Title
Problem Representation in Experts and Novices: Part 1. Differences in the Content of Representation

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 22(22)

ISSN
1069-7977

Authors
Yarlas, Aaron S.
Sloutsky, Vladimir M.

Publication Date
2000

Peer reviewed

Problem Representation in Experts and Novices: Part 1. Differences in the Content of Representation

Aaron S. Yarlas (yarlas.1@osu.edu)
Center for Cognitive Science and School of Teaching & Learning, The Ohio State University
21 Page Hall, 1810 College Road, Columbus, OH 43210 USA

Vladimir M. Sloutsky (sloutsky.1@osu.edu)
School of Teaching & Learning and Center for Cognitive Science, The Ohio State University
21 Page Hall, 1810 College Road, Columbus, OH 43210 USA

Abstract

Two experiments examined the content of novice and expert representations for both surface and deep structural elements of arithmetic equations. Experiment 1, which used a forced-choice categorization task in which surface features of equations (e.g., digits) competed with deep structural principles of mathematics (associativity and commutativity), found that experts were more likely to focus on principles in their judgments than were novices, who focused more often on surface elements. Experiment 2, using a similar task, introduced trials in which only principled elements varied. Novices were able to focus on principled elements in this case, but failed to transfer these representations when surface features were reintroduced. These findings indicate that novices had knowledge of the principles, but that they did not attend to them when competing surface features were present.

Introduction

It has been well established that in various knowledge domains (e.g., physics, mathematics, or chess) experts approach problems in a manner different from that of novices (Chase & Simon, 1973; Chi, Feltovich, & Glaser, 1981; Larkin, 1983; Simon & Simon, 1978; Reed, Ackinclose, & Voss, 1990). In particular, while experts are more likely to focus on hidden relational properties of a problem, novices are more likely to focus on less important surface features of a problem. However, while there is some understanding of the content of mental representation (i.e., of which aspects of information are likely to be represented and which are likely to be left out), the process of construing the representation remains largely unknown. Do people attend to and encode those aspects that are left out, but then discard them, or do they fail to attend to and encode these "irrelevant" aspects?

The current paper (Part 1) focuses both on establishing differences in content of representation for experts and novices within a simple domain (arithmetic) and testing a number of viable explanations that could account for these differences. A subsequent paper (Part 2) focuses on examining differences in the process of construing representations for experts and novices.

There is a large body of literature indicating that in problem solving, reasoning, learning and transfer, and problem categorization, novices tend to focus on surface features rather than on deep relational properties. These effects have been demonstrated in a variety of knowledge domains, including chess (Chase & Simon, 1973), mathematics (Blessing & Ross, 1996; Schoenfeld & Herrmann, 1982; Bassok, 1996, 1997; Novick, 1988; Reed, et al, 1990; Silver, 1981), physics (Chi, et al 1981; Simon & Simon, 1978; Larkin, 1983; Larkin, McDermott, Simon, & Simon, 1980), and computer programming (Adelson, 1984). Similar effects have been observed in a variety of knowledge-lean domains, such as deductive and inductive inference. When presented with deduction problems, untrained reasoners often tended to ignore the argument's logic (i.e., its deep structure) while relying on the argument's surface features, such as content and believability (Evans, Newstead, & Byrne, 1993; Johnson-Laird & Byrne, 1991). When presented with induction and analogy problems, novices and young children also often ignored deep relational structure while relying on the surface features (Gentner, 1989; Holyoak & Koh, 1987).

While there is little disagreement that novices focus on surface features, it remains unclear why novices tend to focus on surface features and not on deep relational properties. One possible explanation of novices' tendency to represent surface features is that novices merely have little knowledge of deep structural relations. However, while this possibility is capable of explaining expert-novice differences in extremely knowledge-demanding domains, such as medical diagnostics, chess, or advanced physics, it falls short of explaining these differences in fairly simple domains, such as elementary mathematics and physics. For example, researchers examining novices' representations in mathematics and physics often drew examples from students' textbooks, thus reasonably assuming that students should be familiar with the deep structure underlying these problems (Chi, et al, 1981; Larkin, 1983; Novick, 1988). The credibility of the lack of knowledge explanation is further undermined by findings that even those novices who receive instruction in a domain often continue to focus on surface features rather than the deep structure of a problem. These has been demonstrated across a variety of knowledge domains, including mathematics (Morris & Sloutsky, 1998) and physics (Kaiser, McCloskey, & Proffitt, 1986; McClosskey, 1983). Finally, the fact that findings on novices' representations in knowledge-lean domains are compatible with those in knowledgerich domains makes the low knowledge explanation even less plausible. At this point, however, lack of knowledge cannot be ruled out as an explanation for differences in problem construal by experts and novices. It is also possible that surface features are used more frequently by novices, and, as a result, they are more available than deep relational properties (cf. Anderson, 1990). Henceforth, we will refer to this possibility as the availability explanation.

Another possibility that appears more credible is that even when novices know about deep relations and are capable of extracting these relations, they still fail to represent these relations because surface features are more prominently present in the problem. In failing to represent relational features, they may either fail to encode relations, or these relational features may lose attentional competition to more salient surface features. However, this representational processing explanation can only be tested if the above-described explanations are eliminated as possibilities. In the current paper, then, the focus is on establishing difference in content of representations of experts and novices within the domain of arithmetic, and then testing the knowledge and availability explanations. If differences between experts and novices are found, and the data are inconsistent with the predictions of the alternative explanations, then the way is cleared to test the representational process explanation.

The goal of the current studies, then, is to establish why experts and novices differ in the content of their problem representations. To achieve this goal, we deemed it necessary to control for knowledge factors, while manipulating representational factors. In controlling for knowledge factors, we (a) used simplified tasks and (b) selected only those deep properties that were well familiar to a wide range of participants. In particular, we selected the commutative and associative properties of arithmetic, because these principles are learned in the elementary school and revisited in the beginning of the middle school (Everyday Mathematics: Teacher's Reference Manual, 1998), and therefore are likely to be familiar to the majority of middle school students and college undergraduates.

In this paper, we present two experiments. In Experiment 1, experts and novices in mathematics were asked to group arithmetic equations. These groupings could be based either on the commonality of surface elements (e.g., digits used, the number of constituent elements in the equations) or on the commonality of a deep mathematical relation (principles of commutativity or associativity). In Experiment 2, we introduced a two-phase grouping task. During the first phase, deep relations were "unmasked," such that surface elements were not varied among the compared equations. During the second phase, the deep relations were "masked" again by reintroducing competing surface elements.

Experiment 1

The goal of this experiment was to validate the principles in question and to eliminate the possibility that expert-novices differences stem from differences in overall intelligence (or age) between novices and experts.

Method

Participants Five samples were selected for the current experiment. The first group, which will be referred to as the “younger children”, contained 20 first- and second-graders taken from an elementary school (*M* = 7.26 years, *SD* = 0.59; 8 girls and 12 boys). The second group, which will be referred to as the “older children”, contained 16 sixth-graders taken from a middle school (*M* = 12.10 years, *SD* = 0.38; 5 girls and 11 boys). Both of these groups were selected from schools located in an upper middle-class suburb of Columbus, Ohio.

The third group of participants consisted of 25 undergraduates in an introductory psychology course at a large Midwestern university who participated for course credit. This group had an average age of 19.78 years (*SD* = 1.38), with 11 women and 14 men.

These three groups of mathematics "novices" were contrasted with a group of mathematics "experts". This group consisted of 20 graduate students in a Mathematics department at the same university who participated for payment of ten dollars. This group had an average age of 28.88 years (*SD* = 6.05), with 7 women and 13 men.

However, differences between "experts" and "novices" were not limited to expertise. Experts were also older and they might represent a self-selected group with respect to an overall ability. Therefore, we deemed it necessary to select a matching group that would be similar to experts in terms of age and overall ability, while differing in the level of expertise. This matching group consisted of 16 graduate students in a History department at the same university who participated for a payment of ten dollars. This group had an average age of 29.93 years (*SD* = 4.67), with 8 women and 8 men.

Materials Five features of arithmetic equations were used in Experiment 1. Two of these features were considered "principled properties", in that they represented deep, relational principles of mathematical operations: the associativity and commutativity principles. The former states that for addition, subtraction, and multiplication, constituent parts can be decomposed and recombined in different ways (e.g., $a + b = [a - c + c] + b$). The latter states that the order of elements is irrelevant for addition and multiplication (e.g., $a + b = b + a$). The other three features were nonprincipled surface features that occur in arithmetic equations: (1) digits (e.g., 6, 3); (2) sign (e.g., -, +); and (3) the number of constituent terms in an equation. The numerical solutions of equations were controlled for by making these solutions either all equal or all equivalently different for each trial.

A forced-choice similarity paradigm was used in this experiment. Participants were presented with three cards at a time, a target card and two test cards, each which had printed on it an arithmetic equation. Participants were instructed to match the problem on the target card to one of the test problems with which they believed it was most similar. Each of the two test problems shared one feature with the target problem, and differed on the feature that the target shared with the other test problem, with all other features held constant. All five features were pitted directly against each other, with the exception of the two principled features, yielding a total of nine feature comparisons. For example, on one of the trial in which commutativity competed with digit, the Target problem was 6 + 3 + 4 = 3 + 4 + 6, the digit test problem was 6 + 3 + 8 = 3 + 4 + 10, and the commutativity test was 7 + 2 + 8 = 8 + 2 + 7.

There were four exemplar arithmetic equations representing each of the nine comparison sets, resulting in a total of 36 trials presented to participants. The numbers used in the arithmetic equations ranged from 1 to 15, and the operations used included addition, subtraction, and multiplication.

Procedure All participants were run individually by a male experimenter into a small, quiet room. Participants were instructed that they would be presented with math problems for which they were to group together problems that were similar. A warm-up trial was used to acquaint participants with the task. For the warm-up trial, the participant was presented with cards containing Gelman and Markman's (1986) blackbird-flamingo-bat figures. The target card, which depicted a blackbird that looked similarly to the bat and dissimilarly from the flamingo, was placed equidistantly below the flamingo and bat cards, which were the test items. The experimenter pointed to each of the two test items, and asked the participant "which of these is more like this," after which the experimenter pointed to the target item. After participants chose one of the test items, the experimenter asked the participant "why did you choose that one?" After the participant's verbal explanation (either based on physical similarity or the commonality of species), the experimenter pointed out that the other test item could have also been chosen based on the other attribute, and made the point that similarity can simultaneously occur across multiple dimensions. All participants showed understanding of this concept and of the task.

Four trials for each of the nine features-principle comparisons resulted in a total of 36 trials, which took approximately 30 minutes. Trial order was determined using a block randomization procedure. The positioning of the test items in relation to the target (i.e., left or right) was counterbalanced across comparison type.

Results and Discussion

The main goal of this experiment was to examine participants' knowledge of principles in question. To achieve this goal, we considered as choices indicating knowledge only those for which the participants' explanation of the choice was consistent with the principle. This was done because participants could select principled test stimuli for a reason that might have nothing to do with the principle in question. Only explanations *directly* referring to the principle in questions were considered choice-consistent. The proportion of consistent choices for each principle is the dependent variable used in the forthcoming analyses.

The degree to which participants in each sample made explanation-consistent principled choices was analyzed using a one-way ANOVA for each principle across samples. Table 1 presents overall percentages of explanation-consistent principled choices aggregated across trials by principles and age groups. The ANOVA for explanation-consistent associativity choices yielded a significant difference among the samples in the proportion of choices made, F (4, 92) = 30.72, MSE = .07, p < .001. The percentage of explanation-consistent associativity choices increases monotonically across the five samples. The ANOVA for explanation-consistent commutativity choices also indicated that there was a significant difference among the samples in the proportion of choices made, F (4, 92) = 23.61, MSE = .08, p < .001. As shown in Table 1, the percentage of explanation-consistent commutativity choices also increases monotonically across the five samples.

Table 1: Means and standard deviations (in parentheses) for percentage of explanation-consistent principled choices in Experiment 1.

	Principle	
Sample	Associativity	Commutativity
Younger children	0.00 (0.00)	0.00 (0.00)
Older children	2.08 (6.04)	9.03 (17.32)
Undergraduates	10.66 (26.15)	20.44 (30.71)
History grads	26.39 (38.89)	36.81 (35.54)
Math grads	80.00 (36.34)	77.22 (35.95)

Bonferroni post-hoc tests (with α = .05) were used to compare the mean proportion of explanation-consistent principled choices for each sample. These tests yielded identical patterns for both the associativity and commutativity principles, indicating that there were not significant differences in the proportion of explanation-consistent principled choices by younger children, older children, and undergraduates, that History graduate students made significantly more explanation-consistent principled choices than younger children, and that Mathematics graduate students made significantly more explanation-consistent principled choices than each of the other four samples.

Results from Experiment 1 point to several important regularities. First, experts were found to consistently represent principles when categorizing arithmetic equations, whereas novices were more likely to focus on surface features rather than on principles; even when novices did focus on principles, they did so inconsistently. Second, very few younger children exhibited knowledge of principles in question. Third, expert-novice differences were not limited to age or general intelligence: history graduate students and math experts, equally aged groups with similar levels of overall intelligence, exhibited large differences in using deep principled features. Thus the experiment allows us to eliminate the possibility that general ability or development account for expert-novice differences.

However, Experiment 1 left an important question unanswered: it remains unknown why many novices failed to focus on deep principled features. One possibility is that novices merely lack knowledge of these principles. A second possibility is that surface elements are more available than deep relational features due to a more frequent use of the former. The goal of Experiment 2 is examine the two possibilities.

Experiment 2

To accomplish the main goal of this experiment (i.e., to distinguish among the above mentioned possibilities), it was necessary to observe whether novices represent principled features when these features do not compete with surface elements. In the current study, then, participants were given a number of trials in which the target problem shared a principled feature with one of the test problems, and shared no unique surface features with the other test problem. We refer to these trials as "unmasked" since principled features are no longer attentionally "masked" by surface elements.

In addition, in the current experiment the "unmasked" trials are followed by "masked" trials equivalent to the trials in Experiment 1, in which the surface elements are reintroduced to compete with principled features in participants' similarity judgments. This will enable the examination of the degree to which representations of principled features will be maintained, or whether the surface features will draw attention away from principled features, such that there is no transfer of representation due to the positive learning set. If the former is true, then it is expected that participants' explanation-consistent principled choices will be more frequent for the subsequent "masked" trials than they were in Experiment 1; if the latter is true, then there should be no difference between the frequency of these choices.

If novices are more likely to make explanation-consistent principled choices in "unmasked" trials, it indicates that they have knowledge of the principles in question, thus undermining the lack of knowledge explanation. If novices are more likely to represent principles in "unmasked" trials but there is no transfer to "masked" trials, this finding would undermine the availability explanation.

Method

Participants Three samples were selected for Experiment 2, each representing a different age group. Two of the groups, the "younger children" and the "older children" used the same participants from Experiment 1; Experiment 2 was conducted approximately four months after Experiment 1 for both samples. The third group of participants consisted of 19 undergraduates in an introductory psychology course at a large mid-western university who participated for course credit. This group had an average age of 21.79 years (*SD* = 6.49), with 12 women and 7 men.

Materials and Procedure The same principled features (i.e., associativity and commutativity) and surface features (i.e., digit, sign, and number of elements) used in Experiment 1 were used in Experiment 2. The same nine comparisons used in the previous experiment were again used here for the last 27 trials (three trials for each of the nine comparisons). In addition, in the current experiment, the first eight trials consisted of 'unmasked' comparisons, thus leading to a total of 35 trials.

For the "unmasked" trials, each of the two principled features (i.e., commutativity and associativity) was compared four times against 'control' problems. For these trials, the two test problems were equivalently similar to the target on nonprincipled features, while one test problem shared a principled feature with the target problem. For example, for an unmasked-commutativity trial, the target equation was $2 + 6 + 8 = 6 + 8 + 2$, the commutativity Test equation was $11 + 1 + 5 = 5 + 1 + 11$, and the control Test equation was $3 + 9 + 5 = 12 + 4 + 1$.

Results and Discussion

We first analyze performance in "unmasked" and "masked" trials across the three groups of novices. For purposes of clarity, we will refer to "masked" trials in Experiment 1 as *Masked 1*, whereas "masked" trials in Experiment 2 will be referred to as *Masked 2*. Again, when analyzing performance, we will focus on the proportion of choices made by participants for only the trials in which mathematical principles were present, and we will consider only those choices for which the participants' explanation of the choice was consistent with the principle. We first present the analyses of Unmasked and Masked 2 trials, followed by comparisons across Masked 1, Unmasked, and Masked 2 conditions.

The degree to which participants in each sample made explanation-consistent principled choices in the Unmasked comparisons was analyzed using a one-way ANOVA for each principle across the three samples. The ANOVAs for explanation-consistent associativity and commutativity choices revealed significant differences among the samples in the proportion of choices made, Fs $(2, 51) > 5.42$, ps $< .01$. As evidenced in Table 2, the percentage of explanation-consistent principled choices increased monotonically with age.

Table 2: Means and standard deviations (in parentheses) for percentage of explanation-consistent principled choices in Experiment 2.

Sample	Trial Type: Unmasked	Trial Type: Masked 2
	Associativity	
Younger Children	0.00 (0.00)	1.17 (5.10)
Older Children	20.31 (29.18)	6.94 (15.11)
Undergraduates	27.63 (36.22)	18.13 (27.27)
	Commutativity	
Younger Children	15.79 (30.29)	1.75 (7.65)
Older Children	67.19 (29.89)	15.28 (22.18)
Undergraduates	90.79 (20.77)	35.09 (36.71)

Bonferroni post-hoc tests (with $\alpha = .05$) were used to compare the mean proportion of explanation-consistent principled choices for Unmasked trials for each sample. For associativity trials, this test indicates only one statistically significant difference among samples, that undergraduates made more explanation-consistent associativity choices than younger children. However, for commutativity trials, all between-sample comparisons were statistically significant. Aggregated across both principles, less than 10% of the younger children's responses were principle-based, while almost 50% of older children's responses and over 60% of undergraduate students' responses were principle-based.

It should be noted that there were large differences in the proportion of participants focusing on commutativity and associativity, with the former being greater than the latter. However, even for associativity, where effects were smaller than for commutativity, around 50% of older children and undergraduates provided at least one explanation-consistent principled choice, thus exhibiting knowledge of the principle in question.

The degree to which participants in each sample made explanation-consistent principled choices in the Masked 2 comparisons was also analyzed using a one-way ANOVA for each principle across samples. The ANOVAs for explanation-consistent associativity and commutativity choices yielded a significant difference among the samples in the proportion of choices made, Fs $(2, 51) > 4.1$, ps $< .05$. Again, as evidenced in Table 2, the percentage of explanation-consistent principled choices increases monotonically with age.

Bonferroni post-hoc tests (with $\alpha = .05$) were again used to compare the mean proportion of explanation-consistent principled choices for Masked 2 trials for each sample. For both principles, this test indicates only one statistically significant difference among samples that undergraduates made more explanation-consistent principled choices than did younger children. These data in conjunction with the results of the Unmasked condition suggest that even when participants knew the principle in question, they often focused on surface features.

Overall proportions of explanation-consistent principled choices in Masked 1, Unmasked, and Masked 2 trials aggregated across the principles and broken down by sample are presented in Figure 1. Participants' explanation-consistent principled choices on Unmasked trials generally increased in comparison to their choices on Masked 1 trials. Younger children gave more explanation-consistent commutativity choices for Unmasked trials than for Masked 1 trials ($t = 2.27$, $p < .05$), though there was not a significant difference in the amount of explanation-consistent associativity choices, which is due to a floor effect. Older children gave more explanation-consistent principled choices for Unmasked trials than for Masked 1 trials for both principles ($t = 2.65$, $p < .02$ for associativity, and $t = 9.8$, $p < .001$ for commutativity). Undergraduates gave more explanation-consistent commutativity choices for Unmasked trials than for Masked 1 trials ($t = 8.59$, $p < .001$), though there was a marginally significant difference in the amount of explanation-consistent associativity choices ($t = 1.81$, $p = .078$). These differences indicate that unmasking increased the proportion of principled choices in all samples.

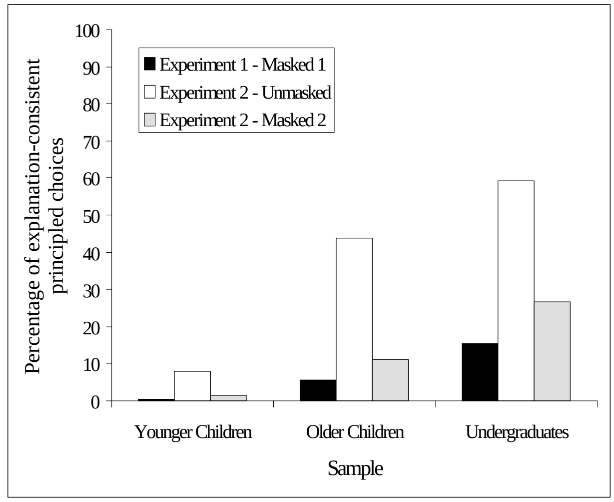

Figure 1. Percentage of explanation-consistent principled choices for each sample for unmasked" and "masked" trials in Experiment 2, and "masked" trials in Experiment 1.

Participants' explanation-consistent principled choices on Masked 2 trials generally decreased in comparison to their choices on Unmasked trials. Younger children gave more explanation-consistent commutativity choices for Unmasked trials than for Masked 2 trials ($t = 2.37$, $p < .05$), though there was not a significant difference in the amount of explanation-consistent associativity choices, which is due to a floor effect. Older children gave more explanation-consistent principled choices for Unmasked trials than for Masked 2 trials for both principles ($t = 2.74$, $p < .02$ for associativity, and $t = 8.15$, $p < .001$ for commutativity). Undergraduates gave more explanation-consistent commutativity choices for Unmasked trials than for Masked 2 trials ($t = 7.23$, $p < .001$), though there was a marginally significant difference in the amount of explanation-consistent associativity choices ($t = 1.81$, $p = .078$). These differences indicate that there was not pure transfer of representations from Unmasked to Masked 2 trials: once principled features had to compete again with surface features, the number of explanation-consistent principled choices decreased markedly.

An important question is whether the transfer led to a significant increase of explanation-consistent principled choices compared to when participants were never exposed to Unmasked trials. That is, whether being exposed to a positive learning set significantly increased subsequent attention to principles. To answer this question, we compared participants' explanation-consistent principled choices on the Masked 1 and Masked 2 trials. While the proportions of explanation-consistent principled choices are somewhat larger for each sample and each principle on Masked 2 trials than for Masked 1 trials (as evidenced in Figure 3), t-tests for each comparison revealed that none of these differences are statistically significant. Thus, the positive learning set of the Unmasked trials had a nonsignificant effect on the degree to which participants represented principled features of mathematics problems.

Overall, results of Experiment 2 indicate that 94% of the middle school participants and 100% of the undergraduate participants exhibited knowledge of principles in question (i.e., provided an explanation-consistent principled choice on at least one trial), focusing on these principles in Unmasked trials. This finding severely undermines the lack of knowledge explanation. At the same time, the increase in younger children's principled choices due to "unmasking" was rather small, which points to a lack of knowledge. However, even in the two older groups, once nonprincipled features were reintroduced, representation of principled properties attenuated to levels similar to Experiment 1, a finding that undermines the availability explanation.

Conclusion

The results of the two reported experiments establish a difference in the content of expert and novice representations for arithmetic problems. These results suggest that the observed differences do not stem from a lack of knowledge of deep principles by novices. The results further suggest that differences in content of problem representations in experts and novices may stem from different processing mechanisms underlying the construal of problem representations in experts and novices. The research presented in Part 2 will focus the examination of the processes of construal of problem representations by expert and novices.

Acknowledgments

This research has been supported by a grant from James S. McDonnell Foundation to the second author.

References

Adelson, B. (1984). When novices surpass experts: The difficulty of a task may increase with expertise. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 483-495.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Bassok, M. (1996). Using content to interpret structure: Effects on analogical transfer. *Current Directions in Psychological Science*, *5*, 54-58.

Bassok, M. (1997). Two types of reliance on correlations between content and structure in reasoning about word problems. In L. D. English (Ed.), *Mathematical reasoning: Analogies, metaphors, and images*. Mahwah: NJ: Lawrence Erlbaum Associates.

Blessing, S. B., & Ross, B. H. (1996). Content effects in problem categorization and problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *22*, 792-810.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55-81.

Chi, M. T. H., Feltovich, P. G., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121-152.

Evans, J. St. B. T., Newstead, S. E.; Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction*. Hove, England: Lawrence Erlbaum Associates.

Everyday Mathematics: Teacher's Reference Manual (1998). Chicago, IL: Everyday Learning.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*, 183-209.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning*. New York: Cambridge University Press.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition*, *15*, 332-340.

Johnson-Laird, P., & Byrne, R. (1991). *Deduction.* Hove, UK: Lawrence Erlbaum.

Kaiser, M. K., McCloskey, M., & Proffitt, D. R. (1986). Development of intuitive theories of motion: Curvilinear motion in the absence of external forces. *Developmental Psychology*, *22*, 67-71.

Larkin, J. (1983). The role of problem representation in physics. In D. Gentner & A. Stevens (Eds.), *Mental models*. Hillsdale, NJ: Erlbaum.

Larkin, J. H., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Models of competence in solving physics problems. *Cognitive Science*, *4*, 317-345.

McClosskey, M. (1983). Naïve theories of motion. In D. Gentner & A. Stevens (Eds.), *Mental models*. Hillsdale, NJ: Erlbaum.

Morris, A. K., & Sloutsky, V. M. (1998). Understanding of logical necessity: Developmental antecedents and cognitive consequences. *Child Development*, *69*, 721-741.

Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*, 510-520.

Reed, S. K., Ackinclose, C. C., & Voss, A. A. (1990). Selecting analogous problems: Similarity versus inclusiveness. *Memory & Cognition*, *18*, 83-98.

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *8*, 484-494.

Silver, E. A. (1981). Recall of mathematical problem information: Solving related problems. *Journal of Research in Mathematics Education*, *24*, 117-135.

Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. S. Siegler, (Ed), *Children's thinking: What develops?* Hillsdale, NJ: Lawrence Erlbaum Associates.