Transitioning From Studying Examples to Solving Problems: Effects of Self-Explanation Prompts and Fading Worked-Out...

Article *in* Journal of Educational Psychology · December 2003

DOI: 10.1037/0022-0663.95.4.774

Alexander Renkl
University of Freiburg

Journal of Educational Psychology
2003, Vol. 95, No. 4, 774–783

0022-0663/03/$12.00 DOI: 10.1037/0022-0663.95.4.774

Transitioning From Studying Examples to Solving Problems: Effects of Self-Explanation Prompts and Fading Worked-Out Steps

Robert K. Atkinson
Arizona State University

Alexander Renkl
University of Freiburg

Mary Margaret Merrill
Louisiana State University at Shreveport

Although research has demonstrated that successively fading or successively removing more and more worked-out solution steps as learners transition from relying on examples to independent problem solving reliably fosters performance on near-transfer tasks—relative to example–problem pairs—this effect is not reliable on far-transfer tasks. To address this, the authors combined fading with the introduction of prompts designed to encourage learners to identify the underlying principle illustrated in each worked-out solution step. Across 2 experiments, this combination produced medium to large effects on near and far transfer without requiring additional time on task. Thus, the instructional procedure is highly recommendable because it (a) is relatively straightforward to implement, (b) does not prolong learning time, and (c) fosters both near- and far-transfer performance.

Worked-out examples typically consist of a problem formulation, solution steps, and the final answer itself. Research indicates that exposure to worked-out examples is critical when learners are in the initial stages of learning a new cognitive skill in well-structured domains such as mathematics, physics, and computer programming (Anderson, Fincham, & Douglass, 1997). Moreover, studies performed by Sweller and his colleagues (e.g., Sweller & Cooper, 1985; for an overview, see Sweller, van Merriënboer, & Paas, 1998) document that learning from worked-out examples can be more effective than learning by problem solving.

Although worked-out examples have significant advantages, their use as a learning methodology does not, of course, guarantee effective learning. Chi and her colleagues (Chi, Bassok, Lewis, Reimann, & Glaser, 1989) noted that examples drawn from college-level physics textbooks often do not include all of the reasons why a certain step in the solution was performed. As a result, the burden of explaining the solution steps rests on the learner. Chi et al. (1989) discovered that learners attempted to establish a rationale for the solution steps by pausing to explain the examples to themselves and that these learners appeared to learn more than those who did not—a phenomenon they termed the *self-explanation effect*.

Self-Explanation Effect

At first, Chi and her colleagues (Chi et al., 1989) postulated that the self-explanation effect principally involved inference generation on the part of a learner. That is, by self-explaining, the learner is inferring information that is missing from a text passage or an example's solution. However, because of some inconsistencies among this view and some of the findings in the self-explanation literature, Chi (2000) revised this initial view by suggesting that the self-explanation effect is actually a dual process, one that involves generating inferences and repairing the learner's own mental model. In the latter process, it is assumed that the learner engages in the self-explanation process if he or she perceives a divergence between his or her own mental representation and the model conveyed by the text passage or example's solution. According to Chi, this new viewpoint extends the inference generation by suggesting that "each student may hold a naive model that may be unique in some ways, so that each student is really customizing his or her self-explanations to his or her own mental model" (p. 196).

According to Renkl (1997), there are four relatively distinct self-explanation styles, two associated with successful problem-solving strategies and two associated with inadequate strategies. Although he found that most learners were actually passive or superficial explainers who did not appear to learn much, Renkl (1997) discovered that the successful learners could be classified as either anticipative reasoners or principle-based explainers. He classified learners who tended to self-explain by anticipating the next step in an example solution, then checking to determine whether the predicted step corresponded to the actual step as

Robert K. Atkinson, Division of Psychology in Education, Arizona State University; Alexander Renkl, Department of Psychology, University of Freiburg, Freiburg, Germany; Mary Margaret Merrill, Department of Psychology, Louisiana State University at Shreveport.

We thank Paige Patterson for her assistance in running participants and coding data in Experiments 1 and 2. A subset of the results from Experiment 1 was reported at the American Educational Research Association, New Orleans, Louisiana, April 2002.

Correspondence concerning this article should be addressed to Robert K. Atkinson, Division of Psychology in Education, Arizona State University, P.O. Box 870611, Tempe, Arizona 85287-0611. E-mail: robert.atkinson@asu.edu

anticipative reasoners and noted that these learners started the learning process with a relatively high level of prior knowledge. Principle-based explainers, on the other hand, tended to identify the essential meaning of a problem by attempting to articulate its goal structure—including the application of operators—while also elaborating on the principle that the example was intended to convey. In contrast to the anticipative reasoners, however, these principle-based explainers had low prior knowledge. Thus, these findings suggest that it is functional to elicit principle-based explanations to learners with low prior knowledge while encouraging anticipative reasoning to learners with high prior knowledge.

Besides findings from correlational studies such as the ones by Chi et al. (1989) and by Renkl (1997), there is also experimental evidence that corroborates the significance of self-explanations when studying examples. In a study conducted by Renkl, Stark, Gruber, and Mandl (1998), learners in an experimental group were informed about the importance of self-explanations before being presented with a live model depicting how to self-explain. The control group, on the other hand, received think-aloud training instead of self-explanation training before the presentation of the instructional examples. According to their results, a short self-explanation training procedure produced an increase in the frequency of self-explanation activities among the learners in the experimental condition and enhanced both near-transfer performance (i.e., problems with the same structure or solution rationale but different surface features such as objects and numbers) and far-transfer performance (i.e., problems with different structure that required the generation of a modified solution procedure).

Similarly, Conati and VanLehn (1999, 2000) created and evaluated a computer-based learning environment designed to support learning from worked-out examples by prompting self-explanations. A tutorial component contained templates that were to be filled in by browser items (physics rules or subgoals in a solution plan) as building blocks of self-explanations. In addition, the tutorial component gave hints as to which aspects needed further self-explanations. Contrary to expectation, however, this environment did not foster learning gains—with the exception of some subgroups that were identified post hoc as profiting from the tool. Similarly, Hausmann and Chi (2002, Experiment 1) did not find positive effects of a computer-based facility where the students were encouraged to contribute their own written self-explanations by typing them into the computer. On the other hand, Aleven and Koedinger (2002) obtained positive results with prompting for self-explanations during problem solving rather than during example study in an intelligent instructional environment. Specifically, they documented that problem-solving practice within such a learning environment can be enhanced with prompting the learners to self-explain by identifying the underlying problem-solving principles. In summary, the findings concerning the use of self-explanation prompting in a computer-based learning environment are mixed and need further investigation.

Fading From Example Study to Problem Solving

It is also important how the instructional materials (examples and problems) are designed (for an overview, see Atkinson, Derry, Renkl, & Wortham, 2000). For instance, pairing examples with practice problems is more effective than exposing learners to either sets of examples only (Trafton & Reiser, 1993) or practice problems only (Sweller & Cooper, 1985). Recently, Renkl, Atkinson, and Maier (2000) proposed a variation of the traditional method of pairing examples with practice problems. They tested whether the two learning modes (i.e., example study and problem solving) could be combined by successively introducing more and more elements of problem solving in example study until learners are solving the problems on their own. In this way, more and more elements of anticipation are introduced when the knowledge level of the learner increases.

According to Renkl et al. (2000), this rationale is useful as a way to structure the transition from studying examples in initial skill acquisition to problem solving in later phases of the learning process. Specifically, they combined problem solving and example study in the following way. First, a complete example was presented (model). Second, an example was given in which one single solution step was omitted (scaffolded problem solving). Then, the number of blanks was increased step by step until just the problem formulation was left, that is, a problem to be solved (independent problem solving). In this way, a smooth transition from modelling (complete example) to scaffolded problem solving (incomplete example) to independent problem solving was implemented.

As a first step for testing this fading procedure, Renkl and his colleagues (Renkl, Atkinson, Maier, & Staley, 2002) explored the effectiveness of fading from example study to problem solving against the traditional method of using example–problem (EP) pairs within a computer-based environment. Across a field experiment and two controlled laboratory experiments, Renkl et al. (2002) found that (a) their fading procedure produced reliable effects on near-transfer items but not on far-transfer items, (b) the number of problem-solving errors generated during the learning phase played a role in mediating the effectiveness of the fading procedure, and (c) it was more advantageous to fade out worked-out solution steps using a backward approach by omitting the last solution steps first instead of omitting the initial solution steps first (i.e., a forward approach).

Because the fading procedure did not produce reliable effects on far transfer, this raises the question as to whether there are other instructional approaches that can be combined with fading from example study to problem solving that can foster far transfer in particular. For instance, although the fading procedure encouraged the learners to generate anticipations on those steps where the solutions were omitted, during the study of worked-out steps there was no instructional means to induce active processing. This might be a major drawback of this learning environment, especially because Renkl (1997) has shown that most learners are passive or superficial self-explainers. On the other hand, a learning environment that combines the procedure with prompts to encourage more active example processing during the study of worked-out steps—for instance, prompts that elicit principle-based self-explanations—might foster far transfer better than fading alone. By combining prompting at worked-out steps with fading (i.e., inducing anticipation), the learners would be required to learn in a fashion considered more favorable according to the findings of Renkl (1997), that is, to elicit principle-based self-explanations in the initial stages of learning and followed by a procedure that induces anticipations.

Overview of Experiments

On the basis of the aforementioned research, it appeared worthwhile to examine whether a learning environment that relies on fading could also incorporate explicit self-explanation prompts to encourage learners to think more deeply about the structure of the worked-out steps, thereby improving subsequent performance on far-transfer problems drawn from well-structured domains. The purpose of the present research was to examine the impact of two instructional approaches—backward fading (BF) and EP pairs—in combination with the presence or absence of self-explanation prompts on transfer across a set of probability tasks. Specifically, the aim of Experiment 1—a laboratory-based experiment—was to examine whether there was a positive learning effect associated with fading and/or self-explanation prompts—straightforward prompts designed to permit the learner to tailor his or her self-explanation according to his or her own mental model of the situation at hand—and whether there was an interaction between the use of fading and the use of self-explanation prompts. Although prior to the experiment we hypothesized that learners assigned to a BF condition would produce significantly more accurate near-transfer solutions than their counterparts assigned to the EP-pairs condition, the impact of self-explanation prompts and the interaction between prompts and type of instruction remained open questions. The goal of Experiment 2 was to replicate the novel findings of Experiment 1 within a more authentic, school-based setting. Across the two experiments, a common set of learning-process and learning-outcome measures were collected. The learning-process measures included accuracy of anticipations during learning and study time. In addition, Experiment 1 included the correctness of the learners' responses to the self-explanation prompts. The learning-outcome measures included correctness of solutions on near-transfer problems and far-transfer problems.

Experiment 1

This experiment was designed to address three research questions: (a) Does the BF produce more favorable learning outcomes than EP pairs? (b) Does the use of self-explanation prompts in comparison with the lack of such prompts lead to better learning outcomes? (c) Is there an interaction between the use of fading (vs. EP pairs) and the use of self-explanation prompts (vs. no prompts)?

Method

Participants and design. The participants of this study were 78 educational psychology and psychology students (27 freshmen, 27 sophomores, 13 juniors, and 11 seniors) at a large, southeastern university. The sample comprised 15 males and 63 females (mean GPA $= 3.07$, mean ACT score $= 21.99$). The participants were randomly assigned in approximately equal proportions to the cells of a 2×2 between-subjects factorial design. The first factor was the characteristics of the instructional material (BF or EP pairs). The second was the presence or absence of self-explanation prompts. Thus, this experiment consisted of four conditions: (a) BF only ($n = 19$), (b) EP pairs only ($n = 19$), (c) BF plus prompting ($n = 20$), and (d) EP pairs plus prompting ($n = 20$).

Pencil–paper materials. The pencil–paper materials included a demographic questionnaire, an overview of the fundamental principles of probability, a nine-item pretest, and a 13-item posttest. The questionnaire asked each learner to provide demographic information (e.g., standardized test scores, number of postsecondary statistics courses in progress or completed) that could be used to judge the learner's prior knowledge in statistics and mathematics in general. The five-page overview of the fundamental principles of probability covered such topics as (a) experiment and sample space, (b) probability of an event, (c) probability of the nonoccurrence of an event (i.e., the principle of complementarity), (d) probability of the linked occurrence of events (i.e., the multiplication principle for independent events), and (e) probability of A and/or B (i.e., the addition principle). The following is an excerpt from the overview's treatment of the topic of experiment and sample space:

> When doing an experiment, different events can occur; for example, a specific ball is pulled out of the ballot box or a specific number appears on a die. The sum of all possible events that can occur during an experiment is referred to as the *sample space*. For instance, if a six-sided die is rolled, the sample space includes the six numbers that can possibly appear on the die. On the other hand, if a coin is flipped, the sample space includes the two events that can occur, "heads" and "tails."

The pretest was designed to assess prior knowledge, and it consisted of nine relatively straightforward probability calculation problems (e.g., "When rolling a six-sided die what is the probability that 2 or 4 will appear?"). The posttest consisted of 13 problems including one very simple warm-up problem, which was ignored in our final analysis. Unlike the pretest, where each item only required the application of one probability principle per item, the 12 posttest items used in the final analysis each involved the coordinated application of several probability principles. Furthermore, the 12 posttest problems consisted of six near-transfer items and six far-transfer items. The near-transfer problems had exactly the same underlying structure (i.e., solutions involved the same set of probability principles applied in exactly the same manner) as several of the examples−problems the learners encountered during the learning phase but differed only in terms of surface characteristics (i.e., cover story, values for the problem parameters). Thus, despite sharing the same solution rationale, the near-transfer items and their structurally similar examples−problems from the learning phase appear different on the surface because they did not share either cover stories or problem values. The following is an example of a near-transfer item that is structurally isomorphic to several of the examples−problems provided during the learning phase:

> Charley needs an egg for cooking but is aware that some of the dozen eggs in his fridge are rather old and probably spoiled. Although he is not aware of it, four eggs are still edible while eight eggs are spoiled. He does know, however, that spoiled eggs unlike fresh eggs float in water. What is the probability that if Charley puts two eggs into the water that the first one floats but the second egg sinks?

On the other hand, far-transfer problems differed from the examples–problems provided during the learning phase with respect to both structure and surface features. That is, the far-transfer problems not only differed from the examples–problems from the learning phase in terms of their cover stories and values for the problem parameters, but also in terms of their solution rationale (i.e., solutions involved the same or similar sets of probability principles applied in different combinations). Thus, for the learner to correctly solve any of the far-transfer problems, he or she had to modify the problem-solving procedures illustrated in items from the learning phase to derive a solution to the novel transfer problems. The following is an example of a far-transfer item:

> 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?

Computer-based learning environment. Director 6.0 (Macromedia, 1997) software, an authoring tool for multimedia productions, was used to create the Windows-based learning environment used in this experiment. The learning environment was originally developed by Renkl (1997), modified by Stark (1999), and finally adapted to the present needs by Robert K. Atkinson. This computer-based learning environment was designed to deliver instruction to learners learning to solve probability word problems. It consisted of a set of worked-out examples and problems from the domain of probability calculation. On the whole, the instructional lesson consisted of two sets of probability tasks where each set consisted of four tasks with the same underlying structure (i.e., solution rationale) but different surface features (i.e., cover stories, values). Across all four tasks, the worked-out examples and problems consisted of exactly three solution steps or subgoals. To assist the learner in distinguishing a problem's subgoals, we visually isolated and labeled each of the three subgoals (e.g., *first solution step*). The following is the cover story from one of the worked examples provided during the learning phase:

> Mrs. Zinfandel purchased 12 bottles of her favorite vintage red wine. Unfortunately, due to improper storage, 4 bottles have turned to vinegar and are undrinkable. What is the probability that the first bottle that Mrs. Zinfandel opens is vinegar but the second one is drinkable?

The learning environment was configurable to run in one of four modes that reflected the four conditions of the present experiment. First, in the BF-only condition, the first task in each of the two sets was a completely worked-out example where all three of the problem's solution steps were sequentially provided to the learner. That is, instead of appearing on the screen as a completely worked-out example, this task appeared at the outset unsolved. The learner moved forward through the example by clicking on a *next* button and watched as each of the three solution steps was successively added—like a learner-paced animation—over a series of pages, with the final page containing the entire solution. Once the learner finished inspecting this worked example, he or she then proceeded to the next page by clicking on a button marked *next problem* located at the bottom of the page. The second task was similar to the first, with one notable exception: The final (i.e., third) solution step was omitted (see Table 1). Instead of presenting this step, the learner was required to anticipate this step on his or her own by typing in the solution into a field located on the screen (see Figure 1), otherwise he or she could not continue. The learning environment was programmed to record the correctness of the problem-solving attempts in a log file for subsequent analysis. After inputting the anticipated answer, the learner clicked the *next* button at which time the correct solution step was displayed for the learner to receive feedback on the correctness of his or her problem-solving attempt (see Figure 2). In the third task of each set, the first solution step was provided and the last two steps were omitted. As with the previous task, the learner was required to input a solution before continuing. Finally, in the fourth task of each set, all three steps were omitted. Thus, the final task of each set was essentially a problem-solving task.

In the EP-only condition, each set comprised two pairs of a completely worked-out example followed by a problem-solving task. In other words, all three solution steps for the first and third tasks of each set were provided, whereas all three solution steps for the second and fourth tasks of each step were omitted. Across the two sets of problems, there were a total of 12 omitted steps—the same number of unsolved solution steps as found in the BF condition—where the leaner was required to anticipate the answers.

The BF-plus-prompting and EP-plus-prompting conditions were indistinguishable from the BF-only and EP-only conditions, respectively, with one notable exception: the presence of self-explanation prompts (see Figure 3). In the two prompting conditions, the learner was encouraged to self-explain each solved solution step by first examining the step and then identifying which principle of probability the step exemplified. To encourage this process, we prompted the learner to select the probability principle—the same principles covered in the introductory material covered in the handout (i.e., probability of an event, the principle of complementarity, the multiplication principle for independent events, or the addition principle)—from a list that appeared to the right of the solved solution step and enter his or her selection before being permitted to continue. Once a principle was selected, the learner's response plus the correct principle appeared below the solved solution step for the learner to scrutinize and check his or her accuracy.

Table 1
Description of Instructional Material

Tasks	Solution step	Condition: Backward fading	Condition: Example–problem pairs
1 and 5	1	Worked[a]	Worked
	2	Worked	Worked
	3	Worked	Worked
2 and 6	1	Worked	Omitted[b]
	2	Worked	Omitted
	3	Omitted	Omitted
3 and 7	1	Worked	Worked
	2	Omitted	Worked
	3	Omitted	Worked
4 and 8	1	Omitted	Omitted
	2	Omitted	Omitted
	3	Omitted	Omitted

Note. Tasks 1–4 were in Set 1, and Tasks 5–8 were in Set 2.
[a] The solution step was provided for the learner. [b] The learner was responsible for solving that particular solution step.

Scoring. Five measures required scoring: correctness of principles, accuracy of anticipations, pretest, near transfer, and far transfer. The learning environment automatically coded the learner's correctness of principles and accuracy of anticipations. In the two prompting conditions, the number of principles that the learners were required to identify was held constant at 12 across both types of instructional material. For each principle that the learners correctly identified, 1 point was awarded (no partial credit), thereby producing a maximum score of 12 for this measure. To create a proportion correct on the measure, we summed the participant's response and divided by 12. With regard to accuracy of anticipations, across all four conditions, the number of unsolved solution steps or anticipations was held constant at 12. For each solution step that the learners correctly anticipated the answer, 1 point was awarded (no partial credit), thereby producing a maximum score of 12 for this measure. The participant's response on this measure was summed and divided by 12, thereby generating a proportion correct on the measure, with values ranging from 0 to 1.

On the pretest, each correct solution was awarded 1 point (no partial credit), thereby generating a maximum score of 9 for the pretest. On the posttest, each problem consisted of three distinct solution steps. For each correct step, 1 point was awarded (partial correct). Thus, if the participants solved each solution step correctly, the problem solution was awarded 3 points. For both the near- and far-transfer measures, 18 was the maximum score that a learner could achieve (e.g., 3 points per problem times 6 near-transfer problems). To create a proportion correct score, with values ranging from 0 to 1, we summed the participants' responses on each measure (i.e., near and far transfer) across all six questions and divided by 18.

Procedure. Small groups of participants, varying in size from 1 to 10, were brought into a laboratory equipped with 10 Windows-based desktop computers (600 mHz, 256 RAM, 15-in. monitors), each located in their own cubicle. The participants were seated at the individual desktops and instructed to work independently of their peers. During

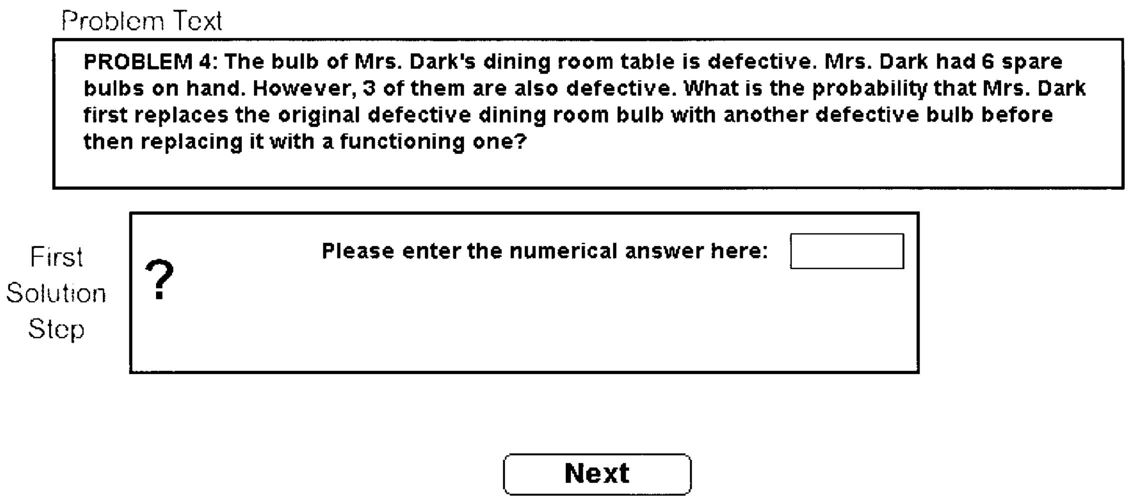

Figure 1. Example with a first missing solution step.

the course of the experiment, they spent approximately 90 min in the laboratory during which time they completed several tasks. First, the participants were asked to fill out a demographic questionnaire. Next, a pretest on prior knowledge in probability calculation was presented. To provide or reactivate basic knowledge that allowed the participants to understand the worked-out examples, we gave the participants an instructional text on basic principles of probability calculation. After reading this instructional text, the participants were to turn their attention to the computer-based learning environment and study the worked-out examples and solve the practice problems provided by the program. The participants were permitted to refer to the instructional text at any point during the computer-based portion of the experiment. During this phase, the experimental variation took place (BF vs. EP pairs, prompting vs. no prompting). The time spent for learning was recorded. Finally, the participants completed a posttest.

Results

Table 2 presents the mean scores and standard deviations for each group on each of the dependent measures. For the dependent measures, a 2×2 analysis of covariance (ANCOVA) was conducted using the pretest as a covariate ($\alpha = .05$; exception: instructional time; see below). Each measure was tested for homogeneity of regression, and the results were found to be nonsignificant—all Fs < 1.

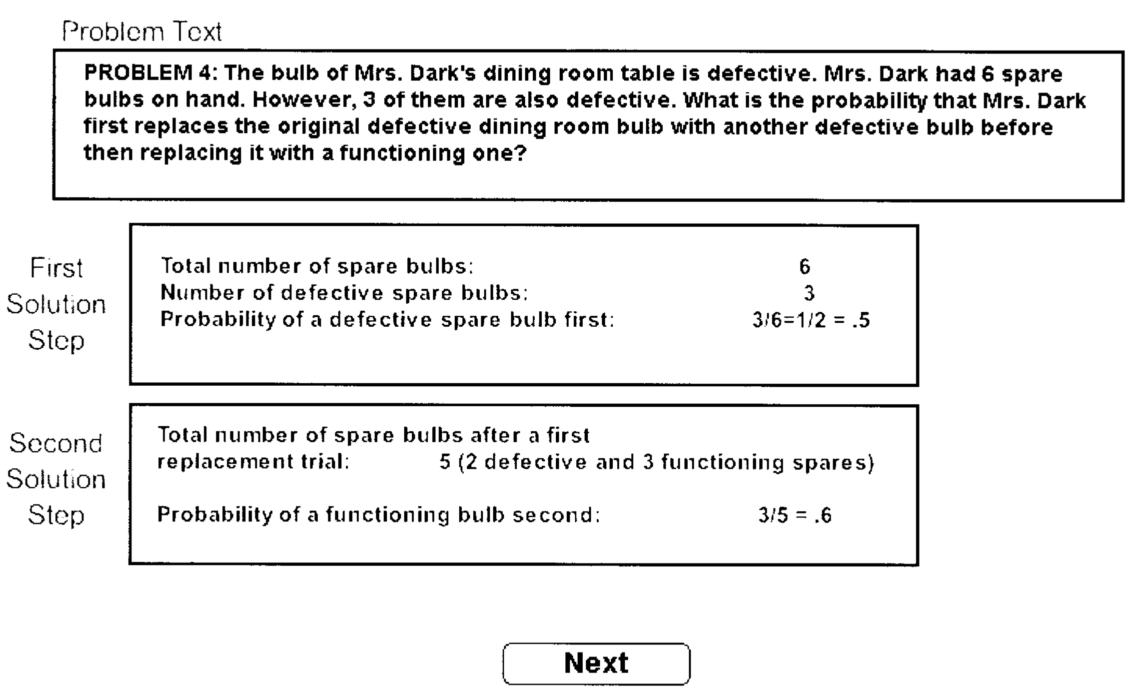

Figure 2. Example with a worked-out second solution step.

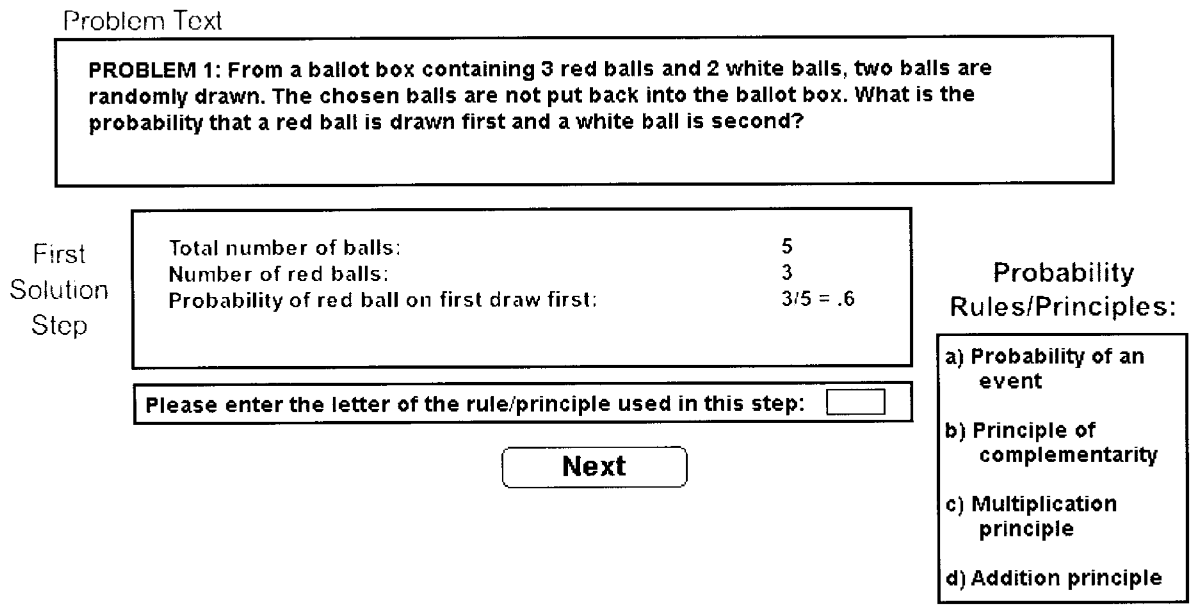

Figure 3. Example with a self-explanation prompt on the first solution step.

Analysis of learning-process measures. An ANCOVA revealed a significant main effect on anticipation for type of instruction, $F(1, 73) = 10.05$, $MSE = 0.07$, $p < .05$. The participants assigned to the BF conditions outperformed their peers in the EP-pairs conditions in terms of accuracy of anticipations. Cohen's f statistic for these data yields an effect size estimate of .33 for accuracy of anticipations, which corresponds to a medium to large effect. There was no significant main effect for prompting, $F(1, 73) = 0.66$, $p = .42$. There was also no interaction between type of instruction and self-explanation prompting, $F(1, 73) = 1.18$, $p = .28$.

To test for the possibility that the advantage of the prompting group could be attributed to time interacting with the instructional material, we conducted an analysis of variance (ANOVA) on instructional time. There was no significant main effect for prompting, $F(1, 74) = 0.01$, $MSE = 77.88$, $p = .95$, or fading, $F(1, 74) = 0.65$, $p = .42$. In addition, there was no significant interaction between these two factors, $F(1, 74) = 0.01$, $p = .94$. For correctness of principles, which only applied to the two prompting conditions (i.e., BF plus prompting and EP plus prompting), an ANCOVA yielded no significant main effect for accuracy of principles, $F(1, 36) = 0.19$, $MSE = 0.03$, $p = .66$.

Analysis of learning-outcome measures. There was a significant main effect for type of instruction material on near transfer, $F(1, 73) = 4.50$, $MSE = 0.05$, $p < .05$, where the participants who were assigned to the BF conditions significantly outperformed their counterparts in the EP conditions. Cohen's f statistic for these data yields an effect size estimate of .23 for near transfer, which corresponds to a medium effect. There was also a significant main effect for self-explanation prompting, $F(1, 73) = 5.01$, $p < .05$, where the participants who were presented with self-explanation prompts outperformed their peers who did not receive the prompts on near transfer. Cohen's f statistic for these data yields an effect size estimate of .25 for near transfer, which corresponds to a

Table 2
Study Time and Scores on Each Measure as a Function of Type of Instructional Material (Experiment 1)

	Example–problem pairs						Backward fading					
	No prompting ($N = 19$)			Prompting ($N = 20$)			No prompting ($N = 19$)			Prompting ($N = 20$)		
Measure	*M*	*SD*	Adj. *M*	*M*	*SD*	Adj. *M*	*M*	*SD*	Adj. *M*	*M*	*SD*	Adj. *M*
Pretest	4.21	2.02		5.35	1.81		4.95	1.76		5.58	2.32	
Correctness of principles				0.82	0.18	0.82				0.85	0.17	0.85
Accuracy of anticipations	0.42	0.29	0.48	0.52	0.30	0.49	0.73[a]	0.30	0.74	0.66[a]	0.33	0.62
Study time	33.05	2.03		31.30	1.97		33.05	1.97		31.58	2.03	
Near transfer	0.39	0.24	0.43	0.61	0.23	0.59[b]	0.58	0.25	0.58[a]	0.69	0.24	0.65[a,b]
Far transfer	0.36	0.18	0.40	0.52	0.20	0.50[b]	0.51	0.19	0.51[a]	0.60	0.18	0.57[a,b]

Note. Adj. = adjusted.
[a] Differs statistically from the example–problem pairs means. [b] Differs statistically from the no-prompting means.

medium effect. There was, however, no evidence of an interaction between the type of instructional material and the presence or absence of self-explanation prompts on this measure, $F(1, 73) = 0.81$, $p = .37$.

The same pattern of effects was evident on the far-transfer measure. As with near transfer, there was a significant main effect for type of instructional material, $F(1, 73) = 5.99$, $MSE = 0.03$, $p < .05$, where the participants in the BF conditions solved significantly more far-transfer problems than their peers assigned to the EP conditions. Cohen's f statistic for these data yields an effect size estimate of .27 for near transfer, which corresponds to a medium effect. Again, there was also a significant main effect for self-explanation prompting, $F(1, 73) = 4.50$, $p < .05$, where the participants who received self-explanation prompts produced significantly more accurate solutions to the far-transfer problems in comparison with their counterparts who did not receive the prompts. Cohen's f statistic for these data yields an effect size estimate of .23 for near transfer, which corresponds to a medium effect. There was, however, no evidence of an interaction between the type of instruction and the presence or absence of self-explanation prompts on this measure, $F(1, 73) = 0.31$, $p = .58$.

Discussion

Is there a positive learning effect associated with fading? The results of this experiment essentially replicate the findings of Renkl et al. (2002). That is, the BF condition was associated with a higher solution rate of near-transfer problems as Renkl et al. (2002) documented in Experiments 1, 2, and 3. Moreover, we provide additional support for the notion that the BF condition produced more accurate solutions on far-transfer problems, an effect that was inconsistent across the experiments in Renkl et al.'s (2002) study. Thus, it appears that the BF procedure can substantially foster both near and far transfer. In addition, the BF procedure resulted in a statistically significant effect on accuracy of anticipations. Finally, the advantage of fading could not be attributed to additional time on task.

Do self-explanation prompts impact learning? In contrast to the results of Conati and VanLehn (2000)—but in accord with Aleven and Koedinger's (2002) findings—a simple prompting procedure can substantially foster both near and far transfer. Hence, the acquisition not only of (relatively simple) rules (i.e., near transfer) but also of understanding (i.e., far transfer) can be fostered by this instructional procedure. It is also notable that the advantage of prompting could be achieved without significantly increasing learning time. This is a particularly important accomplishment in light of the fact that this prompting procedure—one that proved to be both effective and efficient—is a very simple and easy-to-implement feature for computer-based learning environments.

Is there an interaction between the use of fading and the use of self-explanation prompts? According to the results of this experiment, there was no evidence of an interaction between the use of fading and the use of self-explanation prompts on any of the measures. This may be regarded as a positive finding from an educational point of view because both instructional means produced at least medium effects on learning outcomes and were combined without causing any decrement in performance.

In sum, although there is little doubt after this experiment that near and far transfer is fostered by the BF procedure—an effect that has been consistently found across several studies—the effect of prompting, albeit positive, remains in contrast less substantiated. The results of Conati and VanLehn (2000) as well as of Hausmann and Chi (2002) in particular indicate that these results should be interpreted cautiously until they can be replicated. Hence, a sensible next step is a (conceptual) replication of this effect. Furthermore, it is important to investigate whether the findings hold not only for university students but also for school-age students.

Experiment 2

To address the open question that was mentioned in the preceding discussion, we conducted a second experiment. To replicate the findings of Experiment 1 with respect to prompting, identical conditions (BF plus prompting and BF only) were implemented. We did not include EP-pairs groups because (a) the advantages of fading has been well documented in previous studies, and (b) as shown in Experiment 1, fading and prompting do not interact with each other in terms of learning outcomes.

Besides the conceptual replication of the prompting effect, we tested whether our finding also held for school contexts. Specifically, the participants selected for this experiment were high school students. Against this background, this experiment was designed to address one primary research question: Do self-explanation prompts enhance the learning effect associated with fading?

Method

Participants and design. With parent permission, 40 students (18 males and 22 females) from a southern high school volunteered to participate in this study (mean GPA = 3.41). The participants consisted of 7 sophomores, 22 juniors, and 11 seniors, all of whom were currently enrolled in an advanced algebra course. The participants were randomly assigned in equal proportions (20 per condition) to one of two conditions: BF only or BF plus prompting.

Learning environment. The learning environment used in this experiment was similar to the one used in Experiment 1, with one notable exception: The EP-pairs conditions were no longer available. Instead, the learning environment was reconfigured to run in one of two modes that reflected the two conditions of the present experiment, namely, BF only and BF plus prompting.

Instruments. The instruments used in this experiment were the same as those used in Experiment 1.

Scoring. The scoring of the pretest, anticipative accuracy, near-transfer measure, and far-transfer measure was identical to that of Experiment 1.

Procedure. The procedure of this experiment was similar to that of Experiment 1, with one exception: The present experiment was conducted in a computer classroom equipped with 30 work stations located at the high school from which the students were recruited.

Results

Table 3 presents the means scores and standard deviations for each group on each of the dependent measures. For each dependent measure, an ANCOVA was conducted using the pretest as a covariate ($\alpha = .05$; exception: instructional time). Prior to analy-

Table 3
Study Time and Scores on Each Measure as a Function of Type of Instructional Material (Experiment 2)

	Backward fading					
	No prompting			Prompting		
Measure	*M*	*SD*	Adj. *M*	*M*	*SD*	Adj. *M*
Pretest	4.98	2.22		5.35	1.73	
Accuracy of anticipations	.62	.24	.63	.72	.21	.71
Study time	32.95	9.04		30.85	7.13	
Near transfer	.29	.27	.30[a]	.53	.32	.52[a]
Far transfer	.23	.24	.23[a]	.41	.25	.41[a]

Note. Means in a row sharing superscripts differ statistically. Adj. = adjusted.

sis, each measure was tested for homogeneity of regression, and the results were found to be nonsignificant—all Fs < 2.

Analysis of learning-process measures. There was no significant effect on anticipation, $F(1, 37) = 1.72$, $MSE = 0.4$, $p = .20$. To test for the possibility that the advantage of the prompting group could be attributed to time interacting with the instructional material, we conducted an ANOVA on study time. Although the learners assigned to the no-prompting condition spent slightly more time interacting with the material than their prompting peers, the difference was not statistically significant, $F(1, 38) = 0.56$, $MSE = 67.28$, $p = .46$.

Analysis of learning-outcome measures. ANCOVAs were conducted on near transfer and far transfer. There was a significant effect on near transfer, $F(1, 37) = 6.65$, $MSE = 0.07$, $p < .05$, where the participants who were presented with self-explanation prompts outperformed their peers who did not receive the prompts. Cohen's f statistic for these data yields an effect size estimate of .42 for near transfer, which corresponds to a large effect. The ANCOVA conducted on the measure of far transfer was also significant, $F(1, 37) = 5.14$, $MSE = 0.07$, $p < .05$. The participants who were presented with self-explanation prompts outperformed their peers who did not receive the prompts on far transfer. Cohen's f statistic for these data yields an effect size estimate of .37 for near transfer, which corresponds to a large effect.

Discussion

The results of this experiment clearly indicate that the use of self-explanation prompts in combination with a BF example sequence fosters learning. In particular, this combination appears to not only assist learners in solving problems similar to the ones provided during instruction, but, more important, problems that are structurally different from the instructional material. Moreover, this combination produces a large effect, which indicates that it is also of practical relevance. Finally, as in Experiment 1, the prompting effect had no "time cost"; in other words, it fosters the quality, not the quantity, of example processing.

General Discussion

This research provides additional support for the procedure put forth by Renkl et al. (2002) that entails the use of a fading procedure to structure the transition between studying examples in early stages of cognitive skill acquisition to solving practice problems in later stages. Moreover, this research demonstrates that using a BF procedure fosters the acquisition of rules that can be (more or less) directly applied (i.e., near transfer) as well as those that can be flexibly applied (i.e., far transfer).

The most important message of this article is, however, that a BF procedure can be combined with self-explanation prompting to produce an effect that is both statistically and practically significant. Furthermore, this combined procedure does not increase learning time beyond simply fading alone. Instead, this combination appears to positively influence the quality of example processing without increasing learning time, a learning outcome that we consider to be ideal.

Comparisons With Other Studies

One may also be surprised to learn that our very simple prompting procedure—one that required the learner to only select the underlying principle while not requiring elaborated reflections about the task—produced medium to strong effects on both near and far transfer. In the present study, not only were self-explanations prompted, but feedback about the correctness of the self-explanation was given. Hence, the present effect of our self-explanation procedure is probably due to eliciting self-explanations and to the feedback provided with these explanations. This is not a surprising finding given that the importance of feedback on self-explanations was emphasized recently by Aleven and Koedinger (2002). To understand the exact mechanisms that are impacted by prompting, it would be important to conduct follow-up experiments designed to separate the effects of finding the principle from the effect of providing feedback on the accuracy of principle selection.

Another open question emerges from Conati and VanLehn's (1999, 2000) research that found—in contrast to the present study—very restricted positive effects of a computer-based prompting procedure, one that required the student to select principles or subgoals that corresponded to the actual step. One possible explanation for these conflicting results is that Conati and VanLehn's (1999, 2000) presentation of instructional examples imposed high demands on working memory. Specifically, in their learning environment, the problem formulation and the solution, which each consisted of several boxes, could never be seen at once. To track the learning processes, Conati and VanLehn (1999, 2000) required that the learners move the mouse over one box at a time in order to reveal its contents. Thus, at any given point, only part of the solution was available to the learners. As a result, for the learners to understand the problem in its entirety once they reached the last solution step, they needed to maintain all of the preceding steps in working memory because the steps no longer appeared on the screen. In addition, when the learners wanted to self-explain, they first had to use a series of menus to construct their self-explanation. For example, they first had to decide whether to focus on domain principles or on subgoals before using a browser to search several submenus to complete their explanation. Taken together, the learning environment, including type of self-explanation prompting used by Conati and VanLehn (1999, 2000), may have imposed so much processing demands that many learners may have been cognitively overloaded. The presentation mode and the type of prompting in the present study is a relatively simple one that allows the learners to devote much of their cognitive capacity to gaining understanding. The divergence of the present results from the findings of Hausmann and Chi (2002) can be explained by the fact that they just generally encouraged the students to type comments to themselves. In the present study, in contrast, specific prompts were provided throughout the learning phase.

Theoretical Implications

Our findings on the usefulness of a learning environment that combines fading worked-out steps with self-explanation prompts support the basic tenets of one of the most predominant, contemporary instructional models, namely the cognitive apprenticeship approach (Collins, Brown, & Newman, 1989). This approach suggests that learners should work on problems with close scaffolding provided by a mentor or instructor. This approach is characteristic of Vygotsky's (1978) "zone of proximal development" in which problems or tasks are provided to learners that are slightly more challenging than they can handle on their own. Instead of solving the problems or tasks independently, the learners must rely—at least initially—on the assistance of their more capable peers and/or instructors to succeed. According to this approach, the learners will eventually make a smooth transition from relying on modeling to scaffolded problem solving to independent problem solving. In other words, this model advocates the fading of instructional scaffolding during this transition. Correspondingly, our partially worked-out examples provide a scaffold that permits learners to solve problems they could not successfully solve on their own. The instructional scaffolding—in the shape of worked-out solution steps—is gradually faded in our learning environment.

Reflection is also part of the cognitive apprenticeship process (Collins et al., 1989). That is, learners are encouraged to reflect on their problem-solving process and to try to identify ways of improving it. For instance, they are encouraged to reflect on the problems that they have missed and to try to explain how to generate a correct solution, a process that can increase the likelihood that the correct solution procedure will be internalized by the learner. As the present study suggests, one way of promoting this reflection process is to use prompts that induce self-explanations. In sum, our successful implementation of an arrangement that incorporates Vygotskian-based instructional principles (i.e., scaffolding, reflection) provides additional evidence of its value as basis for modern instructional models.

Practical Implications

One may ask whether it is practical to use the instructional procedures analyzed in this article for teaching skills in well-structured domains. Overall, the use of prompts that encourage the learners to figure out the principle that underlies a certain solution step can be recommended for several reasons, including the following: (a) it produces medium to high effects on transfer performance, (b) these effects are consistent across different age levels (university and high school), (c) it does not interfere with fading, (d) it is very easy to implement (even without the help of computer technology), and (e) it requires no additional instructional time. This prompting procedure is, however, not without its drawbacks. Because this procedure is designed to elicit principle-based explanations, it is ideally suited for well-structured domains such as mathematics and physics that contain clearly identifiable domain principles "under" each solution step—as was the case with our probability examples. As one can imagine, not all domains contain such clearly identifiable principles. Hence, it is worth noting that our prompting procedure can only be applied in an unmodified manner when each solution step can be explained by a principle within the domain. If this is not the case, the prompting procedure could be modified so that the explication of goal–operator combination is the focus, that is, at each worked-out step the learner has to explicate which subgoal is achieved. The research program of Catrambone (1996, 1998) has convincingly shown that elaborating on the goal structure of (well-structured) problems fosters the learner's transfer performance. However, whether this type of prompting in a less structured domain would produce comparable learning effects needs to be tested in future studies.

Open Questions for Further Research

Although the present work resulted in several significant educational insights, it also generated a number of new research questions. Two such questions have already been mentioned in the preceding discussion and refer to the following issues: (a) separation of the effects of finding the domain principles and of feedback on this activity and (b) effects of prompting other types of self-explanation, such as explication of goal–operator combinations.

Another interesting question refers to our decision to prompt only at worked-out steps. As the results of Aleven and Koedinger (2002) suggested, prompting self-explanation during problem solving can also foster learning. Against this background, it might be possible that the prompting effect would even be stronger when the learners are required to name the domain principle at each step, irrespective of whether it is worked out or to be solved. On the other hand, giving principle-based explanation at each step may become a redundant activity that contributes little to learning (cf. Pirolli & Recker, 1994).

Finally, another fruitful goal of future research would be to develop and evaluate versions of our fading and prompting procedures that can be used for studying instructional material containing worked-out examples coupled with practice problems from nonmathematized and less well-structured domains (for first steps in this direction, see Schworm & Renkl, 2002). This would enable us to develop and experimentally test instructional procedures that could be used across a wide range of educational tasks.

References

Aleven, V. A., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science, 26,* 147–179.

Anderson, J. R., Fincham, J. M., & Douglass, S. (1997). The role of examples and rules in the acquisition of a cognitive skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 932–945.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. W. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70,* 181–214.

Catrambone, R. (1996). Generalizing solution procedures learned from examples. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1020–1031.

Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General, 127,* 355–376.

Chi, M. T. H (2000). Self-explaining expository texts: The dual process of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science* (pp. 161–238). 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* (pp. 453-494). Hillsdale, NJ: Erlbaum.

Conati, C., & VanLehn, K. (1999). Teaching meta-cognitive skills: Implementation and evaluation of a tutoring system to guide self-explanation while learning from examples. In S. P. Lajoie & M. Vivet (Eds.), *Proceedings of AIED '99, 9th World Conference of Artificial Intelligence and Education* (pp. 297–304). Le Man, France: IOS Press.

Conati, C., & VanLehn, K. (2000). Toward computer-based support of meta-cognitive skills: A computational framework to coach self-explanation. *International Journal of Artificial Intelligence in Education, 11,* 398–415.

Hausmann, R. G. M., & Chi, M. T. H. (2002). Can a computer interface support self-explaining? *Cognitive Technology, 7*(1), 4–14.

Macromedia. (1997). *Director 6.0* [Computer software]. San Francisco: Author.

Pirolli, P., & Recker, M. (1994). Learning strategies and transfer in the domain of programming. *Cognition and Instruction, 12,* 235–275.

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

Renkl, A., Atkinson, R. K., & Maier, U. H. (2000). From studying examples to solving problems: Fading worked-out solution steps helps learning. In L. Gleitman & A. K. Joshi (Eds.), *Proceeding of the 22nd Annual Conference of the Cognitive Science Society* (pp. 393–398). Mahwah, NJ: Erlbaum.

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.

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.

Schworm, S., & Renkl, A. (2002). Learning by solved example problems: Instructional explanations reduce self-explanation activity. In W. D. Gray & C. D. Schunn (Eds.), *Proceeding of the 24th Annual Conference of the Cognitive Science Society* (pp. 816–821). Mahwah, NJ: Erlbaum.

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

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., van Merriënboer, J. J. G., & 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 15th Annual Conference of the Cognitive Science Society* (pp. 1017–1022). Hillsdale, NJ: Erlbaum.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Received August 14, 2002
Revision received January 14, 2003
Accepted May 1, 2003 ■