

Sequence Effects in Solving Knowledge-Rich Problems: The Ambiguous Role of Surface Similarities

Katharina Scheiter (k.scheiter@iwm-kmrc.de)

Department of Applied Cognitive Psychology and Media Psychology, University of Tuebingen
Konrad-Adenauer-Strasse 40, 72072 Tuebingen, Germany

Peter Gerjets (p.gerjets@iwm-kmrc.de)

Multimedia and Hypermedia Research Unit, Knowledge Media Research Center
Konrad-Adenauer-Strasse 40, 72072 Tuebingen, Germany

Abstract

Sequence effects are said to occur whenever the problem-solving performance varies as a function of the order in which problems are solved. We present a framework that explain sequence effects as a result of (a) learning during solving a problem and of (b) transferring the learned content on succeeding problems. In two experiments we studied the ambiguous influence of surface similarities among structurally dissimilar knowledge-rich problems on sequence effects. These experiments demonstrate that surface similarities may either foster performance by enabling learning during problem solving or lead to negative transfer depending on whether problem solvers are already aware of the structural features of the problems prior to solving them or not.

Sequence Effects in Problem Solving

In this paper we investigate the ambiguous role of surface similarities for the occurrence of sequence effects when solving so-called knowledge-rich problems (VanLehn, 1996). Sequence effects occur whenever performance varies as a function of the order in which problems are solved. These effects can be analyzed as the result of (a) learning during problem solving and (b) transfer (Scheiter & Gerjets, 2002). *Learning* refers to a change in the cognitive system of the problem solver (i.e., newly generated or modified knowledge structures) that occurs due to solving a problem. *Transfer* refers to the transmission of these newly generated or modified knowledge structures to a subsequent target problem. Whenever problem orders are compared to each other that differ with regard to how they support these two processes, sequence effects may arise.

Sequence effects for knowledge-lean problems Sequence effects can be well demonstrated for rather context-free tasks whose accomplishment does not presuppose substantial domain-specific prior knowledge (i.e., knowledge-lean problems, cf. VanLehn, 1989). For instance, sequence effects can be found for simple stimulus-response tasks in the taskshift paradigm (Rogers & Monsell, 1995) or for rule induction tasks (Sweller, 1980). In these tasks learning mostly consists in the activation of knowledge structures necessary for the accomplishment of the task and transfer is seen as a more or less automatic process in which higher activated knowledge structures are more readily available when working on a succeeding task.

Sequence effects for knowledge-rich problems These problems differ from knowledge-lean problems in several respects. First, working on knowledge-rich problems requires domain-specific prior knowledge. Second, whereas knowledge-lean problems consist only of a small number of task stimuli that are relevant to their solution (i.e., structural features), knowledge-rich problems are often embedded in a semantically rich context (e.g., algebra word problems). That is, they may contain many surface features that are irrelevant to their solution. As will be stressed in the remainder of this paper, these surface features and their combination with structural features play a central, but ambiguous role for the formation of sequence effects when solving knowledge-rich problems. In particular, surface similarities among succeeding structurally dissimilar problems may either hinder performance by means of negative transfer or improve it by fostering structural awareness and learning during problem solving.

The differences between knowledge-lean and knowledge-rich problems may also influence the two processes that are made responsible for the occurrence of sequence effects. For knowledge-rich problems learning during problem solving may not consist only in the activation of already existing knowledge structures and transfer may not only be an automatic process:

- *Learning during problem solving:* Solving a knowledge-rich problem presupposes a substantial amount of problem comprehension, which is expressed in the problem solver's ability to construct a *situation model* of that problem (Nathan, Kintsch, & Young, 1992). A situation model is a qualitative representation of a problem in terms of concrete objects, events, and relations mentioned in the problem. The situation model can be transformed into a more abstract and quantitative *problem model* by further interpreting the situation model on the basis of problem schemas, which are "templates for organizing problem-relevant information" (Nathan et al., 1992, p. 332). These problem schemas are therefore considered to be the relevant prior knowledge helpful to solve knowledge-rich problems. The problem model that results from this knowledge-based interpretation process mainly comprises structural features that determine how a problem has to be solved. To sum up, learning during solving a knowledge-rich and context-rich problem - as the first factor in the explanation of sequence effects - results in a declarative representation of a problem and its solution that may be

more or less abstract depending on what prior knowledge a problem solver has at his or her disposal. However, there is evidence that even problem schemata themselves are not only abstract, but may include information that is tied to the concrete context of problems (Hinsley, Hayes, & Simon, 1977). In fact, many problem categories (defined by the structural features of the problems) are associated with a specific cover story so that superficial problem features may act as useful cues for retrieving a problem schema from memory. This is, however, helpful only for problems where a cover story activates a problem schema that *corresponds* to the structural features of the problem. Otherwise, the apparent resemblance with regard to superficial but not structural features may lead to performance impairments.

- *Transfer to subsequent problems:* Sequence effects are not only based on learning during problem solving but also rely on the fact that the learned content is applied to a new problem. In contrast to knowledge-lean problems, transfer among knowledge-rich problems is said to be a non-automatic process requiring effort and metacognitive awareness (Weisberg, DiCamillo, & Phillips, 1978; Salomon & Perkins, 1989). This kind of transfer is often referred to as transfer by analogy and consists in a number of different processes (Gentner, 1983; Gick & Holyoak, 1980). In order to solve a target, first a suitable source problem has to be retrieved from memory. Next, elements of the source problem and its solution are mapped onto the target. Finally, based on these mappings a solution for the target problem can be generated. Whether a target problem is successfully solved by means of analogical transfer depends on first whether a problem solver recognizes similarities between the target and an earlier problem-solving episode stored in memory. Second, this similarity judgment must be based on structural similarity, i.e., it must rely on features that are relevant to the solution of the problems. Unfortunately, a wide range of evidence shows that problem solvers often fail to notice spontaneously that they have encountered similar problems before (Gick & Holyoak, 1980). Furthermore, similarity judgments are most often based on superficial features of the problems, which sets the stage for negative transfer (Holyoak & Koh, 1987). Transfer in solving knowledge-rich problems – as the second factor in the formation of sequence effects – therefore crucially depends upon a problem solver’s ability to distinguish among a problem’s surface features and its deep structure.

Funnily enough, it may be a strong resemblance between problems with regard to superficial features that fosters this structural awareness. That is, comparing two problems that share many surface features but differ with regard to structural features may highlight the structural differences between these problems. This effect has been referred to as the “near-miss” effect (Winston, 1975).

To conclude, surface features of knowledge-rich problems and the way they are combined with structural features may either hinder performance (negative transfer) or improve it

by fostering structural awareness and learning during problem solving (near miss). This ambiguity is expressed in two contrasting hypotheses on sequence effects when solving knowledge-rich problems. These hypotheses can be best explained in the context of the experimental materials that will be introduced next.

Problem Sequences for Algebra Word Problems

The test problems that were used for experimentation consisted in nine algebra word problems. There were three problems that stemmed from each of three problem categories, namely work problems, motion problems, and interest problems (Mayer, 1981). Within each problem category the three problems could be solved by applying the same solution procedure. Different problem categories required different solution procedures. For the purpose of experimentally testing the impact of superficial and structural similarities on sequence effects we made use of a method introduced by Blessing and Ross (1996), i.e., we cross-varied the surface and structural features of the word problems (see Table 1).

Table 1: Algebra word problems used for experimentation

		Cover story		
		Work	Motion	Interest
Problem category	Work	Correspond.		
	Motion		Correspond.	
	Interest			Correspond.

Note: Grey-shadowed cells indicate non-corresponding test problems.

For each of the three problem categories there was always one problem that was embedded in the cover story that is typical for problems of that category (i.e., a so-called *corresponding problem*). Additionally, there were two problems of each problem category that were each couched into cover stories that are typical for another problem category (i.e., *non-corresponding problems*). For instance, the non-corresponding problem shown in Table 2 belongs to the problem category of work problems, however, its cover story is very similar to that typical of motion problems.

Table 2: Corresponding and non-corresponding test problems for the work-problem category

Corresponding problem: An electrician needs 3 hours in order to repair a lamp, while his apprentice needs 6 hours to perform the same job. How long would it take them to repair the lamp, if they worked together?

Non-corresponding problem: Jane needs 6 hours to drive to Tom by car, whereas Tom needs 4 hours for driving the reversed way. How long would they need to drive until they meet on the way, if they started at the same time?

Common solution procedure: $1 = (\text{Rate}_1 + \text{Rate}_2) * h$

In order to investigate the ambiguous role of surface features two different problem sequences were constructed for the nine corresponding and non-corresponding test problems. In the *structure-blocked sequence* the test problems were presented according to their category membership resulting in three blocks of problems (work-, motion-, interest-category), so that within each block there were three problems that could all be solved by applying the same solution procedure. On the contrary, in the *surface-blocked sequence* test problems were presented according to their cover stories resulting in three blocks of superficially similar problems (work-, motion-, interest-story). However, within each block each of the three problems had to be solved by a different solution procedure. In both sequences each block of problems began with a corresponding problem (i.e., a problem couched into its typical cover story), followed by two non-corresponding ones (i.e., problems couched into atypical cover stories).

In the remainder of the paper these problems for which we varied the problem-solving sequence will be referred to as isomorphic problems, because they are structurally identical (isomorphic) to examples participants had to study before working on the test problems. In addition to these isomorphic problems participants had to solve three transfer problems that each required an adaptation of one of the known solution procedures. For the transfer problems order was kept constant for all participants.

We investigated two contrasting hypotheses concerning performance differences between the two problem sequences that are illustrated in the next section.

Transfer versus Near Miss?

The transfer hypothesis states that performance for isomorphic problems in the structure-blocked sequence should be superior to performance in the surface-blocked sequence, whereas the near-miss hypothesis assumes the opposite.

Transfer hypothesis The first hypothesis is based on an activation-summation model that is commonly used to describe the retrieval of source problems in analogical transfer (Holyoak & Koh, 1987). According to this model the retrieval of stored problem knowledge to solve a target problem is guided by structural as well as superficial similarities between target and source. That is, any feature of a target problem can act as a source of activation for mental representations that also contain this feature. “Activation from multiple shared features will summate, and if the activation level of a stored representation exceeds some threshold, that representation will become available for further processing” (Holyoak & Koh, 1987, p. 333f.). The most activated representation is used for solving the current problem. Using this model in order to predict performance for the aforementioned problem sequences leads to the assumption that negative transfer among succeeding problems should occur in the case of a surface-blocked problem sequence as is illustrated in Figure 1.

The *surface-blocked sequence* starts with a corresponding problem (indicated by identical indices for problem category PC and cover story CV) that may activate the correct

problem schema. The second problem $PC2_{CV1}$ is a non-corresponding problem from a different problem category than the first problem, but embedded in a cover story typical for problems of problem category 1. Its structural features may activate the correct problem schema 2; however, its surface features activate problem schema 1, which is already highly activated as it has been applied to solve the first problem. Therefore, problem schema 1 may be spuriously used to solve the second test problem – resulting in negative transfer. A correct problem schema is applied again when the next corresponding problem ($PC2_{CV2}$) is encountered at the beginning of the second block.

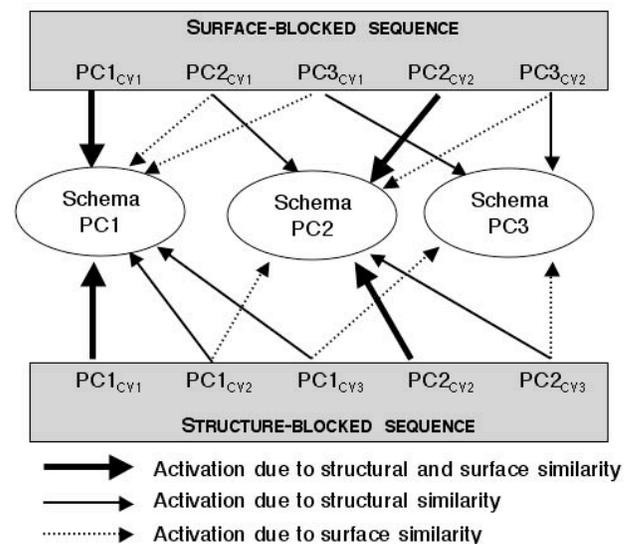


Figure 1: Negative and positive transfer based on activation summation

On the contrary, in the *structure-blocked sequence* the structural features of the second problem $PC1_{CV2}$ cue the already highly activated problem schema 1 that has been used to solve the first (corresponding) problem and that is also suited to solve the second test problem. Additionally, there may be activation of the wrong problem schema 2 due to shared surface features, however, this should not be sufficient to dominate the selection process. In this problem sequence non-corresponding problems within each block benefit from the previous activation of a problem schema by means of positive transfer, whereas performance for non-corresponding problems is impaired in the case of a surface-blocked sequence (i.e., negative transfer). According to the transfer hypothesis sequence effects in favor of a structure-blocked sequence should in particular be observable for non-corresponding compared to corresponding problems. These sequence effects can be explained by differences among the sequences with regard to *enabling positive transfer* among succeeding problems.

Near-miss hypothesis According to the second hypothesis a surface-blocked sequence should result in better performance compared to a structure-blocked sequence. This should be the case because a surface-blocked sequence highlights structural differences among succeeding

problems as these problems share the same cover story and differ only with regard to a few, but crucial features, i.e. structural features (Winston, 1975). On the contrary, in the structure-blocked sequence in which superficially dissimilar problems of one category are presented in succession attention may be directed toward differences with regard to the surface features between these problems and may thereby even impede problem-solving performance. Again, sequence effects should be in particular observable for non-corresponding problems. These sequence effects can be explained by differences among the sequences with regard to enabling *learning during problem solving* (i.e., acquiring knowledge on structural features).

Therefore, the near-miss hypothesis also predicts better performance for transfer problems, if they are solved subsequently to a surface-blocked sequence of isomorphic test problems. Transfer problems require an adaptation of known solution procedures and thus presuppose knowledge on structural features and on their relation to solution steps. The transfer hypothesis, on the other hand, does not allow for any predictions concerning transfer performance.

The two contradicting hypotheses were tested in two experiments described in the following sections.

Experiment 1

Method

Participants Subjects were 40 students (33 female, 7 male) of the Saarland University, Germany, who participated for course credit or payment. Average age was 23.03 years.

Materials and procedure Participants first received a booklet that contained an example problem together with a step-by-step solution for each of the three problem categories used for experimentation. The worked-out examples were couched into their typical cover stories. Participants were given twelve minutes to study the examples carefully. Subsequently, the example booklet was taken away from them and they received another booklet with the nine test problems that were isomorphic to the previously studied examples. Each test problem was presented on a single page and the pages were arranged according to either the surface-blocked or structure-blocked sequence. Participants were instructed to solve the problems in the given order. There were no time limits during testing. When participants had finished working on the nine isomorphs they had to solve the three transfer problems.

Design and dependent measures The problem sequence for the nine isomorphic test problems was varied as a between-subjects variable. The problems were either blocked according to the problem category they belonged to (structure-blocked sequence) or according to their cover stories (surface-blocked sequence). The sequence of the three blocks (either containing structurally or superficially similar problems) was counterbalanced within the two experimental conditions. As performance measures

subjects' error rates for the isomorphic and transfer problems were registered.

Results and Discussion

In all analyses of variances reported here we used math grades and gender as covariates because they were both strongly associated with problem-solving performance. In a first step we analyzed subjects' performance on the isomorphic test problems by an ANCOVA (Figure 2).

As postulated in the near-miss hypothesis subjects who were presented with the surface-blocked sequence outperformed subjects who worked in the structure-blocked condition ($F(1,35) = 4.51$; $MS_E = 305.01$; $p < .05$). Further distinguishing between corresponding and non-corresponding isomorphic problems revealed that this sequence effect in favor of the surface-blocked sequence could only be observed for non-corresponding problems ($F(1,35) = 5.27$; $MS_E = 446.94$; $p < .05$), whereas there were no performance differences for corresponding problems as predicted ($F < 1$).

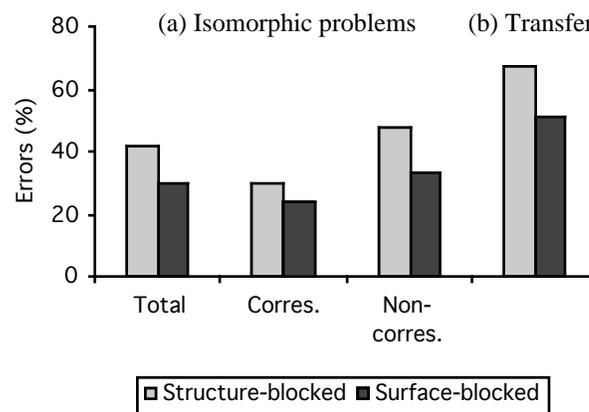


Figure 2: Problem-solving errors for (a) isomorphic problems (total, corresponding and non-corresponding) and (b) transfer problems as a function of problem sequence

Arranging isomorphic problems according to their cover story compared to arranging them according to their structural features not only improved performance on these problems, but also positively influenced subsequent transfer performance ($F(1,34) = 5.17$; $MS_E = 572.02$; $p < .05$). There were no trade-offs between the time needed for problem solving and any of the error measures.

To conclude, performance on the isomorphic as well as on the transfer problems clearly supported the near-miss hypothesis. That is, presenting superficially similar problems successively emphasizes differences between these problems with regard to structural features and thereby fosters problem-solving performance. Additionally, working on problems from different problem categories alternately seems to result in more knowledge on structural features that can be brought into play in order to solve problems that require an adaptation of known solution procedures.

It is important to note that the surface-blocked sequence is superior because it supports a process that seems to be the most crucial prerequisite for transfer, namely recognizing

structural features among problems. This explanation has an interesting consequence as it implies that the observed superiority of the surface-blocked sequence should vanish under certain conditions. In particular, there should be no sequence effect due to near miss when problem solvers are already sufficiently aware of the structural properties of the to-be-solved problems.

In order to address this issue we conducted a second experiment where subjects' attention was directed towards structural features of the to-be-solved problems by means of a categorization task *before* they attempted to solve them either in the structure-blocked or surface-blocked condition.

Experiment 2

Experiment 2 was designed to test the assumption that fostering structural awareness by means of a prior categorization task should reduce the superiority of the surface-blocked sequence. More precisely stated, we expected that participants who were able to categorize the problems according to their structural features would no longer benefit from a surface-blocked sequence in which structural differences are highlighted. On the contrary, it might even well be that due to the fact that good categorizers already know about the problems' structural similarities they might benefit from activation mechanisms in the structure-blocked sequence in which succeeding problems can be solved by the same solution procedure, thereby enabling positive transfer among these problems. However, for subjects who experience difficulties in detecting the structural similarities among problems in the categorization task, the surface-blocked sequence might be another chance to recognize the structural properties of the problems. This superiority of the surface-blocked sequence may however be less accentuated compared to Experiment 1 as it might at least be partially overridden by minor effects of prior problem categorization. In sum, poor categorizers were expected to show a – maybe less accentuated – superiority of the surface-blocked sequence, whereas no sequence effect (or even a superiority of the structure-blocked sequence) was postulated for good categorizers.

Method

Participants Subjects were 40 students (29 female, 11 male) of the Saarland University, Germany who participated for course credit or payment. Average age was 25.03 years.

Materials and procedure The same materials were used for experimentation as in Experiment 1. Participants first studied the booklet containing three worked-out examples for a maximum of twelve minutes. In contrast to Experiment 1 they did not start working on the isomorphic test problems immediately after this learning phase. Rather, participants received a sheet of paper that listed all nine isomorphic test problems, whereby the order of presentation depended on experimental condition. They were instructed to categorize problems according to their mathematical features by assigning the letter 'A' to all problems that could be solved in a similar way, the letter 'B' to all problems that could be

solved in a different way etc. It was stressed that there could be as many categories as they liked and that the number of problems assigned to each category did not need to be identical for all categories. After having accomplished the categorization task participants were told to solve the problems they had categorized before. The list with the problems to be solved and the category assignments produced by the participant were visible during the whole test phase. Contrary to Experiment 1, the test problems were not presented each on a single page, rather participants received only the list and blank paper to write down their solutions. After having solved the isomorphic problems, participants received the three transfer problems.

Design and dependent measures Problem sequence for the nine isomorphic test problems was varied as a between-subjects variable. The problems on the list for the categorization task were either presented in the structure-blocked sequence or in the surface-blocked sequence. The sequence of the three blocks (either containing structurally or superficially similar problems) was counterbalanced within the two experimental conditions. Additionally, we distinguished between good and poor categorizers within each of the experimental conditions (see below).

As performance measures participants' error rates for the isomorphic and transfer problems were registered. Additionally, we determined the performance in the categorization task by computing a structure score according to Quilici and Mayer (1996) that indicates participants' ability to categorize problems according to their structural similarities. This structure score was used for a median split within both problem sequence conditions in order to distinguish between good and poor categorizers.

Results and Discussion

In a first step we analyzed subjects' performance on the isomorphic test problems (Figure 3) by means of an ANCOVA (problem sequence x quality of categorization; math grades and gender as covariates).

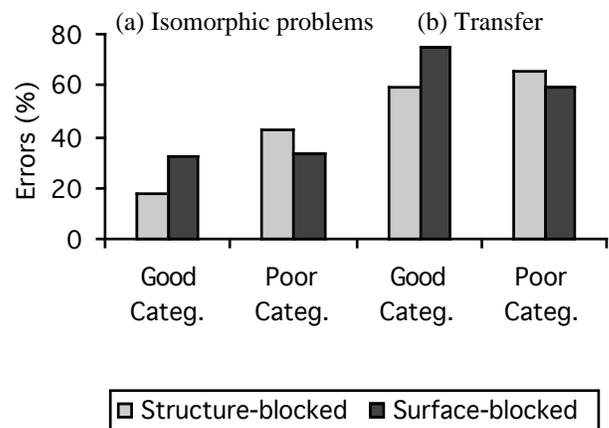


Figure 3: Problem-solving errors for (a) isomorphic and for (b) transfer problems as a function of problem sequence and quality of categorization

Neither problem sequence nor quality of categorization had an impact on problem-solving performance (problem sequence: $F < 1$; quality of categorization: $F(1,31) = 1.11$; $MS_E = 430.51$; $p = .30$). This lack of significant main effects could be traced back to a cross-interaction between the two factors ($F(1,31) = 5.15$; $MS_E = 430.51$; $p < .05$). Participants who performed poorly in detecting structural similarities in the categorization task did not show any sequence effect ($t(18) = 0.83$; $p > .40$; two-tailed), whereas good categorizers performed better in the structure-blocked sequence compared to the surface-blocked sequence. However, the latter difference was not significant due to a lack of statistical power ($t(18) = -1.51$; $p < .15$; two-tailed). Further analyses revealed that as expected this pattern of results held only for non-corresponding problems.

For the transfer problems there was no main effect for problem sequence ($F < 1$). However, good categorizers committed fewer errors in the transfer task than did poor categorizers ($F(1,31) = 5.48$; $MS_E = 581.31$; $p < .05$). Finally, there was the same, but less accentuated cross-interaction as for isomorphic problems ($F(1,31) = 3.66$; $MS_E = 581.31$; $p < .10$). However, the respective paired comparisons were not significant (both $ps > .20$).

To conclude, the quality of problem categorization moderated effects of problem sequence as expected. Not only did the sequence effect in favor of the surface-blocked sequence vanish for good categorizers, but participants who had been able to detect the problems' structural similarities in the categorization task even slightly benefited from a structure-blocked sequence. This can be interpreted as evidence that the structural awareness induced by the categorization task supported positive, activation-based transfer among succeeding structurally similar problems in this sequence. On the other hand, for poor categorizers no sequence effect could be observed. It might well be that the initially expected superiority of the surface-blocked sequence is overridden by minor effects of the categorization task that run into the opposite direction – adding up to a null effect.

Summary

In two experiments we provided evidence for the ambiguous role of surface similarities in the formation of sequence effects. On the one hand, presenting superficially similar, but structurally dissimilar problems successively can foster learning during problem solving by enabling the recognition of structural relations among problems. On the other hand, structure-blocked sequences may foster the transfer process in the case that structural relations among problems have already been detected. Accordingly, we obtained support for the transfer as well as for the near-miss hypothesis depending on whether problem solvers were aware of structural features prior to problem solving or not.

Acknowledgements

This work was supported by a scholarship of the Deutsche Forschungsgemeinschaft awarded to the first author of the paper as a member of the Graduate College for Cognitive Science at the Saarland University, Germany. We thank

Carina Kraemer, Frauke Lancker, Lisa Olbrich, and Tina Schorr for conducting the experiments.

References

- Blessing, S. B., & Ross, B. H. (1996). Content effects in problem categorization and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 792-810.
- Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.
- Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, 12, 306-355.
- Hinsley, D. A., Hayes, J. R., & Simon, H. A. (1977). From words to equations: Meaning and representation in algebra word problems. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, NJ: Erlbaum.
- Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition*, 15, 332-340.
- Mayer, R. E. (1981). Frequency norms and structural analysis of algebra story problems into families, categories, and templates. *Instructional Science*, 10, 133-175.
- Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra-word-problem comprehension and its implications for the design of learning environments. *Cognition and Instruction*, 9, 329-389.
- 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.
- Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124, 207-231.
- Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist*, 24, 113-142.
- Scheiter, K., & Gerjets, P. (2002). The impact of problem order: Sequencing problems as a strategy for improving one's performance. In W. D. Gray & C. D. Schunn (Eds.), *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 798-803). Mahwah, NJ: Erlbaum.
- Sweller, J. (1980). Hypothesis salience, task difficulty, and sequential effects on problem solving. *American Journal of Psychology*, 93, 135-145.
- VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In M. Posner (Ed.), *Foundations of cognitive science*. Mahwah, NJ: Erlbaum.
- Weisberg, R., DiCamillo, M., & Phillips, D. (1978). Transferring old associations to new situations: A nonautomatic process. *Journal of Verbal Learning and Verbal Behavior*, 17, 219-228.
- Winston, P. H. (1975). Learning structural descriptions from examples. In P. H. Winston (Ed.), *The psychology of computer vision*. New York: McGraw-Hill.