentation on this issue is necessary in order to get information on whether different ways of fading have substantially different effects and, if so, which way of fading is the ideal one.

Taken together, this contribution has provided strong evidence for the effectiveness of our "new" rationale for the integration of example study and problem solving. However, in order for us to deeply understand the way this works and to optimize the employment of this rationale, further experiments are necessary.

References

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. W. (under review). *Fostering learning from worked examples: Knows and unknows.*

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173-1182.

Chi, M. T. H. (in press). Self-explaining: The dual process of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in Instructional Psychology*. Mahwah, NJ: Erlbaum.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145-182.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction*. Hillsdale, NJ: Erlbaum.

Gräsel, C., & Mandl, H. (1993). Förderung des Erwerbs diagnostischer Strategien in fallbasierten Lernumgebungen [Promoting the acquisition of diagnostic strategies in case-based learning environments]. *Unterrichtswissenschaft, 21,* 355 – 369.

Merrienboër, J. J. G. van (1990). Strategies for programing instruction in high school: program completion vs. program generation. *Journal of Computing Research, 6,* 265-285.

Merrienboër, J. J. G. van, & De Crook, M. B. M. (1992). Strategies for computer-based programming instruction: Program completion vs. program generation. *Journal of Educational Computing Research, 8,* 212-234.

Mwangi, W., & Sweller, J. (1998). Learning to solve compare word problems: The effect of example format and generating self-explanations. *Cognition and Instruction, 16,* 173-199.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology, 84,* 429-434.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science, 21,* 1-29.

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology, 23,* 90-108.

Stark, R. (1999). *Lernen mit Lösungsbeispielen. Einfluß unvollständiger Lösungsbeispiele auf Beispielelaboration, Motivation und Lernerfolg* [Learning by worked-out examples. The impact of incomplete examples on example elaboration, motivation, and learning outcomes]. Bern, Switzerland: Huber.

Stark, R., Gruber, H., Renkl, A., & Mandl, H. (in press). Instruktionale Effekte einer kombinierten Lernmethode: Zahlt sich die Kombination von Lösungsbeispielen und Problemlöseaufgaben aus? [Instructional effects of a combined learning method: How effective is the combination of worked-out examples and problems to-be-solved?] *Zeitschrift für Pädagogische Psychologie*.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2,* 59-89.

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

Trafton, J. G., & Reiser, B. J. (1993). The contributions of studying examples and solving problems to skill acquisition. In M. Polson (Ed.), *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

VanLehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology, 47,* 513-539.