Title
Designing Sets of Instructional Examples to Accomplish Different Goals of Instruction

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 24(24)

ISSN
1069-7977

Authors
Schorr, Tina
Gerjets, Peter
Scheiter, Katharina
et al.

Publication Date
2002

Peer reviewed

Designing Sets of Instructional Examples to Accomplish Different Goals of Instruction

Tina Schorr (t.schorr@iwm-kmrc.de)
Virtual Ph.D. Program: Knowledge Acquisition and Knowledge Exchange with New Media
Konrad-Adenauer-Str. 40, D-72072 Tuebingen (Germany)

Peter Gerjets (p.gerjets@iwm-kmrc.de)
Dept. of Applied Cognitive Science, Knowledge Media Research Center
Konrad-Adenauer-Str. 40, D-72072 Tuebingen (Germany)

Katharina Scheiter (k.scheiter@imw-kmrc.de)
Dept. of Applied Cognitive Psychology and Media Psychology, Institute of Psychology
Konrad-Adenauer-Str. 40, D-72072 Tuebingen (Germany)

Yiannis Laouris (laouris@cyber.cy.net)
Cyprus Neuroscience and Technology Institute
Promitheos 5, 1065 Nicosia (Cyprus)

Abstract

In this paper we discuss issues of instructional design with regard to different goals of instruction in the context of learning from examples. Two different approaches for identifying suitable instructional methods are considered: First, a cognitive task analysis is presented that examines problem-solving strategies applicable for solving mathematical word problems from a cognitive-modeling perspective. The second approach is based on a review of empirical findings on designing instructional examples. Together, these considerations lead to the selection of two instructional methods that are expected to foster learning with respect to different goals of instruction. This assumption is tested in two experimental studies presented in this paper.

Problem-Solving Strategies for Mathematical Word Problems: A Cognitive Task Analysis

When considering different options of designing instructional material it is important to take into account which instructional goal is to be accomplished. For example, in the case of learning from examples, this goal could consist in enabling a learner to solve problems that are either rather similar or dissimilar to the instructional examples. For solving such problems a learner needs to possess specific prerequisites (e.g., knowledge, time to perform different strategies). In order to enable the learner to acquire these prerequisites by studying instructional material, an instructional designer needs a precise conception of the prerequisites that are needed and therefore have to be imparted by the material. In order to acquire information on how problems in question can be solved and which prerequisites are therefore needed it is helpful for an instructional designer to perform a cognitive task analysis. This can be done by using cognitive modeling as a method of task analysis, i.e., by constructing computer models to simulate a human problem-solver's behavior based on cognitive theories (Gray & Altmann, 2001). Cognitive models are characterized by a high precision that is achieved by the necessity of stating explicit and formalized assumptions in order to get running computer models (Zachary, Ryder, & Hicinbothom, 1998). Because of its high precision a cognitive modeling approach allows for a detailed comparison of different problem-solving strategies (Rittle-Johnson & Koedinger, 2001) as well as for the derivation of instructional methods based on the theoretical assumptions which are specified in cognitive models (Pirolli, 1999).

In the cognitive task analysis presented in this paper we examine two different problem-solving strategies applicable for solving mathematical word problems. A strategy can be characterized as a conditional sequence of subgoals and operators that is suitable to achieve a particular goal (cf. Pirolli, 1999, p. 452). This corresponds to the way strategies are usually represented in cognitive models.

Generally, problem-solving strategies can be classified as search-based, example-based, or schema-based, respectively. Search-based strategies like means-end analysis are appropriate to solve puzzle problems like the Tower of Hanoi (cf. Newell & Simon, 1972) which do not presuppose domain-specific prior knowledge (knowledge-lean problems according to VanLehn, 1989). However, more complex tasks (like solving mathematical word problems) require example-based or schema-based strategies that operate on a rather elaborated knowledge base. Example-based strategies use concrete knowledge about example problems and their solutions. Within this group of example-based strategies different strategies can be distinguished that vary in the extent they make use of example information. Compared to example-based strategies, schema-based strategies use more abstract knowledge representable as generalized, automated problem-solving schemata (e.g. Sweller, van Merrienboër, & Paas, 1998). According to VanLehn (1989, p. 545) a schema consists of „information about the class of problems the schema applies to and information about their solutions". Example-based and schema-based strategies correspond to two main approaches in cognitive science, that is the similarity-based and the rule-based approach (Hahn & Chater, 1998).

In our task analysis we examined two strategies for solving mathematical word problems (i.e., the *keyword-strategy* and the *situation model-strategy*) by formalizing them as executable computer models within the framework of the ACT-R-architecture (Anderson & Lebiere, 1998). In the following paragraphs the two strategies will be outlined according to their subgoal structures.

- The *keyword-strategy* is an example-based strategy that uses concrete knowledge about examples when working on new problems (cf. Sowder, 1988). The strategy is characterized by *bottom-up* processing based on the mechanism of *principle-cueing* (Ross, 1987). The respective ACT-R-model starts with reading a given word problem in order to reach the top goal of solving it. While reading, a text phrase that contains certain keywords (or very similar expressions) can activate these keywords in memory. This activation process in turn may lead to the activation of those examples in memory that contain the respective keywords. These known examples will be retrieved if their activation is sufficiently high (*remindings* according to Ross, 1989) and will then be used to solve the current word problem by applying the same procedure to it that was used in order to solve the examples.
- The *situation model-strategy* is a schema-based strategy that operates on a more abstract and elaborated knowledge base and relies on *top-down* processing (cf. Reusser, 1990). The ACT-R model of this strategy again starts with reading, but at the same time it interprets a given mathematical word problem. On basis of this interpretation-process a situation model can be constructed which represents the situation described in the text in a compressed form (cf. Kintsch, 1998). This situation model is then interpreted in a mathematical fashion by matching it with domain-specific schemas representing different problems categories and their appropriate solutions. Thus, in this strategy a given word problem can be solved by applying the solution specified in a known schema that is selected and instantiated on basis of the situation model of the word problem.

A *comparative evaluation* of the two strategies by applying their cognitive models to solving word problems yielded the following results: The keyword-strategy is convenient for equivalent test problems. These are characterized by a near transfer distance because they belong to the same problem category as the instructional examples and are embedded within the same cover story (Reed, 1999). In order to solve such problems by using the keyword-strategy only a very limited amount of problem-solving time and a small knowledge base, mainly containing superficial keywords, are necessary prerequisites. For the application of the situation model-strategy on the other hand, more prerequisites are needed, in particular more time for problem-solving and a larger and more elaborated knowledge base. This base includes schemas for problem categories and knowledge on structural features of problems that determine the appropriate solution procedure. These higher demands with regard to time and knowledge, however, are accompanied by a good problem-solving performance in isomorphic test problems. These are characterized by an intermediate transfer distance, because they belong to the same problem category as the instructional examples, but are embedded within different cover stories (Reed, 1999).

To conclude, the cognitive task analysis demonstrates computationally that the goal of solving a mathematical word problem can be reached by applying different problem-solving strategies. Thereby, these strategies differ in their processing steps as well as in their prerequisites. In the cognitive task analysis, the latter were specified in terms of problem-solving time and knowledge. However, there are also differences between the strategies with regard to their appropriateness in the context of a specific instructional goal — like solving equivalentversus isomorphic word problems. For equivalent problems, the keyword-strategy is convenient whereas the situation model-strategy is more appropriate in the case of isomorphic problems. Therefore, if the goal of instruction consists in enabling learners to solve equivalent problems, the prerequisites for the keyword-strategy should be imparted by the instructional material. On the other hand, if learners are supposed to learn how to solve isomorphic problems, acquiring prerequisites for the situation model-strategy should be the focus for designing instructional materials. Hence, for an instructional designer it is important to consider the instructional goal in order to ensure that the respective prerequisites are imparted in the materials.

When possessing a precise concept of the prerequisites needed in the context of a specific instructional goal the next question that is to be answered by an instructional designer is how these prerequisites can be imparted by the instructional material. Instructional methods that are suitable in the context of the prerequisites of the two specified strategies may be identified by reviewing empirical findings on designing instructional examples.

Instructional Design: Learning from Examples

The advantages of using examples as learning material have been pointed out in many studies (for an overview cf. Atkinson, Derry, Renkl, & Wortham, 2000). It could be demonstrated that studying examples is of great help for knowledge acquisition and that especially multiple examples support schema induction (e.g., Quilici & Mayer, 1996). However, learners may also experience difficulties when learning from examples. In particular, Ross (1987) found evidence that learners face problems in discriminating between *structural features* of an example (which determine its solution procedure) and *superficial features* describing the example s cover story (which are irrelevant with regard to its solution). Some attempts have been made to counteract such difficulties by improving the design of instructional examples. In order to foster the acquisition of structural features by means of instructional design Atkinson et al. (2000) distinguish between modifying intra-example features, i.e., features concerning the format of a single example, and varying inter-example features, i.e., features related to combinations of multiple examples. An instructional method that bears on inter-examples features and that is supposed to direct learners attention to structural features is the utilization of so-called *structure-emphasizing example combinations* which are contrasted with *surface-emphasizing example combina-*

tions that guide learners attention towards superficial features (Quilici & Mayer, 1996). Both types of example combinations are based on imparting knowledge on multiple problem categories which are each illustrated by multiple examples. In the case of structure-emphasizing example combinations each example of a particular problem category is embedded within a different cover story whereas the same set of cover stories is used across problem categories (Table 1, left). Surface-emphasizing example combinations on the other hand, are characterized by the fact that all examples of a particular problem category are embedded within the same cover story which varies across different problem categories so that problem categories and cover stories are confounded (Table 1, right).

Table 1: Structure-emphasizing (A) and surface-emphasizing (B) example combinations

(A)	*Problem category (PC)*			
	PC1	PC2	PC3	PC4
Cover story (CS)	CS1	CS1	CS1	CS1
	CS2	CS2	CS2	CS2
	CS3	CS3	CS3	CS3
	CS4	CS4	CS4	CS4

(B)	*Problem category (PC)*			
	PC1	PC2	PC3	PC4
Cover story (CS)	CS1	CS2	CS3	CS4
	CS1	CS2	CS3	CS4
	CS1	CS2	CS3	CS4
	CS1	CS2	CS3	CS4

Quilici and Mayer (1996) asked their subjects to categorize new problems after they had studied examples that were either presented as structure-emphasizing or as surface-emphasizing example combinations. A clear superiority of structure-emphasizing example combinations as learning material could be demonstrated for this categorization task.

Structure-emphasizing and surface-emphasizing example combinations seem to be appropriate instructional methods for imparting the prerequisites necessary to apply the two problem-solving strategies discussed. Structure-emphasizing example combinations should help to acquire structural problem features that are required to apply the situation model-strategy which is appropriate for isomorphic problems. On the other hand, surface-emphasizing example combinations should foster the acquisition of (surface-based) keywords that can be used to apply the keyword-strategy that is suitable for equivalent problems. To test this idea we conducted a series of experiments that differ from the studies of Quilici and Mayer (1996) in several respects.

Using an example-based hypertext environment Quilici and Mayer (1996) conducted paper-pencil experiments without measuring the time that was needed to learn with the different example combinations. On basis of our cognitive task analysis we assume that the two sets of example combinations differ with regard to their time demands. The cognitive task analysis supposes an elaborated knowledge base for the situation model-strategy that rests upon structural features and schemata for problem categories. It is expected that the acquisition of such a knowledge base from structure-emphasizing example combinations demands complex cognitive processes. These processes may require more time investment than processes applied to surface-emphasizing example combinations that result in the acquisition of surface-based keywords. To test this hypothesis we implemented our experiments as computer-based experiments using a hypertext system for learning and problem-solving that allows for the concurrent automatic registration of time spent on each page visited by means of logfiles. As a result, average learning times for the two instructional methods can be determined and compared to each other.

Extending the application area to children The results of Quilici and Mayer (1996) are based on adult subjects. However, the examination of the proposed instructional methods is particularly interesting with children as subjects. The reason is that surface-emphasizing example combinations are a common learning material in mathematical school books regardless of the instructional goals. Therefore, instructional implications will immediately arise if structure-emphasizing example combinations proof superior in the context of a particular learning goal.

Using problem-solving tasks as test problems Whereas Quilici and Mayer (1996) tested the performance of their subjects mainly by administering categorizing tasks, we use problem-solving tasks that vary with regard to their transfer distance. The reason for this modification lies in the assumption that the superiority of an instructional method depends on the goal of instruction that is related to a particular transfer distance. Surface-emphasizing example combinations may enable the keyword-strategy that is appropriate for solving equivalent test problems. Structure-emphasizing example combinations may allow for the situation model-strategy that is superior in solving isomorphic problems. Because it is assumed that the prerequisites for the keyword-strategy are imparted by surface-emphasizing example combinations and the prerequisites for the situation model-strategy are facilitated by structure-emphasizing example combinations, the following results are expected: Studying surface-emphasizing example combinations has a positive impact on solving equivalent problems whereas learning with structure-emphasizing example combinations fosters the ability to solve isomorphic problems. But both instructional methods may be not very helpful if the instructional goal is to solve problems with a far transfer distance, i.e., novel problems that are characterized by a different cover story and by a different problem category compared to the instructional examples (Reed, 1999). This should be the case because both instructional methods do not impart flexible knowledge that is needed to solve novel problems.

Hypotheses

Based on the results of the cognitive task analysis as well as on the review of empirical findings on designing instructional examples we derived three experimental hypotheses about the use of structure-emphasizing and surface-emphasizing example combinations as instructional material. (1) *Time demands:* Learners using structure-emphasizing example combinations for learning need more time to study compared to learners using surface-emphasizing example combinations. (2) *Differential effectiveness:* Learners using structure-emphasizing example combinations perform better on isomorphic problems as learners using surface-emphasizing example combinations, whereas the latter show better problem-solving results when working on equivalent problems compared to learners using structure-emphasizing example combinations. (3) *Far transfer distance:* There is no performance difference between learners using structure-emphasizing and learners using surface-emphasizing example combinations when working on novel problems.

To investigate these hypotheses we conducted two experiments using the experimental environment BASICOPERATIONS that is described in the following section.

Experimental Environment

The hypertext environment BASICOPERATIONS used for experimentation is based on the hypertext-system HYPERCOMB developed by Gerjets, Scheiter, and Tack (2000). BASICOPERATIONS deals with the domain of basic arithmetic operations and is divided into a learning and a test phase.

In the learning phase, a learner is presented with 16 worked-out examples one after another in a fixed order. Four different problem categories are illustrated by four worked-out examples each. A problem category is formed by the conjunction of two different basic operations; illustrated are the problem categories (PC1) addition/multiplication, (PC2) subtraction/multiplication, (PC3) addition/division and (PC4) subtraction/division. However, another classification of the worked-out examples can be made with regard to their cover stories. Each of the 16 examples is embedded within one of four different cover stories whereby each cover story is used in four examples. The cover stories deal with (CS1) a family on a hiking trip, (CS2) a girl getting money, (CS3) a school arranging a sports meeting and (CS4) a boy buying food. The presentation of the instructional examples is blocked according to the problem categories in a predefined sequence, i.e., all four examples of one problem category are presented subsequently before the next problem category is illustrated.

Two different versions of BASICOPERATIONS (german version on the web: *karibik.cops.uni-saarland.de/knac/ex2/zypernA_deutsch* and *../zypernB_deutsch*) were used that can be classified as providing structure-emphasizing or surface-emphasizing example combinations according to the manipulation of Quilici and Mayer (1996). In the version with structure-emphasizing example combinations all four examples illustrating one particular problem category differ in their cover stories (cf. Table 1, left); in the version with surface-emphasizing example combinations all examples illustrating one particular problem category are embedded within the same cover story (cf. Table 1, right).

In the test phase, 18 word problems had to be solved one after another in a fixed order. The instructional example combinations were no longer available during testing. According to their transfer distance with regard to the instructional examples presented in the learning phase the test problems comprised equivalent, isomorphic, and novel problems. In order to calibrate the difficulty of the test problems we conducted a baseline study where subjects had to solve the test problems without any instructional information.

Baseline Study

Method

Participants Subjects were 49 third and forth grade pupils of an elementary school in Nikosia, Cyprus.

Materials and procedure Subjects received a booklet and were instructed to solve the 18 aforementioned word problems one after another as well as 20 simple arithmetic calculations that were used to measure basic arithmetic skills. Subjects received no guidance or instructional support.

Dependent measures One error was assigned for each wrong answer in the simple arithmetic calculations as well as in the word problems.

Results and Discussion

The simple arithmetic calculations yielded an average error rate of 30.7%. With regard to the word problems, the average error rate was 48.6%. This baseline performance indicated that the word problems were sufficiently difficult to solve for the children so that an instructional support might influence the results. Therefore, this set of word problems was used in the following experiments without any modifications.

Experiment 1

Method

Participants Subjects were 44 (mainly) third and forth grade pupils of different elementary schools in Nikosia, Cyprus, who participated in the study without payment. Average age was 8.3 years. These subjects had not participated in the baseline study.

Materials and procedure At the beginning of the experiment a pretest was administered to measure basic arithmetic skills. The pretest consisted of 10 simple arithmetic calculations and of 5 simple word problems. After that, a subject started working with BASICOPERATIONS on his or her own using a PC. The experiment ended after the subject had solved the final word problem in the test phase.

Design and dependent measures As a first between-subjects variable the example combinations presented as learning material were manipulated (structure-emphasizing vs. surface-emphasizing). As a second within-subjects variable the transfer distance of the test problems was manipulated by assigning equivalent, isomorphic, and novel problems to subjects. As a first dependent variable we measured the error rates in the pretest. For every wrong answer in the simple arithmetic calculations one error was assigned. With regard to the simple word problems a maximum of two errors for each problem was assigned: one error for applying the wrong basic operation and another error for wrong calculations. Pretest errors were analyzed in order to ensure that the two experimental conditions were comparable with regard to their prior arithmetic skills. Furthermore, time spent on processing example combinations in the learning phase was registered. In the test phase, error rates in the word problems were obtained by assigning one error for each wrong answer (yielding a maximum of 18 errors).

Results and Discussion

The average time spent with working in BASICOPERATIONS over the experimental groups was about 79 minutes, ranging from 44 to 112 minutes. The children had only little experience with solving word problems as reflected by the average error rates in the pretest for both experimental groups with 31.7% for learners in the structure-emphasizing and 38.5% for learners in the surface-emphasizing condition ($t(39) = -1.15$; $p > .20$; two-tailed).

Unexpectedly, there was no significant difference between learners of the two experimental conditions regarding the time spent on the example combinations with an average of 20.2 minutes for learners with structure-emphasizing and 22.7 minutes for learners with surface-emphasizing example combinations ($t(42) = -.85$; $p > .20$; one-tailed). Similarly, with regard to error rates for the word problems, a MANOVA (*example combinations* x *transfer distance*) revealed no main effects nor an interaction (all $Fs < 1$). A comparison of the average error rates for the word problems between subjects from experiment 1 (50.1%) and subjects from the baseline study (48.6%) showed that the instructional materials in experiment 1 had no effects on performance ($t(91) = .26$; $p > .70$; two-tailed). As an explanation for this unexpected finding one might assume that the children were too young to deal with the instructional materials and that those were not appropriate for learners with very low prior knowledge, i.e., for children just starting to learn how to apply basic operations and how to solve mathematical word problems. To examine this assumption we replicated the experiment with older children possessing a higher level of prior knowledge. For practical reasons, this experiment was conducted in Germany.

Experiment 2

Method

Participants Subjects were 51 third and (mainly) forth grade pupils of different elementary schools in the Saarland, Germany, who participated in the study without payment. Average age was 9.1 years.

Materials and procedure Materials and procedure were the same as in experiment 1. However, the materials were translated from Greek into German.

Design and dependent measures The design was exactly the same as in experiment 1 with example combinations (structure-emphasizing vs. surface-emphasizing) as a between-subjects variable and transfer distance (equivalent vs. isomorphic vs. novel word problems) as a within-subjects variable. As dependent variables pretest errors, learning time, and error rates for the test problems were measured.

Results and Discussion

The subjects in experiment 2 spent an average of about 52 minutes on working in BASICOPERATIONS, ranging from 21 to 81 minutes. With regard to the average error rates in the pretest, subjects in experiment 2 showed a significantly higher prior knowledge (10.1% errors) than those in experiment 1 with 35.0% errors ($t(90) = 8.02$; $p < .001$; two-tailed). For subjects in experiment 2, prior knowledge did not differ between the two groups with 9.6% for learners in the structure-emphasizing and 10.6% for learners in the surface-emphasizing condition ($t(49) = -.35$; $p > .70$; two-tailed).

First, we investigated whether the two experimental conditions of experiment 2 differed with regard to the learning time invested in studying the example combinations. As expected in the *time demands hypothesis*, learners in the structure-emphasizing condition spent significantly more time with the provided example combinations (17.5 minutes) than learners in the surface-emphasizing condition with 12.2 minutes ($t(49) = 1.81$; $p < .05$; one-tailed). This increased time demand can be seen as a result of the more complex processing necessary to learn with structure-emphasizing example combinations compared to less complex processes that are elicited by surface-emphasizing example combinations.

Second, we examined whether there was a differential effectiveness of the two sets of example combinations with regard to equivalent and isomorphic test problems (Figure 1).

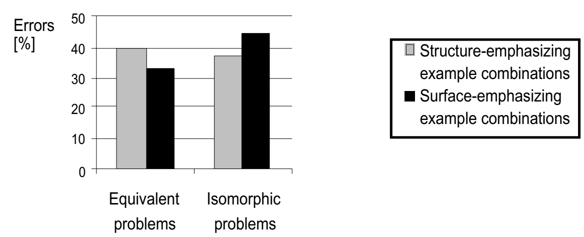

Figure 1: Mean error rates [%] as a function of example combinations and transfer distance

A MANOVA (*example combinations* x *transfer distance)* showed no main effect (all $Fs < 1$), but a significant interaction as expected ($F(1, 49) = 4.11$; $MS_e = 1144.91$; $p < .05$). Subjects who studied structure-emphasizing example combinations performed better on isomorphic test problems than subjects who studied surface-emphasizing example combinations whereas this pattern was reversed for the performance in equivalent problems. In this case the surface-emphasizing condition outperformed the structure-emphasizing condition. This finding provides evidence for the *differential effectiveness hypothesis* which assumes a difference of the two instructional methods with regard to their effectiveness depending on the instructional goal, i.e., transfer distance. Thus, to consider instructional goals may be essential when designing instructional materials.

Third, we tested the assumption that none of the two sets of example combinations supports solving novel test problems. As expected, the error rates for novel test problems did not differ between the two experimental groups with an average error rate of 36.5% for learners in the structure-emphasizing and 34.0% for learners in the surface-emphasizing condition ($t(49) = .30$; $p > .70$; two-tailed). This result provides evidence for the *far transfer distance hypothesis* which assumes that none of the instructional methods is superior in supporting to solve novel problems.

General Discussion

In this paper issues of instructional design were discussed. It could be shown that different set of instructional examples promote the acquisition of different problem-solving strategies and the accomplishment of different instructional goals.

It was argued that instructional goals must be considered when designing instructional materials. In order to specify these goals cognitive modeling as a method of cognitive task analysis was proposed. In the task analysis that was presented in this paper two specific strategies were examined and evaluated. As a result it was argued that problem-solving strategies differ in their prerequisites as well as in their appropriateness in the context of a particular instructional goal. Accordingly, instructional methods that impart the required prerequisites to apply specific problem-solving strategies also differ in their impact on problem-solving performance with regard to different instructional goals.

Based on these considerations two instructional methods, i.e., providing structure-emphasizing or surface-emphasizing example combinations, were examined in two experimental studies. It could be shown that both instructional methods were not appropriate for young learners with very low levels of domain-specific prior knowledge. But for learners possessing higher levels of prior knowledge, both, structure-emphasizing and surface-emphasizing example combinations proved to be appropriate methods to foster learning and problem-solving depending on the current instructional goal.

Acknowledgements

We thank Simon Albers, the Ministry for Educational, Cultural and Scientific Affairs of the Saarland, the Friedrich-Ebert-Stiftung as well as the Deutsche Forschungsgemeinschaft for their support and three unknown reviewers for their helpful comments.

References

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Hillsdale, NJ: Erlbaum.

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

Gerjets, P., Scheiter, K., & Tack, W. H. (2000). Resource-adaptive selection of strategies in learning from worked-out examples. In L. R. Gleitman, & A. K. Joshi (Eds.), *Proceedings of the 22nd Annual Conference of the Cognitive Science Society* (pp. 166-171). Mahwah, NJ: Erlbaum.

Gray, W. D., & Altmann, E. M. (2001). Cognitive modeling and human-computer interaction. In W. Karwowski (Ed.), *International encyclopedia of ergonomics and human factors* (Vol. 1). New York: Taylor & Francis.

Hahn, U., & Chater, N. (1998). Similarity and rules: Distinct? exhaustive? empirically distinguishable? *Cognition, 65,* 197-230.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, MA: Cambridge University Press.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Pirolli, P. (1999). Cognitive engineering models and cognitive architectures in human-computer interaction. In F. T. Durso, R. S. Nickerson, R. W. Schvaneveldt, S. T. Dumais, D. S. Lindsay, & M. T. H. Chi (Eds.), *Handbook of Applied Cognition*. Chicester: Wiley.

Quilici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology, 88*, 144-161.

Reed, S. K. (1999). *Word problems*. Mahwah, NJ: Erlbaum.

Reusser, K. (1990). From test to situation to equation: Cognitive simulation of understanding and solving mathematical word problems. In H. Mandl, E. De Corte, N. Bennett, & H. F. Friedrich (Eds.), *Learning and instruction in an international context* (Vol. II). New York: Pergamon Press.

Rittle-Johnson, B., & Koedinger, K. R. (2001). Using cognitive models to guide instructional design: The case of fraction division. In J. D. Moore, & K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 633-638). Mahwah, NJ: Erlbaum.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 629-639.

Ross, B. H. (1989). Remindings in learning and instruction. In S. Vosnaiadou, & A. Rotony (Eds.), *Similarity and analogical reasoning*. Cambridge, MA: Cambridge University Press.

Sowder, L. (1988). Children's solutions of story problems. *Journal of Mathematical Behavior, 7*, 227-238.

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

VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In M. I. Posner (Ed.), *Foundations of cognitive science*. Cambridge, MA: MIT Press.

Zachary, W. W., Ryder, J. M., & Hicinbothom, J. H. (1998). Cognitive task analysis and modeling of decision making in complex environments. In J. A. Cannon-Bowers, & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training*. Washington, DC: American Psychological Association.