Towards a Generic Self-Explanation Training Intervention for Example-Based Learning

Carolin Busch, University of Freiburg, Engelbergerstr. 41, D-79085 Freiburg, caro.busch@web.de
Alexander Renkl, University of Freiburg, Engelbergerstr. 41, D-79085 Freiburg, renkl@psychologie.uni-freiburg.de
Silke Schworm, University of Regensburg, Universitätsstraße 31, D-93053 Regensburg, silke.schworm@paedagogik.uni-regensburg.de

Abstract: In order to fully exploit the potentials of example-based learning, the learners have to be prompted or trained to self-explain. A restriction of previous training interventions is that they employ the same materials in the training phase as in a subsequent learning phase. Against this background, we developed a computer-based generic self-explanation training that used the topic "fables". In an experiment with university students, we compared two conditions: Example-based learning to scientifically argue with or without a preparatory training intervention on self-explanation (*n* = 29 in each condition). Our generic training intervention fostered self-explanations in a subsequent learning phase on argumentation as well as argumentation skills without increasing learning time. Thus, we have taken a first promising step towards a generic self-explanation training intervention.

Introduction

Worked-out examples consist of a problem formulation and the final solution; typically single solution steps are additionally provided. There has been convincing evidence on the effectiveness of learning from worked-out examples for cognitive skill acquisition for over twenty years (e.g., Paas & van Merriënboer, 1994; Sweller & Cooper, 1985; Zhu & Simon, 1987). Much of this research was performed within the framework of Cognitive Load Theory (CLT) (see Pass, Renkl, & Sweller, 2003; Paas & van Gog, 2006; Sweller, van Merriënboer, & Paas, 1998). When CLT researchers discuss "learning from worked-out examples" (example-based learning), they typically mean that after the introduction of one or more domain principles (e.g., mathematical theorem, physics law), learners should be presented with several examples rather than a single example – as it is commonly the case. Despite this emphasis on examples, CLT researchers acknowledge the importance of requiring learners to solve problems later on in cognitive skill acquisition in order for them to reach proficiency in the domain they are studying (cf. the expertise-reversal effect by Kalyuga, Ayres, Chandler & Sweller, 2003). Hence, example-based learning is a sensible method for *initial* skill acquisition.

Worked-out examples are common place in the instructional material pertaining to well-structured domains such as mathematics or physics. In recent years, there has been a growing number of studies that have shown that example-based learning can also be employed successfully in more ill-structured domains (e.g., Gentner, Loewenstein, & Thompson, 2003: negotiation strategies; Rummel, Spada, & Hauser, 2006: cooperating in a productive way; Schworm, & Renkl, 2006: designing effective learning materials; Schworm & Renkl, 2007: scientific argumentation; van Gog, Paas & van Merriënboer, 2006: troubleshooting).

It is also important to note that example-based learning is only effective if the examples are designed in an appropriate way (see Atkinson, Derry, Renkl, & Wortham, 2000; Renkl, 2005) and if the learners explain the rationale underlying the examples' solutions to themselves (cf. self-explanation effect by Chi, Bassok, Lewis, Reimann, & Glaser, 1989). Hence, it is important to support the learners' self-explanation activities. This is especially true because most learners do not spontaneously self-explain or do so merely in a superficial or passive way (Renkl, 1997).

The relevance of self-explanations in the case of examples from ill-structured domains has been recently investigated, for instance, by Schworm and Renkl (2007). They analysed the effects of prompting self-explanations in an example-based learning environment on scientific argumentation in the sense of Kuhn (1991). The examples consisted of video clips showing exemplary dialogues that contained the crucial elements of Kuhn's argumentation model. Schworm and Renkl tested to what extent the learners were able to identify the argumentation elements and to acquire the skill to apply these elements in arguing in a new content domain. It was found that prompting the learners to identify the argumentation elements in the exemplary dialogues significantly fostered the acquisition of argumentation skills. Prompting was not, however, relevant for acquiring conceptual knowledge about argumentation.

In general, self-explaining examples can be fostered by prompts that are integrated in a learning environment, as just explained (see also Atkinson, Renkl & Merrill, 2003), or by training (e.g., Bielaczyc, Pirolli, & Brown, 1995; Renkl, Stark, Gruber & Mandl, 1998). Both approaches have advantages and disadvantages. Prompting fosters learning without an extra training intervention, but the learning environment has to include well-designed prompts (see Schworm & Renkl, 2007). Training can be employed if a given learning environment without prompts is to be studied later on. A restriction of the training concepts for self-explaining examples that have been tested so far is that they employed the same type of materials in the training phase as in a subsequent learning phase (e.g., Bielaczyc et al., 1995). For example, Renkl et al. (1998) trained half of their participants using examples of (compound) interest calculation in order to prepare them for a later learning phase dealing with the same domain. The self-explanation training of Wong, Lawson, and Keeves (2002) that is at least partially relevant to example-based learning focuses on geometry learning. Other self-explanation training interventions such as the Self-Explanation Reading Training of McNamara (2004) address other learning methods than example-based learning. Against this background, a self-explanation training intervention for example-based learning would be useful that could be used for any domain in a subsequent learning environment.

We developed such a generic short-term training intervention, and we tested its effects employing the example-based learning environment on scientific argumentation developed by Schworm and Renkl (2007). As already described, the examples consisted of video-taped dialogues. Schworm and Renkl have shown that self-explanation prompts that required the learners to analyse these dialogues in terms of the argumentation elements (i.e., principle-based self-explanations) fostered argumentation skills. In the present experiment, we employed this example-based environment without prompts, and we tested whether a generic training intervention fostered self-explanations and learning outcomes when later working in this environment without prompts. More specifically, we addressed the following research questions:

(1) Does a generic self-explanation training intervention foster the number of principle-based self-explanations in a subsequent learning phase?

(2) Does a generic self-explanation training intervention foster argumentation skills?

(3) Does a generic self-explanation training intervention foster conceptual knowledge about argumentation?

Method

Sample and Design

We randomly assigned 58 participants to two experimental conditions (n = 29 in each condition): Example-based learning to argue with or without a preparatory short-term training intervention on self-explanations. The participants were students from a university in South-West Germany (31 female, 26 male, one missing value; mean age: 24.7). As data from a questionnaire show, there was no difference between groups with respect to experience with computer-based learning materials, experience with computers in general, interest to learn about scientific argumentation, or interest to learn about gender differences (which is the topic of the video-taped dialogues). Furthermore, the vast majority of participants stated to have "no prior knowledge" (57 out of 58 participants) about the model of argumentation by Kuhn (1991), only one person stated to have "only marginal knowledge". The study lasted about two hours. The students received 15 Euros for their participation.

Training Intervention

In the experimental condition with training, the participants learned about self-explanation in a computer-based environment that consisted of the following main elements: (1) The learners read about the learning goals of this short-term training intervention. These goals are formulated as questions. The learners are told that they should be able to answer these questions at the end of the training intervention. (2) The learners receive general information about what self-explanations are, why they are important, and what important types of self-explanations can be differentiated. (3) This information is made more concrete by examples from the topic "fables". It is shown that a fable is characterized by several underlying features (e.g., animals as actors, principle of polarisation, "hidden" message) and that the readers have to engage in self-explaining a story in terms of above mentioned underlying features in order to identify the story as a fable (see Figure 1 for a screenshot from this section of the learning environment). (4) The learners practice to categorize several statements to certain types of self-explanations and receive feedback. (5) The learning goals - the questions to which the answer should be known now – are again presented. If the learners think that they cannot answer one or more of these questions, they receive one or more (short) "expert answers".

This training intervention lasted on average about 23 minutes. The control group without training studied a similar computer-based environment in terms of contents (i.e., fables) and structure (i.e., the instructional elements such as presenting learning goals). However the topic of self-explanation was not touched. The participants of the control group needed less time to work through their environment (about 17 min.).

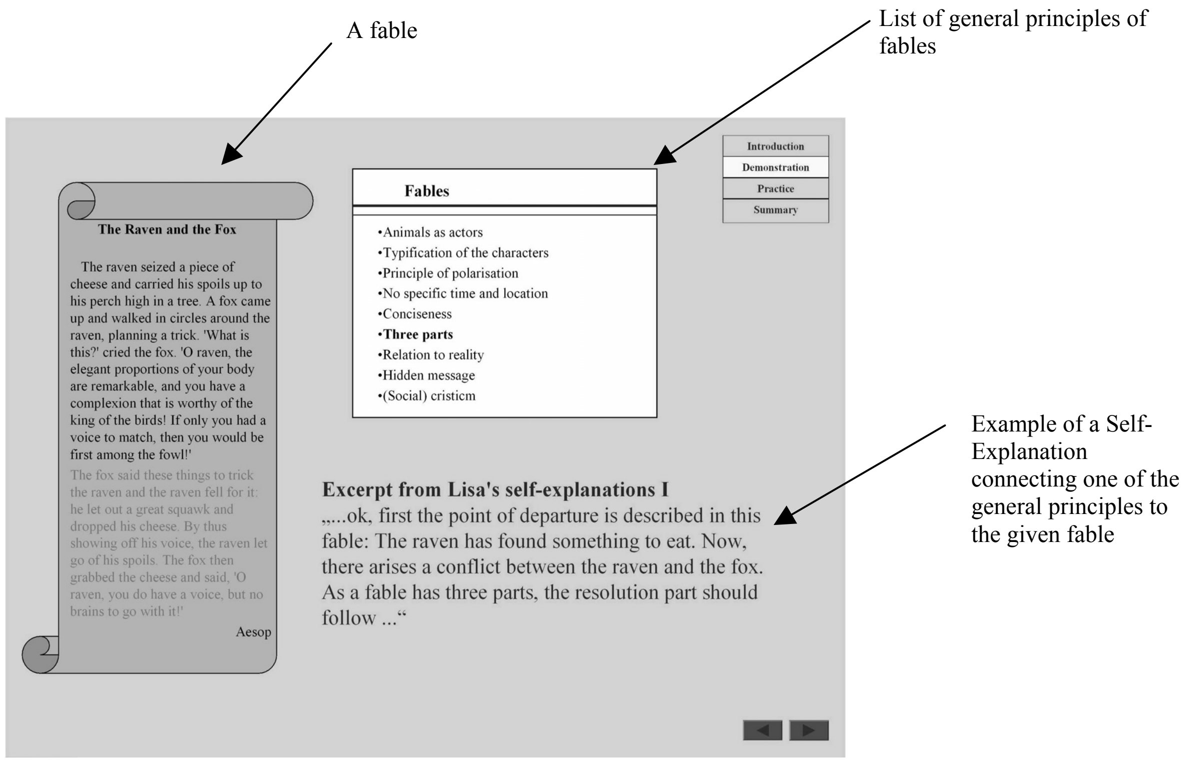

Figure 1. Screenshot from the training intervention (original screens are in colours and in German language).

Procedure

First, the participants filled in a questionnaire on demographic data (about 5 min.). Second, they were trained to self-explain using the topic of fables or just read about fables, depending on the experimental condition (about 20 min.). Third, all participants worked on a shortened version of the learning environment of Schworm and Renkl (2007; about 35 min.). This environment mainly consisted of an introduction to Kuhn's (1991) scientific argumentation model and a number of video examples that demonstrated the elements of this model. The dialogues were dealing with possible causes of gender differences in mathematics and science achievement in school. The participants first watched the whole dialogue (about 6 minutes), afterwards they watched the dialogue twice section per section. Each section of the video exemplified one or two components of the model of argumentation: (a) statement of theory and evidence, (b) statement of alternative theory, (c) counterargument against the original theory and its rebuttal, and (d) rebuttal of the alternative theory (Kuhn, 1991). While the participants watched each dialogue section twice the participants could write down self-explanations or other comments of any type in note boxes (these entries were logged for later analyses) (see Figure 2). However, there were no prompts to do so in any condition. Finally, the participants work on a test that assessed conceptual knowledge about argumentation and argumentation skills (in the sense of Kuhn, 1991; about 50 min.).

Instruments

Coding Scheme for Notes

The notes that the participants typed in the boxes included in the learning environment on argumentation were categorized according to the coding scheme shown in Table 1. We employed six different categories: principle-based self-explanations (relating concrete dialogue elements to the model of Kuhn, 1991), false self-explanations, descriptive statements (referring to the concrete dialogues contents, e.g., about gender differences in mathematics), false descriptive statements, metacognitive statements, and statements that were irrelevant in the present context. Part of the materials (about 15%) was randomly selected and coded by two trained raters. As a measure of reliability, the unadjusted interclass-correlation (ICC) was calculated. An ICC = .975 indicated very high interrater agreement. Thus, the remaining notes were rated by one coder who was blind to the experimental conditions.

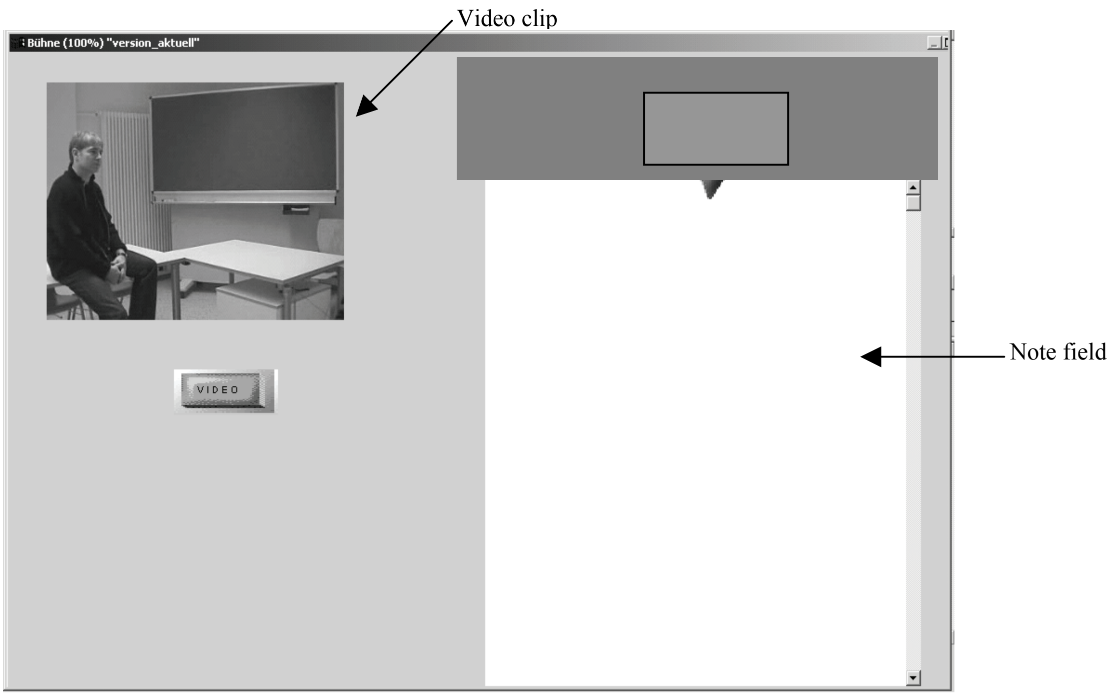

Figure 2. Screenshot from the learning environment on argumentation (original screens are in colours).

Table 1: Categorisation of the Notes that were Written down while Watching the Video Clips.

Type	Example
Principle-based self-explanation	"He puts up an alternative theory."
False self-explanation	"He puts up a counterargument." (actually, it was an alternative theory)
Descriptive statement	"She says, it can´t be the genes."
False descriptive statement	"She says, genes are responsible." (actually, the opposite was claimed)
Metacognitive statement	"I have forgotten the name of those elements."
Side aspects / irrelevant statements	"The guy seems to accept only his own opinion."

Test on Argumentation

The first part of this test assessed the participants' *conceptual knowledge about argumentation* asking two questions in an open format: 1. "What is good argumentation?" 2. "What are the elements of good argumentation?" These two questions had to be answered within 10 minutes. The answers of the questions were categorized according to the goals of the learning environment. A maximum of three points for the first question was assessed if the participants (a) mentioned the difference of theory and evidence, (b) stated the possibility of different perspectives, and (c) recognized the fallibility of their own opinion. A maximum of seven points for the second question had been assessed if the participant enumerated the elements theory, evidence, alternative theory, evidence of the alternative theory, rebuttal of the alternative theory, counterargument, and rebuttal of the counterargument. Summed up, a total of ten points could be reached in this subtest.

The subtest on *argumentation skills* contained six questions in an open format about a newly introduced topic: "What are the causes of increasing violence rates among adolescents?" (Note: German adolescents were meant in this context.) The questions were taken from Kuhn's interview questions (Kuhn, 1991), translated into German, and slightly adapted to the current context:

1. What do you think is the cause of increasing violence?
2. How would you prove that this is the cause?

3. What might somebody else, who does not agree with you, think is the cause of increasing violence?
4. What could you tell her/him to show s/he is wrong?
5. What might somebody else say to show that your opinion about the cause of increasing violence is wrong?
6. What could you tell her/him to show s/he is wrong?

The participants were asked to answer each question within five minutes. The answers were coded according to a coding-system based on the work of Kuhn (1991). Each answer was segmented into simple sentences, independent clauses, and reasonable catchwords. Each of them was coded separately according to the main components of the argumentation model. For a score of the quality of argumentation, the single ratings were aggregated. A totally correct solution was awarded with 9 points. This maximum number of points included (a) a theory and its supporting genuine evidence, (b) a correct alternative theory and its rebuttal, and (c) a successful counterargument and its rebuttal.

Part of the participants' answers to the argumentation-test items (about 15%) was randomly selected and coded by two trained raters. We found an ICC=.785 for the subtest on conceptual knowledge and an ICC=.827 for the subtest on argumentation skills. As the interrater agreement of coding proved to be good, the remaining protocols were coded by one of the coders who was blind to the experimental conditions.

Table 2: Means and Standard Deviations (in Brackets) of the Note Variables, Time-On-Task in the Learning Environment about Argumentation, and the Argumentation Test Scores

	Condition	
	Training	No Training
Principle-based self-explanations	4.28 (3.70)	1.86 (2.62)
Descriptive statements	18.86 (5.34)	21.72 (6.05)
False self-explanations	0.66 (0.81)	0.72 (1.33)
False descriptive statements	0.24 (0.64)	0.10 (0.31)
Metacognitive statements	0.03 (0.19)	0.03 (0.19)
Side aspects	0.59 (0.87)	0.90 (1.37)
Time-on-task on argumentation	34.48 (5.31)	34.19 (5.32)
Conceptual knowledge about argumentation	6.31 (1.83)	5.76 (1.66)
Argumentation skills	5.17 (2.12)	4.03 (1.99)

Results

We used an alpha level of .05 for all statistical tests. As an effect size measure, we used *d*, qualifying values of about 0.5 as medium effects and values of about 0.8 and greater as strong effects (see Cohen, 1988).

Table 2 provides the descriptive statistics of the notes variables, the time-on-task in the learning environment about argumentation, and the learning outcomes scores. With respect to the note variable, it can be seen that principle-based self-explanations and descriptive statements were typed in the note boxes to a substantial degree; all other categories occurred so infrequently so that they were ignored for further analyses (see Table 2). The participants in the training group typed more principle-based self-explanations into the note boxes as compared to the control group, $t(50.37) = 2.868$, $p = .006$ (*t* test for unequal variances). The effect size was strong ($d = .76$). Although, the number of "purely" descriptive statements was descriptively lower in the training group, this difference failed to reach the 5%-level of significance, $t(56) = -1.91$, $p = .061$.

The learners in both conditions spend on average about 34 min. in the learning environment on argumentation (see Table 2). Thus, there was no significant group difference in time-on-task, $t(56) = 0.415$; $p = .679$. With respect to learning outcomes, the training intervention led to better argumentation skills, $t(56) = 2.106$, $p = .040$. The effect size was medium ($d = .55$). We did not, however, find a significant training effect on conceptual knowledge about scientific argumentation, $t(56) = 1.200$; $p = .235$.

Discussion

The research questions can be answered as follows: (1) Our generic self-explanation training intervention fostered the number of principle-based self-explanations in a subsequent learning phase. (2) It also fostered argumentation skills. (3) It did not, however, foster conceptual knowledge about argumentation. The effect on argumentation skills was not due to time-on-task differences in the argumentation environment so that there were no "learning-time" costs of our training intervention. This finding suggests that the quality of the learning processes in the argumentation environment was heightened by our intervention. This claim is also supported by the higher number of principle-based self-explanations in the training group as compared to the control group.

The lack of a training intervention effect on conceptual knowledge is consistent with the findings of Schworm and Renkl (2007). They also found a self-explanation prompt effect only for skill acquisition but not for conceptual knowledge. Thus, the converging results of these two studies show that principle-based self-explanations referring to argumentation elements do not deepen conceptual knowledge; however, they are effective in enhancing skill acquisition. An issue for further research that cannot be answered on the basis of the present data refers to the question of whether training self-explanations and prompting self-explanations are about equally effective or whether one type of intervention is superior.

In any case, the present transfer effect is remarkable because the trained learners had to bridge the "transfer gap" from the topic of "fables" to "scientific argumentation". Even more, within scientific argumentation they also had to bridge the gap between the topic of the video examples presented for learning (i.e., gender difference in mathematics and science achievement in school) and the topic chosen for testing argumentation skills (i.e., increasing violence among (German) adolescents). Such transfer performance after such a short training phase (less than 25 minutes) is unusual (cf. Detterman, 1993: " First, most studies fail to find transfer. Second, those studies claiming transfer can only be said to have found transfer by the most generous of criteria and would not meet the classical definition of transfer" (p. 15)). The astonishing transfer in our study might be due to the fact that many effective instructional elements were included in the short-term training intervention such as presenting learning goals, informing the learners about self-explaining ("informed training"), and practice on self-explanation. In addition, it is important to note that the training intervention itself also followed an example-based learning rationale. Two exemplary fables and many examples of self-explanation utterances were presented. These factors might have contributed to the present transfer effect.

Given the transfer effect from fables to scientific argumentation, can we now claim to have developed a generic training intervention that can prepare students for example-based learning in any domain? Surely not! Beforehand, we have to show that our training intervention produces transfer effects to other domains as well. In addition, our learners were university students. Thus, another interesting question for future research is whether we can also find transfer effects with younger students (e.g., from high school).

In summary, we have taken a first promising step towards a generic self-explanation training intervention. However, further studies have to show how "generic" our training intervention actually is.

References

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.

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

Bielaczyc, K., Pirolli, P., & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction, 13*, 221-252.

Chi, M. T. H., 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, 18*, 145-182.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Detterman, D. L., (1993). The case for the prosecution: Transfer as epiphenomenon. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 1-24). Norwood, NJ: Ablex.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, 95, 393-408.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38,* 23-31.

Kuhn, D. (1991). *The skills of argument.* Cambridge: Cambridge University Press.

McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, *38*, 1-30.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38,* 1-4.

Paas, F., & van Gog, T. (Eds.) (2006). Recent worked examples research: Managing cognitive load to foster learning and transfer [special issue]. *Learning and Instruction, 16* (2).

Paas, F. G. W. C., & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem solving skills: A cognitive load approach. *Journal of Educational Psychology, 86*, 122-133.

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

Renkl, A. (2005). The worked-out-example principle in multimedia learning. In R. Mayer (Ed.), *Cambridge Handbook of Multimedia Learning* (pp. 229-246). Cambridge, UK: Cambridge University Press.

Renkl, A., 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.

Rummel, N., Spada, H., & Hauser, S. (2006). Learning to collaborate in a computer-mediated setting: Observing a model beats learning from being scripted. In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.) *Proceedings of the Seventh International Conference of the Learning Sciences* (pp. 634-640). Mahwah, NJ: Erlbaum.

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

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

Schworm, S., & Renkl, A. (2006). Computer-supported example-based learning: When instructional explanations reduce self-explanations. *Computers & Education, 46*, 426-445.

Schworm, S., & Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology, 99*, 285-296.

Van Gog, T., Paas, F., & van Merriënboer, J J.G. (2006). Effects of process-oriented worked examples on troubleshooting transfer performance. *Learning and Instruction, 16,* 154-164.

Wong, R. M. F., Lawson, M. J., & Keeves, J. (2002). The effects of self-explanation training on students' problem solving in high-school mathematics. *Learning and Instruction, 12*, 233-262.

Zhu, X., & Simon, H. A. (1987). Learning mathematics from examples and by doing. *Cognition & Instruction,* 4, 137-166.