Running Head: REPRESENTATION OF PRINCIPLES

MENTAL REPRESENTATION OF ARITHMETIC PRINCIPLES BY CHILDREN AND PREADOLESCENTS

Aaron S. Yarlas

and

Vladimir M. Sloutsky

The Ohio State University

PAPER IN PROGRESS

Address correspondence to:

Vladimir M. Sloutsky
21 Page Hall
1810 College Avenue
The Ohio State University
Columbus, OH 43210
Phone: (614) 688-5855
Fax: (614) 292-0321
Email: Sloutsky.1@osu.edu

Abstract

Three experiments examined the nature of children (7-8 year olds) and preadolescents' (11-12 year olds) representations for both surface and structural features of arithmetic equations. Experiment 1 used a forced-choice categorization task in which principles of arithmetic (associativity and commutativity) were either pitted against surface features of equations (e.g., the number of constituent terms) or were presented without competing surface features. In the former condition, both children and preadolescent focussed on surface features, whereas in the second condition preadolescents, but not children, focussed on the principled features, thus providing evidence that preadolescents know the principled features. Experiments 2 and 3 used a recognition task to examine the process of problem representation in preadolescents. Results of Experiment 2 indicated that preadolescents encoded both surface and principled features. However, when encoding time for problems was drastically reduced in Experiment 3, preadolescents did not reliably encode principles. Models describing possible processing mechanisms for preadolescent representations are discussed.

MENTAL REPRESENTATION OF ARITHMETIC PRINCIPLES BY CHILDREN AND PREADOLESCENTS

Acquisition of mathematical thinking has proven to be difficult for learners: across age and educational levels, people often tend to exhibit difficulties when performing mathematical reasoning or solving mathematical problems (Hinsley, Hayes, & Simon, 1977; Morris & Sloutsky, 1998; Rittle-Johnson & Alibali, 1999; Ross & Kilbane, 1997; Schoenfeld & Herrmann, 1982). What makes this knowledge domain particularly challenging? And what are the fundamental cognitive mechanisms underlying mathematical thinking? There is a large body of evidence indicating that mathematical thinking depends critically on the following components.

First, there is a knowledge component that includes knowledge of mathematical procedures and algorithms, such as operations on numbers, variables, or functions (Hiebert & Lefevre, 1986; Resnick & Omanson, 1987; Rittle-Johnson & Alibali, 1999; Siegler, 1991; VanLehn, 1996). This component also includes knowledge of the conceptual and relational structure of mathematics, such as knowledge of concepts of number and function, and properties of operations upon them, such as associativity of addition (Case & Okamoto, 1996; Hiebert & Lefevre, 1986; Gelman & Meck, 1986, 1992; Rittle-Johnson & Alibali, 1999).

Second, there is a representation component that includes adequate abstraction and encoding of the relational structure from particular problems (Larkin, 1983; Resnick & Omanson, 1987; Larkin & Simon, 1987). For example, to solve the problem "Bill has eight marbles and Jill has six times more; how many marbles does Jill have?", one has to abstract the problem structure that could be represented as a solvable equation, such as "8 x 6 = ?".

Third, there is an execution component that includes application of appropriate procedures and algorithms to this abstracted structure (see Hiebert & Lefevre, 1986; VanLehn, 1996, for reviews). It has been argued that the application of appropriate procedures and algorithms is trivially easy to execute once these algorithms are well known (cf., Siegler, 1976) and once the relational structure of the problem has been extracted. Therefore, in the reported research, we specifically focus on the first two components: knowledge of the relational structure of arithmetic, and the ability to extract and mentally represent this structure. In particular, we examine knowledge and representation of simple relational (i.e., principled) properties of arithmetic by elementary school children (7-8 year olds) and preadolescents (11-12 year olds).

There is ample research indicating that in problem solving, reasoning, learning and transfer, and problem categorization, both child and adult novices often fail to represent deep structural properties of problems, such as an equation leading to the solution of a problem. They focus instead on a variety of surface features, such as the problem’s story line. This tendency has been demonstrated in a variety of knowledge domains, including chess (Chase & Simon, 1973), mathematics (Blessing & Ross, 1996; Hinsley, Hayes, & Simon, 1977; Schoenfeld & Herman, 1982; Bassok, 1996, 1997; Novick, 1988; Reed, Ackinclose, & Voss, 1990; Ross & Kilbane, 1997; Silver, 1981), physics (Chi, Feltovich, & Glaser, 1981; Simon & Simon, 1978; Larkin, 1983; Larkin, McDermot, Simon, & Simon, 1980), and computer programming (Adelson, 1984).

What accounts for this tendency of novices to focus on surface features of problems at the expense of hidden relational properties? Is it a function of a lack of knowledge, or rather the inability to extract and mentally represent these relational properties?

One possible explanation of novices' tendency to focus on surface features is that they merely have little knowledge of deep structural relations. However, while this possibility is capable of explaining failures to focus on deep relations by young and untrained participants, it fails to explain such failures by older participants when they are presented with problems drawn from fairly simple knowledge domains, such as elementary mathematics. The credibility of this lack of knowledge explanation is further undermined by findings that novices' representations in knowledge-rich domains are compatible with those in knowledge-lean domains, such as induction and analogy problems (Gentner, 1989; Gentner & Toupin, 1986; Kotovsky & Gentner, 1996; Holyoak & Koh, 1987). In these knowledge-lean domains, participants (especially younger ones) were more likely to focus on featural similarity (i.e., the color of compared entities) than on relational similarity (e.g., the symmetry of compared entities).

Another explanation is that even when novices know deep relations, they still fail to extract and mentally represent these relations because the relations are overshadowed by more salient surface features. This representation explanation contains two possibilities: novices may simply fail to encode deep relational properties, or they may initially encode both surface and deep relational features but then discard the latter in favor of more salient surface features.

There are several studies directly examining the knowledge and representational explanations for adult novices (Sloutsky & Yarlas, in press; Yarlas & Sloutsky, in press). In a set of experiments with mathematics experts (doctoral candidates in mathematics) and adult novices (college undergraduates and doctoral candidates in history), Yarlas and Sloutsky (in press) examined participants' representations of structural properties of arithmetic operations. In a forced-choice grouping task, participants were asked to group together arithmetic equations that could be matched either on the basis of a common mathematical principle (e.g., associativity) or on the basis of a common surface feature (e.g., equations that have the same number of addends). While math experts overwhelmingly relied on mathematical principles in their matches, novices relied mostly on surface features. However "unmasking" of principled features by eliminating competing surface features led to a substantial increase in novices' principle-based groupings. It was concluded, therefore, that adult novices knew the principles in question. Why then did they fail to focus on principles when these principles were “masked” by surface features? Did they fail to extract the principles and to encode them? Or were these principles overshadowed by more salient, directly perceptible surface features?

To answer these questions, Sloutsky and Yarlas (in press) conducted several experiments that used a recognition procedure. This procedure affords the creation of a set of foils such that patterns of recognition judgments point to which aspects of a problem have been encoded and stored in memory and which have been left out. While adult novices encoded both surface and principled features, as evidenced by high degrees of accuracy for both types of features in their recognition judgments, their responses were significantly slower when a foil shared a surface feature with a study item without sharing a principle. This delay may be indicative of a response competition between the deep and surface features, pointing to a relative difficulty for adult novices to inhibit the salient surface feature and correctly reject the foil. In short, even when adult novices knew the principled features and encoded these features, their attention to these features was affected by salient surface features. At the same time, mathematics experts did not exhibit this response competition.

Research on children’s mental representations of arithmetic problems have indicated that elementary school children have knowledge of some relational features and are capable of encoding and attending to these relational features when these features do not compete with surface features. For example, Canobi, Reeve, and Pattison (1998) examined childrens’ conceptual understanding and procedural use of the commutative and associative properties of addition. These researchers presented 6-8 year old children with pre-solved arithmetic equations (e.g., $a + b = c$), then asked children to solve a following problem that was either related to the first problem by an arithmetic principle of commutativity (i.e., b + a = ?) or was unrelated to the first problem (d + e = ?, where the correct solution was also "c"). These researchers found that children solved problems that were related by commutativity faster than problems that were unrelated (although there was no speed-up for problems related by associativity). Additionally, children self-reported using commutativity approximately half of the time when solving related problems, although they reported use of associativity with far less frequency. These results indicate that young children have knowledge of simple arithmetic principles such as commutativity, and are capable of encoding and representing these principles. Other studies have demonstrated that children of similar ages and even as young as 3 years have knowledge of commutativity, and are capable of using problem-solving strategies that incorporated the commutative principle (Baroody, Ginsburg, & Waxman, 1983; Cowan & Renton, 1996; Sophian, Harley, & Manos Martin, 1995).

Therefore, while it is reasonable to expect that even young children know at least some principles of arithmetic, several questions remain unanswered. What are the processes underlying problem representation in child and preadolescent novices? Which surfaces features are considered important by (i.e., are most salient for) child and preadolescent novices? And are there developmental differences in processes underlying mental representation?

The goal of the current research is to answer these questions by providing a detailed examination of knowledge and representation of principles of arithmetic in elementary school children and preadolescents. To achieve this goal, we designed three experiments that controlled for knowledge factors while manipulating representational factors. Knowledge was controlled by using simplified tasks and by selecting only those relational principles that should be well familiar to young mathematics students. Specifically, we selected the associative and commutative properties of arithmetic, because these principles are learned in 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 preadolescents, if not younger children. The associative property 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 commutative property states that the order of elements is irrelevant for addition and multiplication (e.g., a + b = b + a).

In this article, we describe three experiments that examine how children and adolescents represent arithmetic problems. In Experiment 1, we examined whether or not elementary school children and preadolescents know the principles in question, whereas in Experiments 2 and 3, we examined how participants who know the principles extract, encode, and represent these principles.

EXPERIMENT 1

The main goal of this experiment was to examine knowledge of principled relations of arithmetic (i.e., associativity and commutativity) by children and preadolescents. In the current experiment, principles either competed with surface features of arithmetic problems, or were presented alone without competing surface features. Children and preadolescents were asked to group arithmetic equations across three phases. In the first phase, these groupings could be based either on the commonality of surface features (e.g., numbers used, the number of constituent elements in the equations) or on the commonality of a principled relation. We refer to trials in this phase as “masked” trials, due to the fact that the principled relations are masked by the salient surface features. In the second phase, principles were "unmasked," such that surface features were not varied among the compared equations. Following this Unmasked phase, in the third phase the principled relations were "masked" again by reintroducing competing surface features. In this experiment, we were particularly interested in performance in the Unmasked phase compared to the two Masked phases.

Should participants group equations based on principled relations, it would indicate that they have knowledge of these relations. At the same time, if participants fail to group equations based on principled relations even on Unmasked trials, it would suggest that the participants do not know principles in question. Additionally, the experiment will afford examination of the saliency among particular types of surface features in arithmetic problems for each age group.

Participants

Two samples were selected for Experiment 1, each representing a different age group. The first group, which will be referred to as the “children” group, included 20 first- and second-graders (M = 7.26 years, SD = 0.59; 8 girls and 12 boys) who were selected from 3 mixed-grade classrooms at an elementary school located in a suburb of Columbus, Ohio. The second group, which will be referred to as the “preadolescents” group, included 16 sixth-graders (M = 12.10 years, SD = 0.38; 5 girls and 11 boys) who were selected from 3 classrooms at a middle school located in the same suburb of Columbus, Ohio. These participants were selected on the basis of returned parental permission forms. As described below, the experiment involved two separate sessions, with the second session conducted approximately four months after the first session for each sample. All participants participated in both sessions, with the exception of one boy from the “children” sample who participated in the first session but was unavailable for the second session.

Materials

The experiment includes three phases that took place over two sessions. All materials across all phases and sessions were identical for both age groups.

A forced-choice similarity paradigm was used for all three phases in this experiment. For each trial within each phase, 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 a mathematical "expert" would think was more 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. The numbers used in the arithmetic equations ranged from 1 to 14, and the operations used included addition, subtraction, and multiplication. The numerical solutions of all three equations in triads were controlled for by making these solutions either all equal or all different for each trial.

The first session included the Masked 1 phase, during which five features of arithmetic equations were used. Two of these features were considered principled features, such that they represented deep, relational principles of mathematical operations: the associativity and commutativity principles. The other three features were nonprincipled surface features that occur in arithmetic equations: (1) numbers (e.g., 7, 11); (2) sign (e.g., -, +); and (3) the number of constituent terms in an equation. On some trials, the Target equation shared a principled feature with one of the Test equations, and shared a surface feature with the other Test question. Thus, for these trials, principled features were "masked" by surface features of equations (see the top six rows in Table 1 for relevant examples). On other trials, the Target equation shared one type of surface feature with one of the Test equations, and shared another type of surface feature with the other Test equations (see the three bottom rows in Table 1 for relevant examples). All five (2 principled and 3 surface) features were pitted directly against each other, with the exception of the two principled features, yielding a total of nine feature comparisons. Each of the nine comparisons were presented for four times (one practice trial and three experimental trials) resulting in a total of 36 trials (9 practice trials and 27 experimental trials).

The second session occurred approximately four months after the first session. The second session included two phases: the Unmasked phase, and the Masked 2 phase. In the Unmasked phase trials, each of the two principled features were pitted against "control" problems, which did not share unique surface features with the Target. Thus, for these trials, both Test equations were similar to the Target on surface features, while one Test equation shared a principled feature with the Target equation. 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 = 1 + 5 + 11, and the control Test equation was 3 + 10 + 4 = 12 + 4 + 1. Principles were therefore "unmasked" during this phase, in that principles were not competing with surface features. There were a total of 8 trials during the Unmasked phase, with 4 trials for each of the two principled features.

Immediately following the Unmasked phase was the Masked 2 phase. The Masked 2 phase was identical to the Masked 1 phase, in that principled features were again "masked" by competing surface features. The Masked 2 phase consisted of the 27 experimental trials presented in the Masked 1 phase, with 3 exemplars of each of the 9 comparisons presented in Table 1. Thus, the second session consisted of 35 trials: 8 trials in the Unmasked phase and 27 trials in the Masked 2 phase.

Procedure

The procedure for both samples was identical for both sessions, with the exception of a minor change in the cover story that will be described below. For each session, participants were taken individually by a male experimenter into a small, quiet room within their school building.

For the first session, participants in the "children" group were presented with a drawing of a cartoon cat named "Zippy." They were told that Zippy was a smart cat and a good math student. They were further told that Zippy had been given a homework assignment by his teacher to match math problems on cards together based on which problems were more similar to each other, but that after completing the assignment, Zippy dropped the cards and they became unmatched. The child was then asked to help Zippy by matching the equations together the way that Zippy had. Participants in the preadolescent group were told a similar cover story, with the identity of the fictional confederate changed to a smart math student named "Chris". The gender of "Zippy" and "Chris" was never identified. The remaining procedure was equivalent for both samples.

To acquaint participants with the forced-choice similarity task, they were presented with a warm-up trial that included Gelman and Markman's (1986) blackbird-flamingo-bat pictures. The Target card depicted a blackbird that was perceptually similar to the bat and dissimilar from the flamingo. The experimenter asked the participant which of the Test items was more like the Target item, and why they made their choice. All participants were able to answer the posed questions.

Participants were then presented with the 36 Masked 1 phase trials, which took approximately 30 minutes. Trial order was randomized and the positioning of the test items in relation to the target (i.e., left or right) was counterbalanced across comparison type. In addition to recording which test problems were selected by the participants, their verbal explanations for their similarity judgments were also recorded by the experimenter.

The second session took place approximately 4 months after the first session for each group. The same procedure used for the first session was used for the second session (including the cover story, instructions, blackbird-flamingo-bat example, and conduction of trials), with the exception that there were now 35 trials (8 trials in the Unmasked phase followed by 27 trials in the Masked 2 phase). Because all participants had been acquainted with the task in session 1, and because "unmasking" was believed to make principles transparent, practice trials were not used in the second session.

Results and Discussion

In this section, we will focus on participants' choices across the three phases. Recall that one of the objectives 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 explanation-consistent choices. For example, when choosing a commutativity test item, an explanation such as “In both equations, they have the same numbers on both sides, just in different orders” was considered an explanation-consistent choice, while an explanation such as “Both equations have a ‘7’ in them” was not considered an explanation-consistent choice. The percentage of explanation-consistent choices for each principle across the principle-feature comparison trials is the dependent variable used in the forthcoming analyses.

The percentages of explanation-consistent principled choices across age groups and phases, collapsed across principle, are presented in Figure 1.

A 3 * 2 repeated-measures ANOVA was conducted for explanation-consistent choices for each principle, with phase as a repeated-measure and age as a between-participants factor. For the ANOVA involving explanation-consistent associativity choices, all effects were statistically significant, including the phase-by-age interaction (F (2, 66) = 7.17, MSE = .01, $p < .005$) and the main effects for phase (F (2, 66) = 7.72, MSE = .01, $p < .005$) and age (F (1, 33) = 7.35, MSE = .03, $p < .02$). Post-hoc t-tests with Bonferroni adjustments were used to more closely examine differences between age groups and among phases. Across the two age groups, independent-samples t-tests indicated no significant differences for explanation-consistent associativity choices between children (M = 0.00%, SD = 0.00%) and preadolescents (M = 2.08%, SD = 6.04%) for Masked 1 trials, t (34) = 1.55, $p > .5$, nor were differences found for explanation-consistent associativity choices between children (M = 1.17%, SD = 5.10%) and preadolescents (M = 6.94%, SD = 15.11%) for Masked 2 trials, t (33) = 1.57, $p > .3$. However, for Unmasked trials, a significant difference was found for explanation-consistent associativity choices between children (M = 0.00%, SD = 0.00%) and preadolescents (M = 20.31%, SD = 21.98%), t (33) = 3.04, $p < .02$. Paired-samples t-tests with Bonferroni adjustments were used to examine differences across phase within each age group. For children, there were no significant differences in explanation-consistent associativity choices across the three phases, all ts (18) < 1. For preadolescents, however, while there was not a significant difference between the Masked 1 and Masked 2 phases (t (15) = 1.45, $p > .5$), there was a marginally significant difference between the Masked 1 and Unmasked phases (t (15) = 2.65, $p = .054$) and a significant difference between the Masked 2 and Unmasked phases (t (15) = 2.74, $p < .05$).

To more closely inspect knowledge for associativity, in addition to examining the proportions of choices, we also examined the number of participants who made explanation-consistent associativity choices at least once in the 3 phases, as opposed to those who never made such a choice. This measure provides a good estimate of participants' knowledge of principles in question because without this knowledge it would be impossible to make even a single explanation-consistent principled choice. In the children group, only 1 participant (of 19) made at least one explanation-consistent associativity choice, while in the preadolescent group, 7 participants (of 16) made at least one explanation-consistent associativity choice. The analysis of this data indicated that preadolescents were significantly more likely to make at least one explanation-consistent associativity choice than were children, χ^2 (1, N = 35) = 7.30, $p < .01$.

Taken together, these data indicate that the younger group did not make explanation-consistent associativity choices regardless of whether the principle was masked or unmasked by surface features. Preadolescents also generally failed to make explanation-consistent associativity choices when the principle was masked by surface features; however, when associativity was unmasked, preadolescents were significantly more likely to base their choices on this principle. Therefore, it seems reasonable to infer that the children lacked knowledge of the associativity principle, while at least some preadolescents appear to have knowledge of this principle.

The same analyses were conducted for explanation-consistent commutativity choices. For this 3 * 2 ANOVA, all effects were statistically significant, including the phase-by-age interaction (F (2, 66) = 17.71, MSE = .03, $p < .001$) and the main effects for phase (F (2, 66) = 53.85, MSE = .03, $p < .001$) and age (F (1, 33) = 20.65, MSE = .08, $p < .001$). Again, post-hoc t-tests with Bonferroni adjustments were used to more closely examine differences between age groups and among phases. Across the two age groups, no significant difference was found for explanation-consistent commutativity choices between children (M = 0.00%, SD = 0.00%) and preadolescents (M = 9.03%, SD = 17.32%) for Masked 1 trials, t (34) = 2.34, $p > .05$. However, a significant difference was found for explanation-consistent commutativity choices between children (M = 15.79%, SD = 30.29%) and preadolescents (M = 67.19%, SD = 29.89%) for Unmasked trials, t (33) = 5.03, $p > .001$, and a marginally significant difference was found between children (M = 1.75%, SD = 7.65%) and preadolescents (M = 15.28%, SD = 22.18%) for Masked 2 trials, t (33) = 2.49, $p = .053$. Paired-samples t-tests with Bonferroni adjustments were again used to examine differences across phase with each age group. For children, there were no significant differences in explanation-consistent commutativity choices among the three phases, all ts (18) < 2.4, all ps > .05. For preadolescents, again there was not a significant difference between the Masked 1 and Masked 2 phases (t (15) = 1.38, $p > .5$), but there were significant differences between both Masked (1 and 2) and Unmasked phases, both ts (15) > 8, ps < .001.

As with associativity, we also examined the number of participants who made explanation-consistent commutativity choices at least once, as opposed to those who never made such a choice. In the children group, 5 participants (of 19) made at least one explanation-consistent commutativity choice, while in the preadolescent group, 15 participants (of 16) made at least one explanation-consistent commutativity choice. The analysis of these data indicated that preadolescents were significantly more likely to make at least one explanation-consistent commutativity choice than were children, χ^2 (1, N = 35) = 16.13, $p < .001$.

These patterns of results are similar to those found with the associativity principle, indicating that the children also generally lacked knowledge of commutativity, but that preadolescents do have knowledge of this principle and readily attend to it when it was not competing for attention with salient surface features. While the younger group did not generally make explanation-consistent commutativity choices regardless of whether the principle was masked or unmasked by surface features, preadolescents were likely to make a principle-based choice when commutativity was unmasked. The lack of focussing on commutativity by children is surprising given previous research indicating that children of this age and younger are capable of recognizing commutativity in arithmetic problems (Baroody et al., 1983; Canobi et al., 1998; Cowan & Renton, 1996; Sophian, et al., 1995). Possible explanations for the disparity between these two sets of findings will be addressed below in the General Discussion.

Given that participants in both age groups overwhelmingly made choices based upon surface features in the two Masked phases, we analyzed differences in the frequency of explanation-consistent choices among the three surface features for each age group. An omnibus Chi-square test yielded significant differences among proportions of explanation-consistent choices across the three surface features for children, χ^2 (2, N = 1404) = 507.25, $p < .001$. An analysis of the standardized residuals indicated that children were more likely to choose a Test equation that shared numbers with the Target equation (in 87.6% of all trials in which a Test equation shared numbers with the Target equation) than they were to make a choice based upon shared number of constituent elements (24.8%) or shared sign (23.5%), both zs > 20.1, ps < .001. At the same time, children based their choices on shared number of constituent terms and sign with similar frequency, $z < 1$. The analysis of preadolescents' choices indicated a different pattern of preferences among surface features than had been found for children. The omnibus Chi-square for preadolescents yielded significant differences among proportions of explanation-consistent choices across the three surface features, χ^2 (2, N = 1152) = 56.76, $p < .001$. Unlike with children, the analysis of standardized residuals indicated that preadolescents were more likely to choose a Test equation that shared number of constituent terms with the Target equation (78.9% of trials in which numbers were used as a surface feature) than they were to make a choice based upon shared numbers (54.4%) or sign (58.9%), both zs > 4.9, ps < .001, while their choices based on numbers and sign were made with similar frequency, z = 1.08, p > .05.

To directly examine developmental differences in the salience of surface features, proportions of participants' choices were cross-tabulated and subjected to a Chi-square analyses. This analysis indicated that children were more likely to make explanation-consistent choices based on shared numbers than were preadolescents (χ^2 (1, N = 852) = 116.87, p < .001). At the same time, preadolescents were more likely than children to make explanation-consistent choices based on the shared number of constituent terms (χ^2 (1, N = 852) = 247.19, p < .001) and the shared sign (χ^2 (1, N = 852) = 110.36, p < .001). Thus, while children were most likely to attend to individual numbers within equations, preadolescents were most likely to attend to the number of constituent terms in equations, and more likely to attend to sign than were children.

Results of Experiment 1 point to several important regularities. First, very few participants in the children group knew of principles in question, as evidenced by their lack of explanation-consistent principled choices even when the principles were "unmasked". At the same time, preadolescents did demonstrate knowledge of these principles, as evidenced by their increased attention to principles when the principles were unmasked. Second, when principles were "masked" by salient surface features, even preadolescents failed to attend to the principles. Indeed, even immediately following the Unmasked phase, during which preadolescents made explanation-consistent principled choices on almost half of all trials, when surface features were reintroduced in the Masked 2 phase, the number of preadolescents' explanation-consistent principled choices reverted back to the level of Masked 1 trials. Finally, analysis of participants' surface feature choices indicate that children greatly relied on shared numbers in their similarity judgments, while preadolescents' choices were more often driven by shared number of constituent terms in equations. This tendency may be indicative of developmental differences in units of arithmetic thinking: while for elementary school children a single number is such a unit, for preadolescents it is the whole equation.

Having established that preadolescents knew principles in question, we deemed it necessary to examine why they failed to focus on these principles in their judgments during the two Masked phases of the forced choice task. First, it is possible that preadolescents did not notice (and thus did not encode) these principles when the principles were pitted against surface features. Second, it is possible that preadolescents' noticed and encoded the principles, but discarded them in favor of surface features. Experiment 2 was designed to distinguish between these possibilities.

EXPERIMENT 2

The goal of the current experiment was to distinguish between two possible mechanisms underlying preadolescents' failure to represent principles of commutativity and associativity. Recall that one possibility was that preadolescents fail to encode principled properties, whereas another possibility was that preadolescents do encode principled properties, but these properties lose attentional competition to more salient directly perceptible features.

Experiment 2 uses a recognition procedure to examine these possibilities. Note that patterns of recognition responses are indicative of which features of equations are encoded and committed to memory and which features are left out. In the study phase, participants, all of whom were preadolescents, were given a set of arithmetic equations similar to those used as Target items in Experiment 1. These equations all involved a principled property, either commutativity or associativity. In addition, these equations all used consistent levels of two surface features: 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' equations, four types of 'new' equations were presented as foils. Half of these foils, which we refer to as 'Feature +' foils, maintained the same levels of 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 the four combinations of foils: Feature + /Principle + (F+/P+), Feature + /Principle - (F+/P-), Feature - /Principle + (F-/P+), and Feature - /Principle - (F-/P-).

A crucial assumption of this paradigm is that participants will not be able to encode and store the specific equations presented in the study phase, due to the fact that these equations are not very highly differentiable from each another. Rather, it is expected that participants will have a categorical memory for these items based on their surface and/or principled features, such that they will reject items that violate these categories and accept those that do not. The accuracy of this assumption would be evidenced by findings that participants are unable to differentiate among F+/P+ foils and Old targets, and thus would inaccurately judge F+/P+ foils to be "Old."

Patterns of accuracy and latencies afford a direct examination of the processes that underlie preadolescents' problem representations. However, while the range of possible processing models is rather large, we consider only those models that are compatible with prior findings. For example, it seems highly implausible that preadolescents encode only deep relational features. Therefore, all considered models assume that participants encode surface features of arithmetic equations.

Those models that we deemed most plausible are presented graphically in Figure 2. If preadolescents encode only surface features (Model 1), they should respond "Old" when surface features are present and they should respond "New" when surface features are absent, regardless of the presence or absence of principles. On the other hand, if preadolescents encode both surface and deep features (Models 2 and 3), they should respond “Old” when both surface and principled features are present (i.e., F+/P+ foils and Old targets) and respond “New” when surface and/or principled features are absent (i.e., F-/P+, F-/P-, and F+/P- foils).

Several previous studies indicated that adult novices were found to exhibit a surface feature-principle response competition, such that response latencies were longest for foils in which surface features were present but principles were absent (Sloutsky & Yarlas, in press). Therefore, if preadolescents experience the same surface feature-principle response competition as did adult novices (as represented in Model 3), then it should require more time for them to correctly respond to the foil where surface features are present and principled features are absent (i.e., F+/P- foils) than for their responses to foils where surface and principled features are present.

Assuming that preadolescents access features in a serial manner, then each additional step in processing leads to increases in latencies. We can therefore derive testable predictions from each of the models presented in Figure 2; these predictions are presented in Table 2. Note that these predictions are qualitative, in that they do not specify particular levels of accuracy or latency across foils, but rather to patterns of recognition responses and directions among latencies for foils.

From Table 2, it can be seen that the F+/P- foil is the critical foil in differentiating among the three proposed models. If participants encode only surface features (Model 1), then they should judge F+/P- foils to be "Old." However, if participants encode both surface and principled features (Models 2 and 3), they should judge F+/P- foils to be "New." Additionally, patterns of latencies will afford differentiation between the three models. If preadolescents encode only surface features (Model 1), then there should be no differences in latencies among foils. However, if preadolescents encode and access both surface and principled features, with surface features being accessed first (Model 2), F- foils should generate faster responses than F+ foils, and therefore the following pattern of latencies should be observed: Old targets = F+/P+ = F+/P- > F-/P+ = F-/P-. Further, if preadolescents encode and access both surface and principled features, and experience response competition (Model 3), then latencies should adhere to the following pattern: F+/P- > Old targets = F+/P+ > F-/P+ = F-/P-. The goal of Experiment 2 is thus to test processing predictions presented in Table 2 in order to determine which of the proposed processing models for preadolescents is most accurate in predicting their encoding and access of surface and principled features.

Method

Participants

This study included 19 preadolescents, who were sixth-graders selected from 3 classrooms at a middle school located in the same suburb of Columbus, Ohio. Participants had a mean age of 12.36 years (SD = 0.36) and included 6 girls and 13 boys. These participants were selected on the basis of returned parental permission forms.

<u>Materials and Procedure</u>

All participants were run individually in a quiet room in their school building. Stimuli were presented on a laptop 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 30 arithmetic equations, which they had been instructed to memorize. All 30 equations incorporated the mathematical operation of addition, included 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 black type on a white background for ten seconds, with a two-second interval between each equation, during which only the white background was seen. The order of presented equations was randomized.

A distraction phase followed the study phase for the purpose of clearing participants' short-term memory. For the distractor task, participants were presented with 90 letters, for which they had been instructed to indicate whether the letter was a vowel (by pressing the "Z" key on the keyboard) or a consonant (by pressing the "M" key). Each letter was centered and presented in black type on a white background. This phase took approximately three minutes.

Following the distraction phase was the recognition phase. Participants were instructed that they would be presented with a number of arithmetic equations, some that had been presented to them earlier and some that had not been presented earlier. They were further instructed to indicate whether each equation had been presented earlier or not by pressing the "Z" key on the keyboard if the item was presented earlier, or the "M" key if it had not been presented earlier.

There were a total of 60 equations presented in the recognition phase. Each equation was centered and presented in black type on a white background. The order of equations presented in this phase was randomized. These equations fell into five categories, with 12 exemplars for each category. The first category contained “Old targets”, which consisted of equations that had been randomly selected from those that had been presented in the study phase. The remaining four categories were foils, in that they contained new problems that had not been presented in the study phase. One type of foil consisted of "Feature + /Principle +" (F+/P+) equations that used surface features from the same categories as the original equations (i.e., numbers between 1 and 9, and either 5 or 6 numbers), and used either the commutativity or associativity principle (e.g., 6 + 1 + 4 = 1 + 4 + 6 and 9 + 6 = 5 + 4 + 6). A second type of foil consisted of "Feature + /Principle -" (F+/P-) equations that used surface features from the same categories as the original equations but did not use either the commutativity or associativity principle (e.g. 3 + 1 + 9 = 2 + 5 + 6 and 4 + 9 = 3 + 8 + 2). A third type of foil consisted of "Feature - /Principle +” (F-/P+) equations that used surface features that violated the categories used in the original equations (i.e., numbers greater than 9, and either 4 or 7 numbers), but still incorporated either the commutativity or associativity principle (e.g., 7 + 12 = 12 + 7 and 5 + 1 + 15 = 3 + 2 + 1 + 15). A fourth type of foil consisted of “Feature - /Principle –“ (F-/P-) equations that used surface features that violated the categories used in the original equations and did not incorporate either the commutativity or associativity principle (e.g., 8 + 3 + 12 = 5 + 1 + 4 + 13 and 6 + 11 = 2 + 15).

Results and Discussion

In this section, 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, latencies for both correct and incorrect responses for these foils were used in the analyses.

Participants' recognition judgments were highly accurate for most of the foils, as evidenced by significantly greater than chance acceptance of Old targets (M = 75.08%, SD = 20.59%) and rejection of F-/P+ (M = 96.77%, SD = 6.82%) and F-/P- (M = 93.78%, SD = 9.17%) foils (all ts (18) > 5.3, ps < .001) and slightly, though not significantly above chance rejection of F+/P- foils (M = 62.53%, SD = 29.43%, t (18) = 1.9, p = .08). They mostly false alarmed, however, on F+/P+ foils (M = 27.20%, SD = 18.38%), as indicated by less than chance level accuracy, t (18) = - 5.4, p < .001. This latter finding supports the assumption that participants relied upon categorical memory, rather than memory for specific items, in their recognition: 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. These results indicate that these participants took the task seriously and were providing rather systematic responses.

Effect sizes of the two types of features on responses (i.e., Cohen d's) were computed to determine the degree to which surface features and principled features contributed to participants' recognition judgments. This analysis yielded a very large effect size due to surface features (d = 1.50) and a medium effect size due to principled features (d = .48). Thus, while surface features had a larger effect on participants' recognition judgments, principled features had a substantial effect on their responses as well.

Percentages of "Old" responses for participants are presented in the left-hand side panel of Figure 3. A one-way repeated measures ANOVA points to significant differences among foils for participants (F (4, 72) = 71.39, MSE = .03, p < .001). Paired-samples post-hoc t-tests with Bonferroni adjustments indicated the following order of "Old" responses across foils: F-/P+ = F-

/P- < F+/P- < F+/P+ = Old targets, all ts (18) > 4.47, all ps < .002 for differences. Thus, participants gave the fewest "Old" responses for foils in which surface features were absent, and were less likely to give "Old" responses when either surface or principled features were absent than when both surface and principled features were present. These findings are inconsistent with predictions of Model 1 (see Table 2). Rather, the patterns of results indicate that participants encoded both surface and principled features of arithmetic problems and used both of these features in their recognition judgments, as represented in Models 2 and 3 (Figure 2).

The right-hand side panel of Figure 3 presents participants' latencies to correct responses across F-/P- (M = 1856 ms, SD = 977 ms), F-/P+ (M = 1602 ms, SD = 628 ms), F+/P- (M = 3762 ms, SD = 1821 ms) foils, and Old targets (M = 2801 ms, SD = 1176 ms), and all responses for F+/P+ foils (M = 3161 ms, SD = 1338 ms). These measures were also subjected to a one-way repeated measures ANOVA, which indicated significant differences among the foils, F (4, 68) = 21.20, p < .001. Paired-samples post-hoc t-tests with Bonferroni adjustments indicated that F- foils were faster than all other foils, all ts < -5.2, all ps < .005. A planned comparison was used to determine whether the feature-principle response competition that had been found in adult novices was evident in preadolescents (Model 3, Figure 2). Indeed, latencies for F+/P- foils were significantly greater than for Old targets, t (17) = 2.53, p < .05.

The pattern of latencies (i.e., F+ foils being slower than F- foils, and F+/P- foils eliciting the slowest responses) in conjunction with patterns of responses support Model 3 (Figure 1) as the most accurate of the proposed models in describing feature processing by preadolescents. In particular, the observed pattern of responses indicate that preadolescents encode both principled and surface features, whereas the pattern of latencies indicate that, in the course of recognition, they access surface features prior to accessing principled features. Finally, the delayed response to F+/P- foils may be indicative of a tendency to answer "Old" to a F+ foil and a suppression of this tendency, or a feature-principle response competition.

However, Experiment 2 leaves an important issue unresolved. While this experiment indicates that preadolescents encode both surface and principled features, and elucidates the order of feature access, it does not reveal the order of feature encoding. The order of access might be indicative of the order of encoding, but it provides only suggestive evidence. Thus, although results of Experiment 2 suggest that preadolescents encode surface features first, we deemed it necessary to generate more direct evidence. One approach that can be used to investigate this question is to radically reduce exposure time to equations in the study phase. If under these conditions preadolescents will encode only one kind of feature, this would represent strong evidence that encoding is a serial process and that this feature is encoded first. For example, if preadolescents encode surface features prior to principled features, then a sufficiently shorter encoding time would result in only surface features being encoded. This manner of feature processing, which is depicted in Model 1 (see Figure 2), would be manifested by F+ foils being judged as "Old" and F- foils being judged as "New." To provide a such an examination of the issue of encoding order, we conducted Experiment 3.

EXPERIMENT 3

Method

Participants

This study included sixteen preadolescents, who were sixth-graders selected from 3 classrooms at a middle school located in the same suburb of Columbus, Ohio. Participants had a mean age of 12.34 years (SD = 0.30) and included 9 girls and 7 boys. These participants were selected on the basis of returned parental permission forms.

Materials and Procedure

The materials and procedures for this experiment were identical to those used in Experiment 2, with one critical exception. In the current experiment, each of the 30 equations in the study phase was presented on the screen for 1.5 seconds, as opposed to 10 seconds as in Experiment 2.

Results and Discussion

As for our discussion of results for Experiment 2, 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. Again, for all foils except F+/P+, we compared latencies for correct responses only.

Participants again made highly accurate recognition judgments for most of the foils, exhibiting significantly greater than chance acceptance of Old targets (M = 74.48%, SD = 12.35%) and rejection of F-/P+ (M = 88.45%, SD = 15.22%) and F-/P- (M = 90.62%, SD = 10.03%) foils, all ts (15) > 7.9, ps <. 001. On the other hand, accuracy for F+/P- foils were slightly, though not significantly below chance (M = 38.54%, SD = 25.44%, t (15) = -1.8, p = .09), and again preadolescents mostly false alarmed on F+/P+ foils (M = 36.65%, SD = 13.88%), as indicated by less than chance level accuracy, t (15) = - 3.9, p < .005. An independent-samples t-test was conducted to determine whether accuracy for F+/P- foils dropped significantly from levels observed in Experiment 2 due to the reduction in encoding time. This analysis yielded that accuracy for this foil dropped significantly in the current experiment as compared to accuracy in Experiment 2, t (33) = 2.55, p < .02.

As in the previous experiment, effect sizes of the two types of features on responses (i.e., Cohen d's) were computed to determine the degree to which surface features and principled features contributed to participants' recognition judgments. This analysis again yielded a very large effect size due to surface features ($d = 1.67$), but in contrast to that in Experiment 1, a very small effect size due to principled features ($d = .07$). Thus, given the shorter encoding time, surface features still had a large effect on participants' recognition judgments, while principled features had a negligible effect on their responses.

Percentages of "Old" responses for participants are presented in the left-hand side panel of Figure 4. A one-way repeated measures ANOVA points to significant differences among foils for participants, F (4, 60) = 59.38, $MSE = .03$, $p < .001$. Paired-samples post-hoc t-tests with Bonferroni adjustments indicated the following order of "Old" responses across foils: F-/P+ = F-/P- < F+/P- = F+/P+ = Old targets, all ts (15) > 7.04, all ps < .001 for differences. Thus, in contrast to Experiment 2, where participants gave fewer "Old" responses to F+/P- foils than for F+/P+ foils and Old targets, in the current experiment there were not significant differences in the percentages of "Old" responses for these three foils. Rather, participants were likely to judge items as "New" when surface features were absent and as "Old" when surface features were present, regardless of the presence or absence of principled features. This pattern of responses is consistent with predictions of Model 1 (see Table 2).

The right-hand side panel of Figure 4 presents participants' latencies to correct responses across F-/P- (M = 1694 ms, SD = 791 ms), F-/P+ (M = 1499 ms, SD = 517 ms), F+/P- (M = 2291 ms, SD = 1121 ms) foils, and Old targets (M = 1984 ms, SD = 748 ms), and all responses for F+/P+ foils (M = 2095 ms, SD = 751 ms). These measures were also subjected to a one-way repeated measures ANOVA. This analysis indicates significant differences among the foils, F (4, 56) = 7.91, $p < .001$. Paired-samples post-hoc t-tests with Bonferroni adjustments indicated that F-/P+ foils were faster than F+/P-, F+/P+, and Old targets (all ts < -3.3, ps < .05), with no other differences being statistically significant. Thus, there was much less variation in latencies among foils than in Experiment 2. This general lack of differences in latencies is best predicted by Model 1 (see Table 2).

Taken together, the results of Experiment 3 indicate that with a greatly reduced encoding time, preadolescents generally encoded only surface features, and not principled features, of the equations, as represented in Model 1 (Figure 2). Their recognition judgments were driven almost solely by the presence or absence of surface features, with principled features contributing little unique variance to their choices. Additionally, there was much less distinction among latencies for foils in the current experiment than in Experiment 2. This lack of difference among latencies would be predicted from Model 1, in which all decisions are based on a single step, though not from Models 2 and 3, in which decisions regarding F+ foils and Old targets requires a two-step decision.

General Discussion

Results of the three reported experiments are as follows. First, both children and preadolescents focus on surface features rather than deep relational features when both are present in arithmetic equations, with children focussing mostly on particular numbers and preadolescents focussing most often on number of constituent terms, and focussing more often on sign than did children. Second, preadolescents were more likely than children to demonstrate knowledge of the principles in question. Third, when encoding time was sufficient, preadolescents encoded both surface and principled features of arithmetic equations, with surface features being encoded and accessed prior to principled features. And fourth, patterns of preadolescents' recognition responses point to a response competition between surface features and deep principles.

The first two findings point to important differences in arithmetic knowledge between children and preadolescents. Not only do preadolescents know more than children, they also focus on more mathematically important surface features. Indeed, while particular number and the number of constituent elements in the equation have little mathematical importance, the sign has some mathematical importance (e.g., addition and multiplication have different properties, such as multiplication being distributive and addition being non-distributive). In addition, these findings also indicate that for children, a single number is the most salient unit of arithmetic (they consider only properties of isolated numbers, such as number magnitude), whereas for preadolescents the most salient unit is the whole equation (these participants consider properties of equations, rather than properties of isolated elements).

While preadolescents did generally exhibit knowledge of associativity and commutativity, the performance of children indicated a lack of knowledge of either of these principles. These findings diverge from earlier reported results, which indicated that children of this age group knew the commutative property (Barody, et al., 1983; Canobe, et al., 1998; Cowan & Renton, 1996). One possible reason for such a divergence is different task demands. For example, in the study by Canobe, et al. (1998), principled properties did not compete with surface features, whereas in present study they did. In addition, in Canobe, et al. (1998) children had to merely notice the presence of commutativity, whereas in the present task participants not only had to notice commutativity, but also to explicitly verbalize the presence of this principle. It is likely that noticing the presence of the commutative principle in a particular problem is easier than to explicitly refer to this principle because commutativity is a perceptually-detectable relation. Therefore, it seems that the currently reported task used more stringent criteria for knowledge of relational principles. One way of resolving this issue in future research is to expand the recognition paradigm used in Experiments 2 and 3 to elementary school children.

The third and fourth findings point to important regularities in preadolescents' processing of arithmetic problems. In particular, preadolescents encode both surface and principled features of arithmetic problems, and they encode both types of features in a serial manner, with surface features encoded first. In addition, preadolescents also access both types of features in a serial manner with surface features accessed first. Finally, preadolescents exhibited a feature-principle response competition, thus performing consistent with predictions of Model 3 (Figure 1) when given a sufficient amount of encoding time. Note that the critical component distinguishing Model 3 from Model 2 is the presence of response competition between salient surface features and less salient principled features. In the course of response competition, which is evidenced by increased response times to F+/P- foils, participants suppress an attractive "Old" response to the F+/P- foil and reject this foil. However, they may not be able to suppress salient features under more demanding task conditions. In particular, it is possible that in more resource demanding tasks, such as the task presented in Experiment 1, deep relational features in novices may lose competition to salient surface features. This loss may have manifested itself in the tendency to focus on surface features while ignoring deep relational features. Recall that this tendency was prominently present in Experiment 1, as well as in different tasks and knowledge domains (Chase & Simon, 1973; Chi, et al., 1981; Gentner & Toupin, 1986; Kotovsky & Gentner, 1996; Larkin, 1983; Simon & Simon, 1978). The issues of how task demands and response competition affect attention to relational principles may require additional research using demanding tasks, such as problem solving, and problems where surface features would be directly pitted against deep principled properties.

Results of the three experiments elucidate some aspects of mathematical competence. In particular, challenges in acquisition of mathematics may stem from the fact that important mathematical principles represent hidden relations that are often overshadowed by salient surface features. Because both encoding and access of features are serial processes with surface features accessed prior to relational features, it is more likely that relational (and not surface) features would be overlooked. Furthermore, the shorter the time of exposure, the higher the probability that relational features would be ignored. It also seems likely that the probability of overlooking of a relational feature is a function of salience and the number of surface features competing with the relational feature. Therefore, it could be predicted that the more salient the surface features, the more difficult it is to extract and represent important mathematical relations. For example, if asked which of the two sequences exhibits faster growth, (1) 1/32, 1/16, 1/8, 1/4 or (2) 100, 400, 700, 1000, participants should answer that the second sequence grows faster. The reason for such response, according to present research, is that both the number magnitude and the size of each step are much larger in the second sequence than in the first one. As a result, these salient surface features may overshadow the hidden relational feature, such as the rate of change. Therefore, one way of attracting attention to hidden relational features of mathematics is by reducing the salience of surface features (e.g., by using letter notation instead of numbers). Of course, participants would focus on hidden relational features, only if they have knowledge of these features.

We also believe that the observed feature-relation response competition may not be limited to the domain of mathematics, and that it is a rather domain-general phenomenon. In particular, there is preliminary evidence that similar response competition exists in such diverse domains as propositional logic problems (Sloutsky & Yarlas, 2000) and metaphors (Sloutsky & Hasson, 2000). If the observed competition between surface and relational features is a general phenomenon, then it is reasonable to expect that it should affect performance across multiple tasks and domains. For example, the well documented tendency to ignore relational similarity in favor of featural similarity in analogical reasoning (Gentner, 1989; Gentner & Toupin, 1986; Kotovsky & Gentner, 1996) may also stem from feature-relation response competition, similar to the one reported in the present research. Of course, additional research is needed to test these contentions.

While several issues touched upon in this discussion require further research, the present studies allow us to reach the following conclusions:

(1) When surface and relational features compete for attention, children and preadolescents focus on surface features of arithmetic equations, while ignoring deep principled features, such as commutativity and associativity.

(2) Under these task conditions, preadolescents, but not young children, exhibited knowledge of the principled features when competing surface features were eliminated.

(3) Children and preadolescents focus on different surface features of arithmetic equations, with preadolescents most often focussing on properties of equations, and children most often focussing on properties of individual elements of equations.

(4) Preadolescents encode and access both surface and principled features of arithmetic problems. These features are encoded and accessed in a serial manner, with surface features encoded and accessed prior to principled features.

(5) Preadolescents exhibit response competition between principled and surface features of arithmetic problems. This response competition may make the task of focussing on deep principled features more difficult.

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.

Baroody, A. J., Ginsburg, H. P., & Waxman, B. (1983). Children's use of mathematical structure. Journal for Research in Mathematics Education, 14, 156-168.

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, (pp. 221-246). 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.

Canobi, K. H., Reeve, R. A., & Pattison, P. E. (1998). The role of conceptual understanding in children's addition problem solving. Developmental Psychology, 34, 882-891.

Case, R., & Okamoto, Y. (1996). The role of central conceptual structures in the development of children's thought. Monographs of the Society for Research in Child Development, 61(1-2).

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.

Cowan, R., & Renton, M. (1996). Do they know what they are doing? Children's use of economic addition strategies and knowledge of commutativity. Educational Psychology, 16, 407-420.

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

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 (pp. 29-74). 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 (pp. 171-189). Hillsdale, NJ: Erlabaum.

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, (pp. 199-241). New York: Cambridge University Press.

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

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), Conceptual and procedural knowledge: The case of mathematics (pp. 1-28). Hillsdale, NJ: Erlbaum.

Hinsley, D. A., Hayes, J. R. & Simon, H. A. (1977). From words to equations: Meaning and representation in algebra word problems. In P. A. Carpenter & M. A. Just (Eds.), Cognitive processes in comprehension (pp. 89-106). Hillsdale, NJ: Erlbaum.

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

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, (pp. 75-98). 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.

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

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.

Resnick, L., & Omanson, S. F. (1987). Learning to understand arithmetic. In R. Glaser (Ed.), Instructional psychology (Volume 3) (pp. 41-96). Hillsdale, NJ: Erlbaum.

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

Ross, B. H., & Kilbane, M. C. (1997). Effects of principle explanation and superficial similarity on analogical mapping in problem solving. Journal of Experimental Psychology: Learning, Memory, & Cognition, 23, 427-440.

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.

Siegler, R. (1976). Three aspects of cognitive development. Cognitive Psychology, 8, 481-520.

Siegler, R. (1991). In young children's counting, procedures precede principles. Educational Psychology Review, 3, 127-135.

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? (pp. 325-348). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sloutsky, V. M., & Hasson, U. (2000). Feature-relation response competition in metaphor comprehension. Manuscript in preparation.

Sloutsky, V. M., & Yarlas, A. S. (in press). Problem representation in experts and novices: Part 2. Underlying processing mechanisms. Manuscript accepted to the Proceedings of the XXII Annual Conference of the Cognitive Science Society.

Sloutsky, V. M., & Yarlas, A. S. (2000). Expert and novice differences in representational processing of logical arguments. Manuscript in preparation.

Sophian, C., Harley, H., & Manos Martin, C. S. (1995). Relational and representational aspects of early number development. Cognition & Instruction, 13, 253-268.

VanLehn, K. (1996). Cognitive skill acquisition. Annual Review of Psychology, 47, 513-539.

Yarlas, A. S., & Sloutsky, V. M. (in press). Problem representation in experts and novices: Part 1. Differences in the content of representation. Manuscript accepted to the Proceedings of the XXII Annual Conference of the Cognitive Science Society.

Author Note

This research was supported in part by a grant from James S. McDonnell Foundation to the second author. Correspondence concerning the manuscript should be addressed to Vladimir M. Sloutsky (sloutsky.1@osu.edu) or Aaron S. Yarlas (yarlas.1@osu.edu), 21 Page Hall, 1810 College Avenue, The Ohio State University, Columbus, OH 43210.

Table 1.

Examples of equations used for the nine feature comparisons in Masked phases of Experiment 1.

Features shared by Test 1 vs. Test 2 with the Target	Examples		
	Test 1	Test 2	Target
Number vs. Associativity	6 + 3 + 4 = 8 + 1 + 2 + 2	10 + 1 + 2 = 5 + 5 + 1 + 2	6 + 3 + 4 = 2 + 4 + 3 + 4
Sign vs. Associativity	6 + 7 = 9 + 3 + 1	9 – 3 = 1 + 8 – 3	5 + 4 = 3 + 2 + 4
Number of Constituent Terms vs. Associativity	6 + 2 = 3 + 4 + 1	7 + 1 + 8 = 7 + 1 + 4 + 4	9 + 4 = 9 + 3 + 1
Number vs. Commutativity	6 + 3 + 8 = 3 + 4 + 10	5 + 2 + 7 = 2 + 7 + 5	6 + 3 + 4 = 3 + 4 + 6
Sign vs. Commutativity	5 + 8 = 8 + 5	4 X 2 = 2 X 4	6 + 4 = 7 + 3
Number of Constituent Terms vs. Commutativity	6 + 2 + 10 = 7 + 8 + 1	4 + 11 = 11 + 4	5 + 7 + 3 = 3 + 7 + 5
Number vs. Sign	7 – 2 = 6 – 1	5 + 8 = 9 + 4	7 + 2 = 6 + 3
Number vs. Number of Constituent Terms	7 + 5 + 1 = 8 + 4 + 1	9 + 2 = 6 + 5	7 + 5 = 8 + 4
Sign vs. Number of Constituent Terms	2 + 5 + 4 = 3 + 7 + 1	10 – 1 = 11 – 2	7 + 6 = 8 + 5

Table 2

Patterns of responses and latencies predicted by alternative models for preadolescents

Components of Processing	Foil Types and Patterns of Responses				
	Old targets	F+/P+	F+/P-	F-/P+	F-/P-
Encoding	Response Patterns				
Surface Features only (Model 1)	Old	Old	Old	New	New
Surface Features and Principles (Models 2 and 3)	Old	Old	New	New	New
Access	Latencies				
Surface Features only (Model 1)	Fast	Fast	Fast	Fast	Fast
Surface Features and Principles (Model 2)	Slow	Slow	Slow	Fast	Fast
Surface Features and Principles with Response-Competition (Model 3)	Slow	Slow	Slowest	Fast	Fast

Figure Captions

Figure 1. Percentages of explanation-consistent principled choices across age groups and phases in Experiment 1.

Figure 2. Three proposed putative processing models underlying preadolescents' encoding and access of surface features and principled features in the recognition procedure.

Figure 3. Proportion of preadolescents' "Old" responses and latencies (in milliseconds) across foil types in the recognition phase of Experiment 2. Error bars represent Standard Errors of the Mean. F+/P- = Feature +/Principle -, F+/P+ = Feature +/Principle +, F-/P- = Feature -/Principle -, F-/P+ = Feature - /Principle +, and Old = Old targets.

Figure 4. Proportion of preadolescents' "Old" responses and latencies (in milliseconds) across foil types in the recognition phase of Experiment 3. Error bars represent Standard Errors of the Mean. F+/P- = Feature +/Principle -, F+/P+ = Feature +/Principle +, F-/P- = Feature -/Principle -, F-/P+ = Feature - /Principle +, and Old = Old targets.

Figure 1

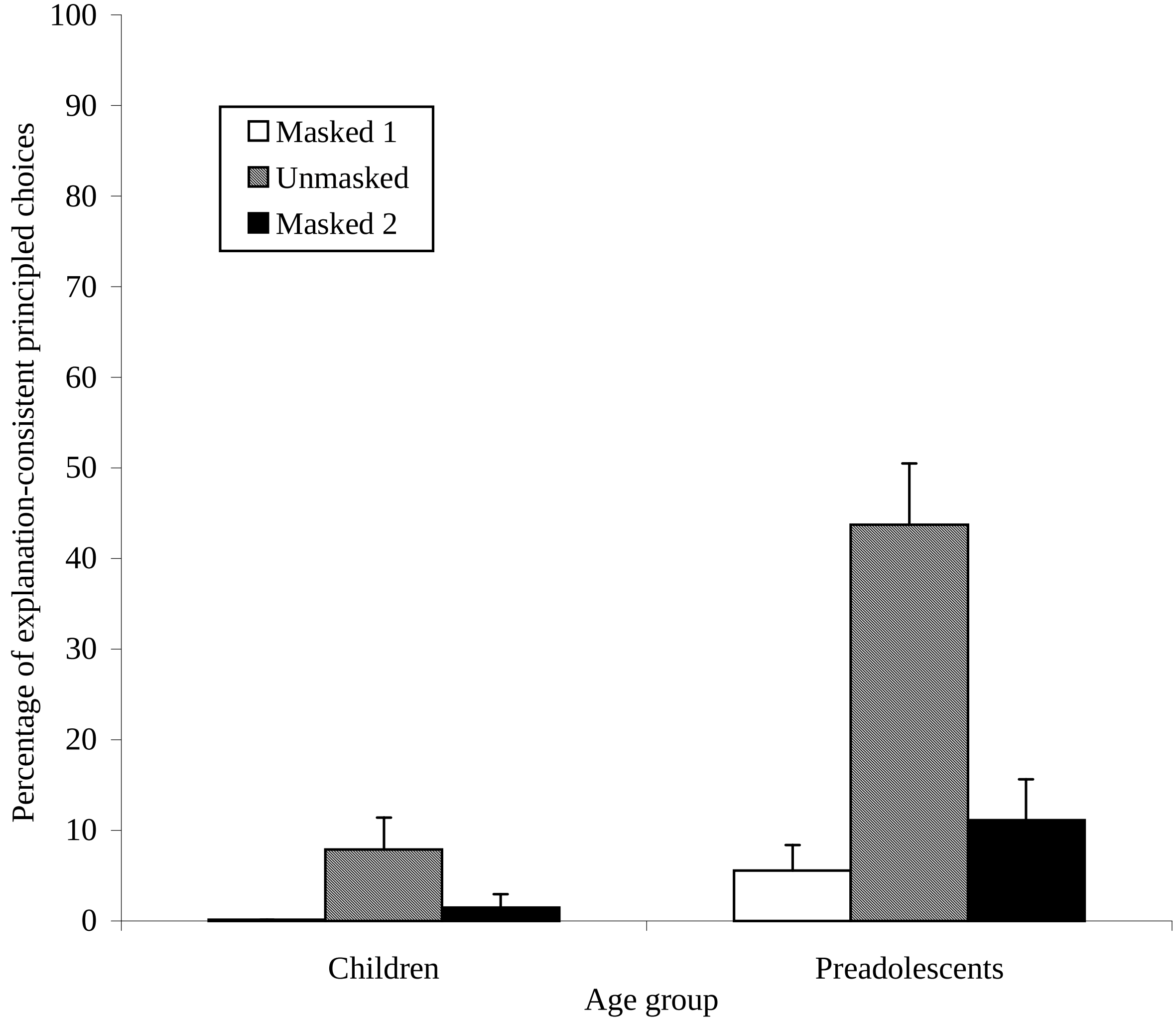

100
90
80
70
60
50
40
30
20
10
0
Percentage of explanation-consistent principled choices
Masked 1
Unmasked
Masked 2
Children
Preadolescents
Age group

Figure 2

Model 1 Model 2 Model 3

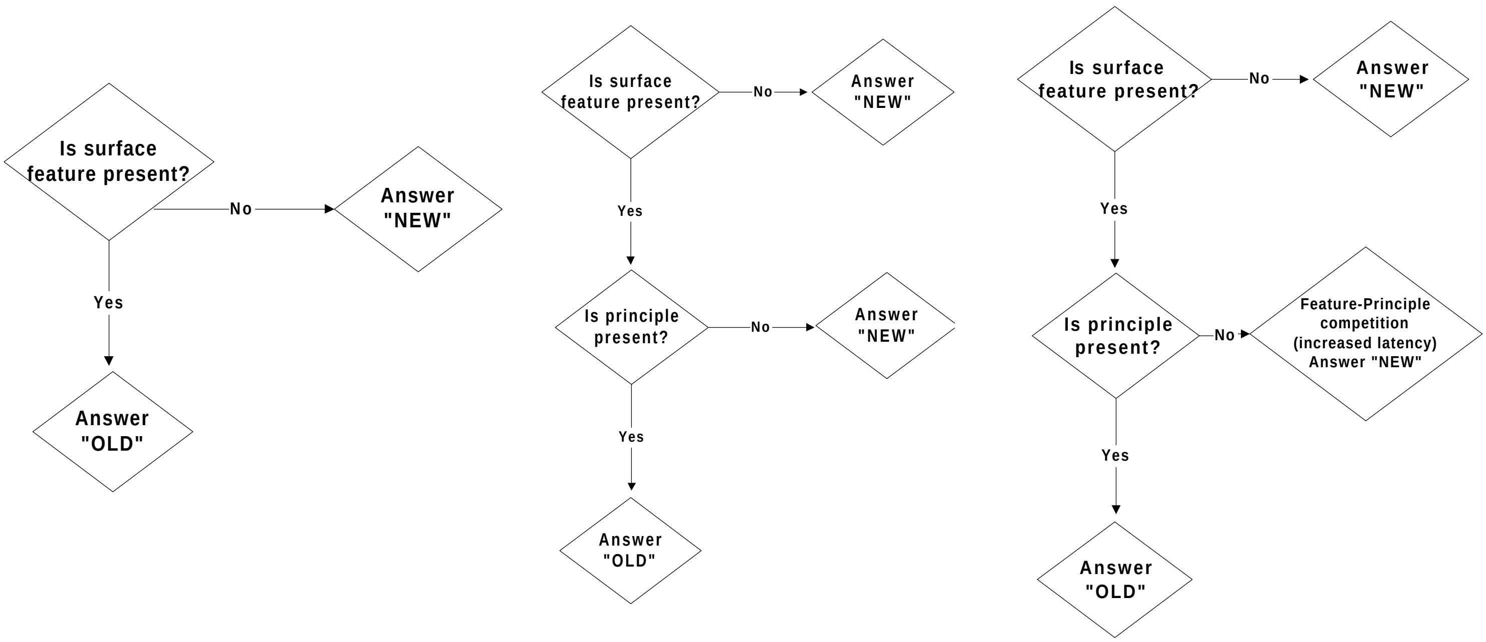


Figure 3

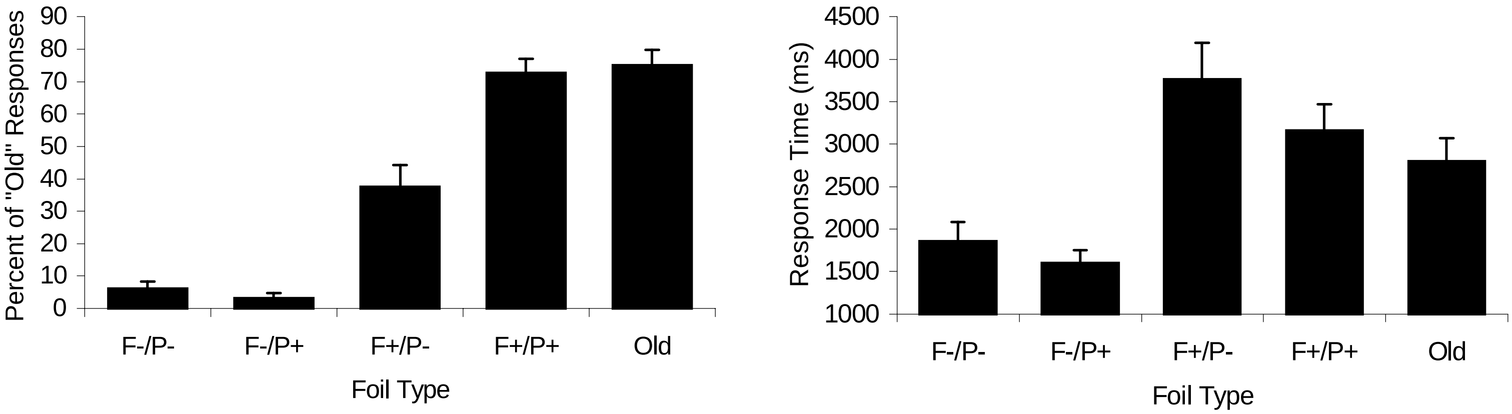
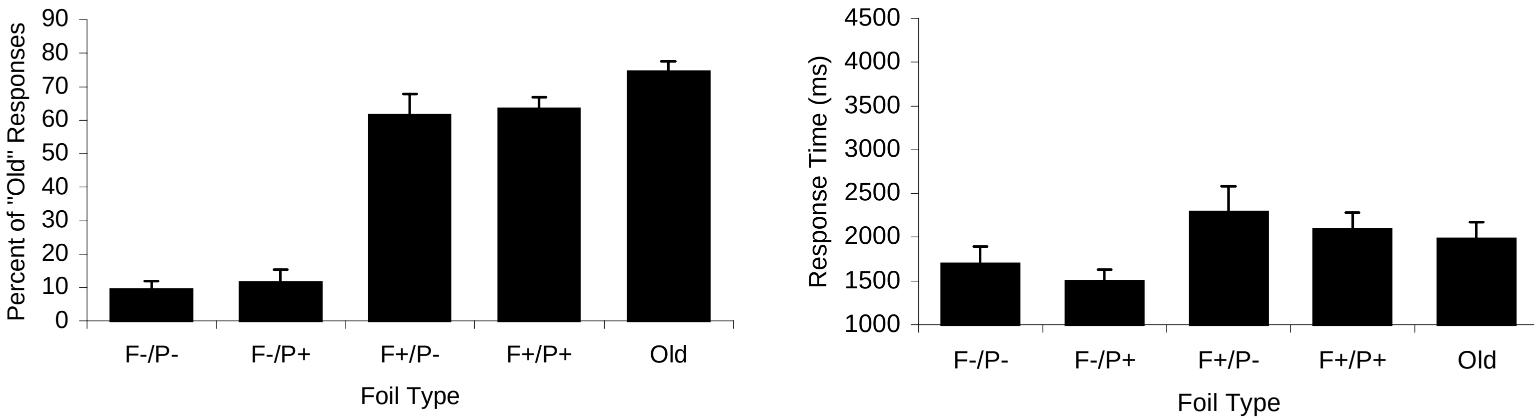

Percent of "Old" Responses
90
80
70
60
50
40
30
20
10
0
F-/P-
F-/P+
F+/P-
F+/P+
Old
Foil Type
Response Time (ms)
4500
4000
3500
3000
2500
2000
1500
1000
F-/P-
F-/P+
F+/P-
F+/P+
Old
Foil Type

Figure 4


Percent of "Old" Responses
90
80
70
60
50
40
30
20
10
0
F-/P-
F-/P+
F+/P-
F+/P+
Old
Foil Type
Response Time (ms)
4500
4000
3500
3000
2500
2000
1500
1000
F-/P-
F-/P+
F+/P-
F+/P+
Old
Foil Type