Title
Problem Representation in Experts and Novices: Part 2. Underlying Processing Mechanisms

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

Authors
Sloutsky, Vladimir M.
Yarlas, Aaron S.

Publication Date
2000

Peer reviewed

Problem Representation in Experts and Novices: Part 2. Underlying Processing Mechanisms

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

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

Abstract

It has been well established that experts and novices focus on different aspects of problems, with novices focusing more on surface features rather than on deep principled features of a problem. What is less clear are the mechanisms that underlie these differences in construal of problem representation. The current study, which uses an 'old/new' recognition procedure, examines expert and novice representation of arithmetic equations in which the deep relational properties (i.e., principles of commutativity and associativity) were well known to both groups. Results indicate that both novices and experts encode both surface and principled features in the same serial manner, with surface features preceding principled features for both. At the same time, only for novices and not for experts, surface features compete with deep features, thus requiring additional resources to inhibit this attentional competition.

Introduction

Mental representation is a central component of several fundamental cognitive processes, including categorization, reasoning, decision making, and problem solving. For example, the way an entity is categorized depends on the content of an organism's mental representation regarding this entity and the similarity of this representation to a composite representation stored in memory (Estes, 1994; Nosofsky, 1988). In addition, the way people reason from propositions and what they infer from these propositions depends on the manner in which these propositions are mentally represented (Byrne, 1989; Johnson-Laird & Byrne, 1991; Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni, 1999). Finally, the content of a mental representation determines the approaches and strategies people use when they attempt to solve problems (Kaplan & Simon, 1990; Larkin & Simon, 1987; Newell & Simon, 1972). Of course, the content of mental representation may depend on knowledge of the conceptual and relational structure of the domain, and transformational procedures and algorithms (Anderson, 1982; 1990; Case & Okamoto, 1996; Gelman & Meck, 1986, 1992; Hiebert & Lefevre, 1986; Rittle-Johnson & Alibali, 1999). For example, the problem " Bill has eight marbles and Jill has six times more" would be represented as "8 x 6 = ?", only if the person has knowledge of and can abstract basic multiplication algorithms.

As noted above, there is a distinction between the content of a mental representation (or what is represented) and the process of construing this content (or what is attended to, encoded, and stored). The process of construing mental representations remains largely unknown, and is the focus of this paper. However, there are several important regularities that have been established with respect to the content of mental representation that are important for the study of the process of construing of mental representation.

In Part 1 of this paper (Yarlas & Sloutsky, 2000) and elsewhere (Yarlas & Sloutsky, 1999), we describe a large body of literature indicating that in problem solving, reasoning, learning and transfer, and problem categorization, novices and experts construe representations that differ in their content. In particular, novices tend to focus on surface features of the problem, whereas experts tend to focus on deep relational features (e.g., Chase & Simon, 1973; Chi, Feltovich, & Glaser, 1981; Gentner & Toupin, 1986; Kotovsky & Gentner, 1996; Larkin, 1983; Simon & Simon, 1978; Yarlas & Sloutsky, 1999). These effects have been demonstrated across a variety of knowledge-rich and knowledge-lean domains.

However, in spite of these well-established expert-novice differences, it remains unclear what accounts for these differences. Do differences occur because experts have knowledge of deep relational properties and novices do not? Do they occur because novices are less intelligent or younger than experts are, and they cannot grasp deep relational properties? Do experts and novices differ in processes underlying the construal of a problem representation? Or do differences stem from a combination of these factors?

In Part 1 of this paper (Yarlas & Sloutsky, 2000), we focused on expert-novice differences in the content of mental representations. It was demonstrated that when tasks are sufficiently simple and deep relational properties are well known, neither differences in knowledge, intelligence, nor development can fully account for the observed differences between novices and experts. In a series of experiments designed to distinguish among these possibilities, tasks were constructed that included principles of arithmetic familiar to novices, and surface features that were completely superfluous with respect to deep relational features. In particular, they asked participants varying in age and degree of expertise to sort mathematical equations that could have common surface elements (e.g., commonality of numbers or the same number of constituent addends in the equation) or common deep mathematical principles (e.g., commutativity or associativity). Results indicated that only mathematics experts consistently focused on principles, whereas novices, regardless of age and intelligence, focused mostly on surface features. However, elimination of surface features led to substantial increase in focusing on principles. Interestingly, the reintroduction of surface features reduced participants' focus on principles to their original low levels. These and other manipulations allowed us to argue that differences between novices and experts stem from differences in processes underlying the construal of a problem representation. However, if novices have knowledge of the principles in question yet still fail to represent them, then several questions arise about processes underlying problem representations in novices and experts. Do novices initially encode both deep and surface features, but later discard the deep relational properties, or do they simply fail to encode the deep relational properties? And what are the processing mechanisms underlying problem representations in experts: do experts encode and discard surface features, or do they ignore these features from the very beginning?

To answer these questions, we used an 'old/new' recognition paradigm in the current experiment. This paradigm affords the creation of a set of foils, such that patterns of hits and false alarms point to which aspects of problems have been encoded and committed to memory and which aspects have been left out. In the study phase, participants were presented with a set of arithmetic equations. These equations all utilized a principled property, either associativity or commutativity. 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., $a + b + c = b + c + a$). In addition, these equations all used consistent levels of two surface elements: all equations used numbers ranging between 1 and 9, and all used either 5 or 6 numbers in the equation. In the recognition phase of the experiment, in addition to 'old' items, four combinations of 'new' equations were presented as foils. Half of these foils, which we refer to as 'feature +' foils, maintained the same levels of surface features as used in the learning phase (i.e., numbers ranging between 1 and 9, and either 5 or 6 numbers in the equation), while the other half of the foils, which we refer to as 'feature -' foils, violated these categories (i.e., numbers greater than 9, and either 4 or 7 numbers in the equation). Also, half of the foils, which we refer to as 'principle +' foils, maintained the use of one of the two principled properties, while the other half, which we refer to as 'principle -' foils, did not use any principled properties in the equation. The two levels of the two kinds of properties (feature being either + or -, and principles being either + or -) were fully-crossed, thus creating four combinations of foils: feature + /principle + (F+/P+), feature + /principle - (F+/P-), feature -/principle + (F-/P+), and feature -/principle - (F-/P-). For example, for the equation $5 + 3 + 6 = 3 + 6 + 5$ in the study phase, the following foils were presented in the recognition phase: (1) $5 + 3 + 6 = 3 + 6 + 5$ (Old), (2) $7 + 4 + 2 = 4 + 2 + 7$ (F+/P+), (3) $5 + 3 + 6 = 3 + 4 + 7$ (F+/P-), (4) $11 + 9 = 9 + 11$ (F-/P+), and (5) $14 + 7 = 9 + 12$ (F-/P-).

The goal of this paper is to elucidate processes underlying problem representations in novices and experts. In this article, we consider and test a number of possible processing models for both novices and experts, which are summarized in Table 1.

Table 1: Summary of considered processing models

Novice Model 1	Encode only surface features with no encoding of deep structural features
Novice Model 2	Encode both surface and deep structural features; attentional competition between surface and structural features, with surface features winning
Expert Model 1	Encode only deep structural features with no encoding of surface features
Expert Model 2	Encode both deep structural and surface features; attentional competition between structural and surface features, with structural features winning
Expert Model 3	Encode both deep structural and surface features; no attentional competition

For this task, if novices encode only surface features and not relational features, they should rapidly respond "Old" when surface features are present and they should rapidly respond "New" when surface features are absent (Novice model 1). Similarly, if experts encode only principles and not surface features, they should rapidly respond "Old" when principles are present and they should rapidly respond "New" when principles are absent. If either group encodes both principles and features, they should exhibit more complex patterns of responses (Expert model 1).

There is preliminary evidence (Yarlas & Sloutsky, 1999) that novices do encode both surface and deep features, but discard the latter in the course of attentional competition (Novice model 2). However, while processing mechanisms underlying problem representations in novices require further clarifications, these mechanisms in experts remain unclear. One possibility is that experts start construing problem representations from deep rather than from surface (Expert model 2). An alternative possibility is that experts construe representations in a manner similar to that of novices, except that there is no attentional competition in experts (Expert model 3). Of course, it is also possible that experts construe representations in a parallel manner, in which case their response latencies should exhibit small or no differences across the foils.

The alternative response patterns derived from the models summarized in Table 1 are presented in Table 2. These predictions are based on the following two assumptions: (1) both experts and novices process properties of problems in a serial manner and (2) each additional step in processing leads to increase in latencies. Both assumptions were previously corroborated using this task with novices (Yarlas &

Table 2: Patterns of responses and latencies predicted by alternative models for novices and experts

	Foil Types and Patterns of Responses				
Models of responses	Old targets	F+/P+	F+/P-	F-/P+	F-/P-
Novices Model 1 (Response type)	OLD	OLD	OLD	NEW	NEW
Novices Model 1 (Latency)	Fast	Fast	Fast	Fast	Fast
Novices Model 2 (Response type)	OLD	OLD	NEW	NEW	NEW
Novices Model 2 (Latency)	Slow	Slow	Very Slow	Fast	Fast
Experts Model 1 (Response type)	OLD	OLD	NEW	OLD	NEW
Experts Model 1 (Latency)	Fast	Fast	Fast	Fast	Fast
Experts Model 2 (Response type)	OLD	OLD	NEW	NEW	NEW
Experts Model 2 (Latency)	Slow	Slow	Fast	Slow	Fast
Experts Model 3 (Response type)	OLD	OLD	NEW	NEW	NEW
Experts Model 3 (Latency)	Slow	Slow	Slow	Fast	Fast

Sloutsky, 1999). Because of these assumptions, the parallel processing model is absent from Table 1; however we do not discount the possibility of parallel processing in experts. Note that predictions presented in Table 2 are qualitative, in that they do not specify accuracy or latency across the conditions, but rather point to (a) patterns of recognition responses and (b) directions of differences in latencies.

Note that the tables have two critical components. First, in novices, responses to F+/P- foils afford either corroboration or elimination of Model 1 for novices (see Table 1), whereas in experts, responses to F-/P+ foils afford corroboration or elimination of Model 1 for experts (see Table 1). Second, within experts and novices, patterns of differences in latencies afford the selection of the more plausible model as well as the description of specific processing components. Specifically, latencies in experts' responses to F+/P- items will allow for discriminating between Model 2 and Model 3 for experts. In short, patterns presented in the table should allow us to distinguish between processing models in novices and experts presented in Table 1.

Method

Participants

Two samples, representing novices and experts, were used in this study. The novice group included twenty-three undergraduates in an introductory psychology course at the Ohio State University who participated for partial course credit. This sample had an average age of 19.2 years (*SD* = 0.9 years), with 12 women and 11 men. The expert group included twelve graduate students in the Mathematics Department at the same university who participated for a payment of twenty dollars. This sample had an average age of 27.6 years (*SD* = 5.8 years), with 3 women and 9 men.

Materials and Procedure

The materials and procedures used in this study were identical for participants in both the novice and expert samples. All participants were run individually with stimuli presented by a personal computer using SuperLab software (Cedrus Corporation, 1999).

The experiment consisted of three phases: the study phase, the distraction phase, and the recognition phase. In the study phase, participants were presented with thirty arithmetic equations, which they had been instructed to memorize. All thirty equations used addition, used numbers ranging from 1 to 9, contained either 5 or 6 numbers, and used either the associative or commutative principle (half for each). Each equation was centered and presented in dark type on a white screen for ten seconds, with a two-second interval between each, during which only the white background was seen. The order of equations was randomized across participants.

A distraction phase followed the study phase for the purpose of clearing participants' short-term memory. For the distraction task, participants were presented with ninety letters, for which they had been instructed to indicate whether the letter was a vowel or a consonant. This phase took approximately three minutes.

Following the distraction phase was the recognition phase. Participants were told that they would be presented with a number of arithmetic equations, some of which had been presented to them earlier and some of which had not been presented earlier, and that they were to decide whether each equation was 'old' or 'new'. There were a total of sixty equations presented in the recognition phase. The order of equations presented in this phase was randomized across participants. There were five categories of foils, with twelve exemplars for each category. Recall that these foils included: (1) Old targets that had been presented earlier in the learning phase, (2) F+/P+ equations, which used similar surface features and used either the commutativity or associativity principle as in the original equations, (3) F+/P- equations, which used similar surface features as the original equations but did not use either the commutativity or associativity principle, (4) F-/P+ equations, which used surface features different from those used in the original equations but used either the commutativity or associativity principle, and (4) F-/P- equations, which used surface features different from those used in the original equations and did not use either the commutativity or associativity principle.

Results and Discussion

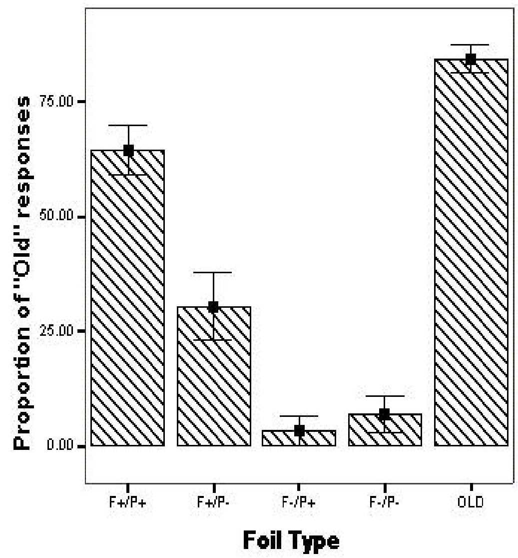
In this section, we will first discuss the accuracy of recognition and latencies of responses for novices, and then for experts. For each group, we will first examine overall accuracy of response to the foils (i.e., correct acceptance of Old targets and correct rejection of all foils). We will then compare participants' "Old" responses and latencies across the foil types. Note that for all foils except F+/P+, we compared latencies for correct responses only. Because we expected a large number of false alarms for F+/P+ foils, for these foils, latencies for both correct and incorrect responses were used in the analyses.

Novices exhibited high overall accuracy for most of the foils, correctly accepting Old targets and correctly rejecting F-/P+, F-/P-, and F+/P- foils. They mostly false alarmed, however, on F+/P+ foils. The latter finding is expected because F+/P+ foils were categorically indistinguishable from Old targets, since both surface features and principled features present in Old targets were also present in F+/P+ foils. More specifically, results indicate that accuracy rates (i.e., hits for Old Targets and correct rejections for the other foils) for F+/P- (M = 0.69, SD = 0.35), F-/P- (M = 0.93, SD = 0.20), F-/P+ (M = 0.97, SD = 0.16), and Old targets (M = 0.84, SD = 0.15) were significantly higher than chance (all ts(22) > 9.4, ps<.001), whereas for F+/P+ (M = 0.36, SD = 0.26) accuracy was significantly lower than chance, $t(22) = -6.4$, $p < .001$. These results indicate that these participants took the task seriously and were providing rather accurate responses.

Percentages of "Old" responses and latencies for novices are presented in Figure 1. A one-way repeated measures ANOVA points to significant differences among foils for novices (F (4, 88) = 53.9, MSE = 542.7, p < .0001. Paired-samples t-tests indicated the following the following direction in the proportion of "Old" responses: Old targets > F+/P+ > F+/P- > F-/P+ = F-/P-, all ts(22) > 3, all Bonferroni adjusted ps < .05 for differences.

Novices' latencies to different foils are also presented in Figure 1. These measures were also subjected to a one-way repeated measures ANOVA. The analysis indicates significant differences among the foils, F (4, 76) = 15.48, p < .001. Planned comparisons revealed that F+/P- latencies were significantly higher than those for Old targets, t(20) = 3.4, p < .005, whereas latencies of F-/P- and F-/P+ foils were significantly lower than those of the Old targets, ts(21) > 3.5, ps< .005.

These data allow us to rule out Model 1 presented in Table 1 -- novices did not base their responses solely on the presence or absence of surface features. When surface features were absent (F-/P- and F-/P+ foils) participants produced fast and accurate "New" responses; however, when surface features were present, novices did not always produce "Old" answers. Rather, novices' responses were mediated by the presence or absence of principled features. In particular, when both surface and principled features were present (Old targets and F+/P+ foils) novices generally responded "Old". These responses were slower than those for F-/P- and F-/P+ foils. Finally, when surface features were present but principles were absent (F+/P- foils), participants in general accurately rejected these foils, but latencies for these correct rejections were significantly higher than latencies for Old targets. These findings support the notion of the attentional competition between the two types of features (see Table 1, Novice model 2), pointing to a relative difficulty for participants to inhibit the salient surface feature and reject the foil. Of course, these data raise an interesting question of whether or not experts would also exhibit attentional competition between deep relational and surface features.



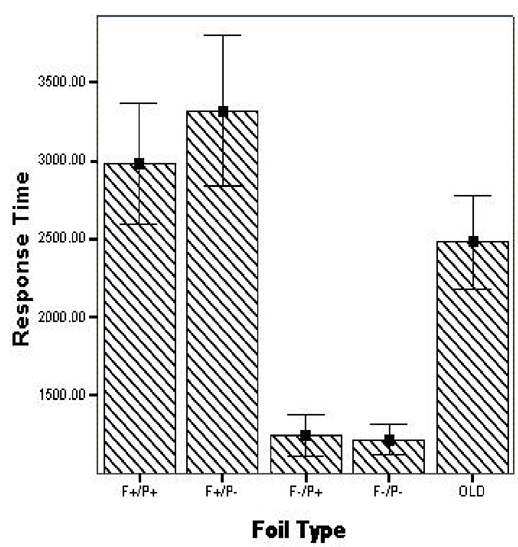


Figure 1. Proportion of novices' "Old" responses and response times (in milliseconds) across foil types in the recognition phase.

Similarly to novices, experts exhibited high overall accuracy for most of the foils, correctly accepting Old targets and correctly rejecting F-/P+, F-/P-, and F+/P- foils. They too mostly false alarmed, however, on F+/P+ foils. More specifically, accuracy rates (i.e., hits for Old Targets and correct rejections for the other foils) for F+/P- (M = 0.96,

SD = 0.06), F-/P- (M = 1.00, SD = 0.00), F-/P+ (M = 0.98, SD = 0.04), and Old targets (M = 0.90, SD = 0.05) were significantly higher than chance (all $ts(11) > 58$, $ps<.001$), whereas for F+/P+ (M = 0.15, SD = 0.09) accuracy was significantly lower than chance, $t(11) = - 5.1$, $p < .001$. These results indicate that experts also took the task seriously and provided rather accurate responses.

Percentages of "Old" responses and latencies for experts are presented in Figure 2. A one-way repeated measures ANOVA points to significant differences among foils for experts (F (4, 44) = 768.5, MSE = 34.2, $p < .0001$. Paired-samples t-tests indicated the following the following direction in the proportion of "Old" responses: Old targets = F+/P+ > F+/P- = F-/P+ = F-/P-, all $ts(22) > 23$, all Bonferroni adjusted $ps < .0001$ for differences.

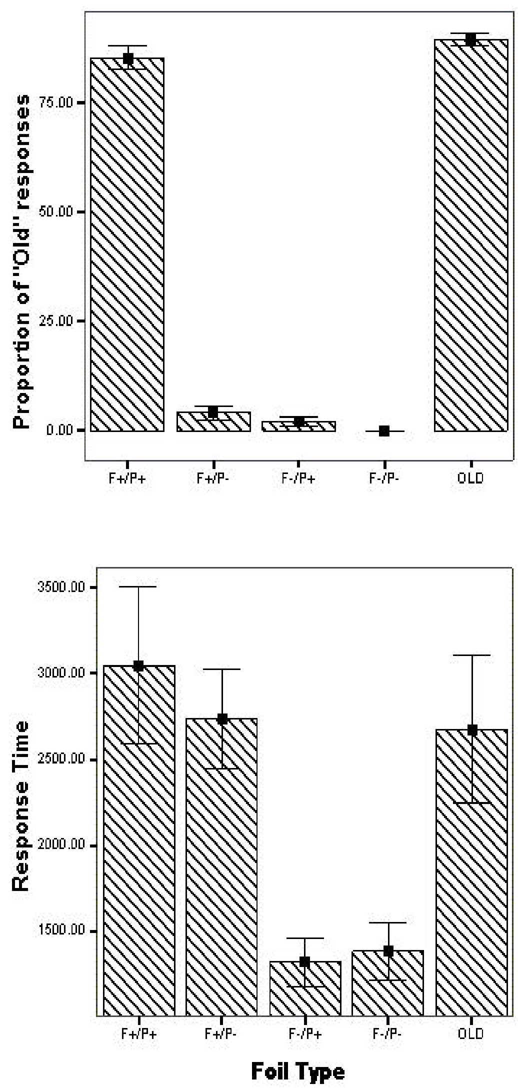


Figure 2. Proportion of experts' "Old" responses and response times (in milliseconds) across foil types in the recognition phase.

Experts' latencies to different foils are also presented in Figure 2. These measures were also subjected to a one-way repeated measures ANOVA. The analysis indicates significant differences among the foils, F (4, 44) = 18.60, $p < .001$. Planned comparison revealed that, in contrast to novices, F+/P- latencies for experts were not significantly different from those for Old targets, $t(11) = 0.2$, $p = .85$, but that latencies for F-/P- and F-/P+ foils were again significantly lower than those of the Old targets, $ts(11) > 4$, $ps< .005$.

The analysis of hits and false alarms allows us to eliminate Model 1 of expert responses presented in Table 1. Indeed, according to this model, experts should have responded "New' when principles were absent, and respond "Old" when principles were present. However, the F-/P+ foils almost invariable generated "New" responses, thus eliminating Model 1. Similarly, the analysis of latencies affords the elimination of Model 2. Recall that according to this model, experts should have more rapidly answered "New" when the principle was absent than when the feature was absent. However, the observed findings are consistent with Model 3 and not with Model 2, given that F-P+ foils were rejected faster than F+P- foils. Therefore, results of the experiment support Model 2 for novices and Model 3 for experts.

These findings point to important processing similarities and differences in experts and novices. First, both experts and novices exhibited serial processing. In addition, when construing problem representations, both experts and novices encode features first. At the same time, only novices experience competition between salient surface features and less salient deep principles. For the majority of novices, well known deep principles end up winning the competition; however, the competition takes time and effort. At the same time, experts represent both deep and surface features of the problem and do not experience such attentional competition. Recall that the experiment employed a very simple recognition task. In more resource demanding tasks, such as categorization, reasoning, or problem solving, deep relational features in novices may lose attentional competition to salient surface features. This loss would manifest itself in novices' tendency to focus on surface feature, while ignoring deep relational features (Chase & Simon, 1973; Chi, Feltovich, & Glaser, 1981; Gentner & Toupin, 1986; Kotovsky & Gentner, 1996; Larkin, 1983; Simon & Simon, 1978; Yarlas & Sloutsky, 1999).

The results have several potential implications. First, they lead to a better understanding of expertise, indicating that expert-novice differences persist even with most simple tasks (it is reasonable to expect that more complex tasks would result in more dramatic expert-novice differences). Second, the results have important educational implications, suggesting that salient surface features may deter rather than promote learning.

Conclusion

The reported findings indicate that even when a task is very simple, experts and novices construct problem representations differently. While both experts and novices encode deep as well as surface features of the problem, only for novices and not for experts, surface features compete with deep features, thus requiring additional resources to inhibit this attentional competition. These findings may or may not hold for less familiar deep principles or more complicated tasks. However, these results allow us to conclude that even when a task is very simple and deep principles are well known, experts and novices differ in processes underlying the construal of problem representations.

Acknowledgments

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

References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 369-406.

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

Byrne, R. M. (1989). Suppressing valid inferences with conditionals. *Cognition*. *31*(1), 61-83.

Case, R., & Okamoto, Y. (1996). The role of central conceptual structures in the development of children's thought.

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.

Estes, W. K. (1994). *Classification and cognition*. New York: Oxford University Press.

Gelman, R., & Meck, E. (1986). The notion of principle: The case of counting. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics*. Hillsdale, NJ: Erlbaum.

Gelman, R., & Meck, E. (1992). Early principles aid initial but not later conceptions of number. In J. Bideaud, C. Meljac & J. Fischer (Eds.), *Pathways to number: Children's developing numerical abilities*. Hillsdale, NJ: Erlbaum.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science*, *10*, 277-300.

Hiebert, J., & Levevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics*. Hillsdale, NJ: Erlbaum.

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

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, *67*, 2797-2822.

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M. S., & Caverni, J.-P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review*, *106*, 62-88.

Kaplan, C. A., & Simon, H.A. (1990). In search of insight. *Cognitive Psychology*, *22*, 374-419.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, *67*, 2797-2822.

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

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65-99.

Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*, 54-65.

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

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, *91*(1), 175-189.

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.

Yarlas, A. S., & Sloutsky, V. M. (1999). Representation of arithmetic principles by novices: Knowledge or attention? Manuscript under review.

Yarlas, A. S., & Sloutsky, V. M. (2000). Problem representation in experts and novices: Part 1. Differences in the content of representation. *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*.