COGNITIVE SCIENCE Vol 22 (4) 1998, pp. 471–516 ISSN 0364-0213

On the Role of Mathematics in Explaining the Material World: Mental Models for Proportional Reasoning

DANIEL L. SCHWARTZ
Vanderbilt University

JOYCE L. MOORE
University of Iowa

Contemporary psychological research that studies how people apply mathematics has largely viewed mathematics as a computational tool for deriving an answer. The tacit assumption has been that people first understand a situation, and then choose which computations to apply. We examine an alternative assumption that mathematics can also serve as a tool that helps one to construct an understanding of a situation in the first place. Three studies were conducted with 6th-grade children in the context of proportional situations because early proportional reasoning is a premier example of where mathematics may provide new understanding of the world. The children predicted whether two differently-sized glasses of orange juice would taste the same when they were filled from a single carton of juice made from concentrate and water. To examine the relative contributions and interactions of situational and mathematical knowledge, we manipulated the formal features of the problem display (e.g., diagram vs. photograph) and the numerical complexity (e.g., divisibility) of the containers and the ingredient ratios. When the problem was presented as a diagram with complex numbers, or "realistically" with easy numbers, the children predicted the glasses would taste different because one glass had more juice than the other. But, when the problem was presented realistically with complex numbers, the children predicted the glasses would taste the same on the basis of empirical knowledge (e.g., "Juice can't change by itself"). And finally, when the problem was presented as a diagram with easy numbers, the children predicted the glasses would taste the same on the basis of proportional relations. These complex interactions illuminate how mathematical and empirical knowledge can jointly constrain the construction of a new understanding of the world. We propose that mathematics helped in the case of successful proportional reasoning because it made a complex empirical situation cognitively tractable, and thereby helped the children con-

Direct all correspondence to: Daniel L. Schwartz, Deptartment of Psychology and Human Development, Box 512-GPC, Vanderbilt University, Nashville, TN 37203; E-Mail: Dan.Schwartz@Vanderbilt.edu.

struct mental models of that situation. We sketch one aspect of the mental models that are constructed in the domain of quantity—a preference for specificity—that helps explain the current findings.

Mathematics plays a central role in modern science. It provides leverage for understanding the material world and for making predictions about worlds that have yet to be experienced. It is an age-old question why mathematics, a self-contained symbolic system, has such powerful empirical extension (Davies, 1992). The Pythagoreans, for example, proposed that the world and knowledge are governed by the same mathematical relations. Consequently, the fit between mathematics and the world of experience is assured (Changeux & Connes, 1995). More recently, psychologists have gathered evidence and developed theories that explore the relationship between mathematics and empirical experience. Case and Okamoto (1996), for example, propose that the integration of core empirical and numerical conceptual structures explains many developmental findings. More generally, a common research approach, relevant to the fit between mathematics and the empirical, is to explore the innate and experiential bases of quantitative understanding. Some researchers have documented early mathematical competencies such as enumeration. (e.g., Gelman & Brenneman, 1994; Wynn, 1992). This early appearance could suggest that mathematics provides leverage on empirical phenomena because people necessarily view the world through a cognitive filter of mathematical structure (e.g., an innate "accumulator," Meck & Church, 1983). Alternatively, others who are more interested in the development of formal quantitative skills have explored how mathematical knowledge evolves through social, symbolic, or physical interactions (e.g., Ben-Zeev, 1995; Cobb, Yakel, & Wood, 1992; Moore, 1993). This could suggest that mathematical knowledge maps onto the world because the structure of understanding is induced or constructed from an inherently "mathematical" environment (cf. Putnam, Lampert, & Peterson, 1990).

An alternative approach, distinct from tracing the developmental trajectory of quantitative knowledge, is to examine how people's problem solving incorporates properties of empirical and mathematical worlds. One might manipulate people's use of mathematical knowledge, empirical knowledge, or a combination of the two for a single situation (e.g., Ahl, Moore, & Dixon, 1992; Bassok & Holyoak, 1989; Heller, Post, Behr, & Lesh, 1990). In this way, one may examine the interaction of empirical and mathematical knowledge, and perhaps, make inferences about the mental structures that allow this interaction. This is the approach taken here. We present 6th-grade children on the cusp of formal reasoning with a problem that may be solved either on the basis of everyday experience or on the basis of proportional relations. By changing features of the problem, we invite children to reason based on empirical experience, or a combination of mathematics and empirical experience. We use the studies to support the proposal that mathematics provides leverage on the material world because it is used as a tool that makes a complex situation cognitively tractable, and thereby helps people construct mental models of that situation.

I. TWO VIEWS ON THE ROLE OF MATHEMATICS

To a large extent, contemporary research concerning how people apply mathematics to situations has viewed mathematics primarily as a computational tool for deriving an answer. The common premise has been that people first understand the structure of a situation, and then choose which computational approach to apply. At the risk of over generalization, it seems safe to say that most investigations of mathematics application adopt what we will call the EQM frame:

Empirical Situation ---> Qualitative Schema ---> Mathematical Procedure

According to EQM, people ideally interpret an empirical situation with a qualitative schema, and the schema in turn determines the numerical procedures they apply. We use qualitative schema as a convenient label for many different proposals. It can stand for the "intuitive" understanding of the ordinal or interval quantities of a situation (e.g., cold, warm, hot; Ahl, Moore, & Dixon, 1992), the representation of an empirical transformation and its effect (e.g., reshaping clay, Siegler, 1981), or even a more abstract and general representation like a part-whole schema (e.g., Resnick, 1992; Kintsch & Greeno, 1985). Despite the real differences between these meanings of qualitative, most research adopts the EQM frame in which mathematics comes after qualitative understanding.

In the developmental literature, for example, Inhelder and Piaget (1958) argued that children come to understand proportions through qualitative pathways rather than mathematical ones. They observed that "the subject first wants to isolate the [empirical] conservation for the same result... so that he can find the proportions, whereas he could have started from the operant relationships and their proportions in order to come to the idea of a potential compensation [i.e., empirical conservation]" (p. 220).

In the educational literature, there are powerful reasons to emphasize qualitative situational knowledge as a way to anchor the learning of mathematical knowledge (e.g., Cognition and Technology Group at Vanderbilt, 1997; Greeno, 1989). Many authors, for example, propose that students should activate "real-world" knowledge of a situation, so they will learn and use mathematics appropriately (e.g., Baranes, Perry, & Stigler, 1989; Hardiman, Wells, & Pollastek, 1984; Karplus, 1981). Supporting this idea, Ahl, Moore, and Dixon (1992) showed that reasoning about a qualitative version of a temperature mixing task (e.g., combining hot and cold water) improved performance on a subsequent numerical version of the task, but not vice versa. They subsequently claimed that "intuitive understanding sets an upper bound for generating [mathematical] solutions in novel or uninstructed domains..." (Dixon & Moore, 1996, p. 252).

Much of the adult problem-solving literature also adopts the EQM frame. This research often emphasizes the retrieval and similarity-finding processes that determine how individuals "choose" which schema or computations to apply to a given instance. One research line, for example, has investigated whether particular computational procedures are cued by particular semantic types (e.g., Hinsley, Hayes, & Simon, 1977). Bassok and Olseth (1995), for example, showed that people did not apply a formula to a discrete and continuous context with the same frequency even those contexts were mathematically isomorphic (e.g., ice delivery rate vs. ice melting rate). Another approach has shown that surface cues like the expression "all together" can cause people to retrieve specific computational methods such as addition—often times incorrectly when they have not evoked a mediating qualitative representation (Cummins, Kintsch, Reusser, & Weimer, 1988; Ross, 1989; Silver, 1986). Yet another approach has looked at analogical processes to explain how people find the structural correspondences between a novel and known situation (Novick, 1988; Reed, 1993). Greeno, Smith, and Moore (1993) summarize the standard view of analogical transfer: "In problems involving use of a formula, we suppose that the mental representation would include a symbolic [qualitative] schema that represents the pattern of quantitative information in the initial learning problems, along with a representation of the formula.... Transfer will occur if the pattern of quantities in the transfer problem is recognized according to the same schema that was used in initial learning" (p. 143). Notice that the qualitative schema does the work in understanding the new situation, not the mathematical formula.

Inherent to the EQM frame is the assumption that mathematics does not help explain the material world. If anything, qualitative understanding of the world helps explain mathematics, as in the case of using physical manipulatives to teach mathematics. We would like to propose, however, that mathematics does play a role in explaining the material world and that people do sometimes use their mathematical knowledge to help mediate powerful empirical insights (Flavell, 1972). Siegler (1981), for example, observed that in the domain of conservation, "Children seem to use their knowledge about the effects of transformations on number to learn about the transformations' effects on liquid and solid quantities" (p. 62). With respect to education, physicist Richard Feynman (1965) stated, "... it is impossible to explain honestly the beauties of the laws of nature in a way that people can feel, without their having some deep understanding in mathematics. I am sorry but this seems to be the case" (p. 39). And in the arena of problem solving, Sherin (1996) indicated an influence of mathematics by showing that algebra leads to an understanding of physics based on balance and equilibrium, whereas programming languages lead to a physics of process and causation.

One demonstration of how mathematics can shape situational understanding is that it can influence the empirical features that one considers. Imagine, for example, that people are deciding between the purchase of three- and four-legged stools (aesthetics aside). Because people have strong mathematical knowledge about additive relationships, they might model the situation emphasizing that four legs are greater than three legs. Consequently, they might decide that the extra leg makes a sturdier stool. Topologists, however, might model a stool in terms of the shape made by the legs. Three legs make a triangle, four legs make a square. Given this encoding, the topologists might infer that a three-legged stool is preferable because a triangle always makes a plane; it is never necessary to put a folded napkin under one leg to make the stool rest evenly on the floor.

In the stool example, mathematical and empirical knowledge interact to support inferences. We propose that this interaction often develops through the "on-line" construction of a mental model that captures the readily modeled information. In a domain that does not involve quantitative information, for example, Vosniadou and Brewer (1993) demonstrated how children capitalize on available information like overheard facts and perceptions to construct a coherent, albeit idiosyncratic, mental model of the earth. We suggest that in quantitative domains people also capitalize on quantitative information, to the extent allowed by their mathematical tools, to create coherent models in working memory. Of course, we cannot claim that this always occurs (e.g., Gelman, 1982; Hegarty, Mayer, & Monk, 1995; Roazzi & Bryant, 1993). Rather, this is how mathematical knowledge provides initial leverage for understanding the world. It helps identify and align relationships that make it possible to pull together different pieces of information into a tractable problem representation. We investigate this claim in the context of proportional reasoning.

II. REASONING ABOUT PROPORTIONAL SITUATIONS

The study of proportional reasoning and the related domain of rational number has produced a great variety of experimental tasks and factors, as well as categorizations of different types of rational numbers, solution strategies, elements of understanding, methods of instruction and developmental sequences (e.g., Carpenter, Fennema, & Romberg, 1993). Stated simply, a proportion is an equality between two ratios. Rather abstract questions involving proportions include determining whether 2/3 equals 4/6, finding the missing value x in the problem $x/3 = 4/6$, and determining whether the fraction 2/3 will necessarily become greater with an increase in the numerator, or denominator, or both. Proportions are difficult, in part, because of the need to consider a number of relations jointly in working memory (Case, 1978; Halford, Wilson, & Phillips, in press). For example, comparing 2/3 and 1/3 is relatively easy because one can "get away with" only considering the relationship between the numerators. Children as young as five can solve this sort of problem. But, they would fail on a problem comparing 2/5 and 1/2. Comparing 2/5 and 1/2 is more difficult because one needs to take the denominators into explicit consideration as well. There are different methods for handling the multiple possible relationships in 2/5 and 1/2, including cross-multiplying, re-computing the fractions so they have equivalent denominators or numerators, or converting the within-fraction relationship into a single construct (e.g., a percent) which in turn reduces the between-fraction comparison to a single, more manageable relationship.

To further complicate matters, when proportions are placed in an empirical context, people not only need to consider at least four distinct quantities and their potential relationships, they also need to decide which quantitative relationships are relevant. Cramer, Post, and Currier (1993), for example, report that a high percentage of adults incorrectly applied proportional reasoning to the following problem: "Sue and Julie were running equally fast around a track. Sue started first. When she had run 9 laps, Julie had run 3 laps. When Julie completed 15 laps, how many laps had Sue run?" (p. 159). This problem does not require proportional reasoning because Sue and Julie were running at the same speed; when Julie had reached 15 laps by running 12 more laps, Sue had also run 12 more laps making her total 21. As another example, consider mixing 1 oz. of orange concentrate and 2 oz. of water as compared to mixing 2 oz. concentrate and 4 oz. water. If the question is which mixture will taste stronger, people should compare ratios to determine that they are equivalent. But, if the question is which one will make more, then the appropriate quantitative relationship shows that they are not "equivalent." This is part of the challenge of connecting mathematics and the world: not only do people often need to coordinate numerous quantitative relationships, they also need to decide which mathematics is appropriate for which relationships. In many cases, people's qualitative knowledge determines which mathematical operations they should apply. But, as we will show, in some cases, mathematics helps people decide which empirical relationships are important. It does this by making it possible for people to model, and thereby understand, the possible relationships.

In the current experiments, we ask children to decide whether large and small glasses that have the same ratios of orange concentrate and water will taste the same (cf. Noelting, 1980a, 1980b). We use a proportion problem for two main reasons. First, the development of proportional reasoning represents a case where the successful use of mathematics can provide a gateway for understanding the empirical world in new and powerful ways. By investigating children who are on the cusp of proportional reasoning, we can evaluate the interaction between mathematical and empirical knowledge. Children's facility with proportional reasoning should not be very great, nor should their knowledge of juice mixtures. Therefore we may examine whether, and under what conditions, they manage to combine limited mathematical and empirical preparedness to come to a new understanding.

The second reason for focusing on a proportion task is primarily theoretical. Previous descriptions of mental models in the domain of mathematical understanding have tended to be analog representations of perceivable systems such as linearly ordered actions (Case & Okamoto, 1996), spatial arrays (Huttenlocher, Jordan, & Levine, 1994), or physical devices (Hatano & Osawa, 1983). If all mental models were to depend on perceptual analogs, it would limit their range of application considerably. Consider, for example, comparing the hues of two containers of pink paint that have been made from quantities of red and white paint. Although people can compare the pink hues perceptually, this does not mean they are reasoning about proportional relations (Lesh, Post, & Behr, 1988). To reason proportionally, people must separate the red and white visual components. This, however, removes the visible quality of interest; namely pink. If mental models must be analogs of perception, we suspect that the challenge of imagining "a relation of relations" (Piaget, Grize, Szeminska, & Bang, 1977) in the form of two pinks, two whites, and two reds would be prohibitively high. More generally, being able to model a higher-order equivalence through perception (e.g., speed) may not provide sufficient constraint to support the discovery of a structure that can explain that equivalence (e.g., distance over time). Although knowing the correct answer certainly helps, the backward mapping from answer to explanation may be quite difficult when the explanatory elements are not differentiated and articulated in one's knowledge of the answer (e.g., Karmiloff-Smith, 1979). Thus, it is important to investigate whether it is possible to describe characteristics of mental models that are not dependent on perceptual analogs and that can include structures that could support proportional reasoning.

Traditionally, the proportional reasoning tasks that have been used to investigate children's conceptual development require the consideration of quantities to achieve a correct answer. The quantitaties can be numerical and require the use of a ratio scale. Or, the relationships can be "qualitative" and require the use of ordinal and/or interval scales (e.g., small and large containers; Dixon & Moore, 1996; Heller et al., 1990; Spinillo & Bryant, 1991). A small change to this traditional problem characteristic may help illuminate the relationship between mathematical and empirical knowledge. Imagine a carton of orange juice made from 40 oz. of water and 24 oz. of concentrate. Also imagine a 4 oz. and a 7 oz. glass filled from the carton. Will the two glasses of orange juice taste the same? This scenario has two primary solution paths. The first is to rely solely on prior empirical experience: the juice in the glasses comes from the same carton, so it should taste the same. The second is to rely on a combination of experience and mathematics: the ingredient ratio is invariant across the containers and ingredient ratio determines flavor. For an adult, the empirical solution is obvious, and the mathematical solution explains the empirical results perfectly. For a child, neither solution may be obvious. Harel, Behr, Lesh, and Post (1994), who devised this task, found that many children predicted that the larger glass would taste stronger because it has more juice.

Although it seems odd that children would reason about the quantities of juice when the experiential answer should be so easily retrieved, it is a mistake to presume that the children were temporarily tricked into thinking that the larger glass would taste stronger. Weak experiential knowledge does not always assert itself, especially when constructing an explanation (Schooler & Engstler-Schooler, 1990; Schwartz & Hegarty, 1996). To help demonstrate that mathematical knowledge can influence the experiential knowledge people use, we conducted a pilot study using a problem more suited to adults: *There are three light switches in one room and three light bulbs in another room. Each switch controls one and only one bulb. When a switch is up, its corresponding light bulb is on. Assume that there is no way to see between the rooms without taking a trip. How many trips between the rooms will be necessary to map the switches to their respective light bulbs?* Of the 20 people asked, 19 decided that it would take 2 round trips. Telling the adults that, in fact, it only takes one trip did not help. The adults had turned the problem into a binary search. Their binary model had the effect of influencing which empirical features the adults brought to bear on the problem. So, paralleling the case of the children who could not "recall" the relevant fact that juice quantity has no effect on taste, the adults could not "recall" the relevant fact that a light bulb generates heat. To solve the problem in one trip, one can turn on switch A for a few minutes, turn switch A off, turn on switch B, and then go to the light bulb room. The lit bulb is connected to switch B, the warm bulb is connected to A, and the remaining bulb is connected to C. In terms of the current proposal, for both the orange juice and light bulb problems, people use their mathematical tools with the available quantitative information to help construct a model of a particular structure. Once constructed, the model shapes the inferences they can draw, right or wrong.

In the orange juice task, children, in their efforts to accommodate the numerical information, may model that one glass has more than the other. At the same time, in their efforts to accommodate the experiential properties, the children may model the juice as it is experienced. As experienced, juice is a unitary entity rather than an entity composed of concentrate and water. As a consequence, their models would include a manifest inequality between the quantities of juice in each glass and no structures for inferring an equality. This is because their models would not include the necessary information for constructing a ratio; they omitted the relationship between the separate ingredients. As a result, they might infer that different amounts of juice lead to different amounts of taste, much like different amounts of clay lead to different weights. In some cases, mathematics, such as subtraction, helps people model the world successfully, and in some cases it leads to models that, upon subsequent testing, turn out to be inadequate.

III. EXPERIMENTAL FACTORS

For children in the 6th grade, neither the correct empirical nor proportional solution path is so obvious that they will use it exclusively (Harel et al., 1994). Instead, it should be possible to manipulate children's use of empirical and mathematical knowledge. To do this in these experiments, we manipulated two dimensions of the juice scenario: physical realism and quantitative complexity. Figure 1 provides two examples of combinations of these dimensions.

Physicality

Prior research in both imagery and mathematical reasoning has shown that the physicality or photo-realism of a problem display can influence the models people construct. Realistic presentations often lead to simulations of empirical experience, whereas diagrams lead to analytic comparisons of static quantities. Schwartz (1995; Schwartz & Black, 1996a), for example, asked adults to determine whether marks on hinges and gears would meet if the mechanisms were put in motion. When the display of the mechanisms was photo-realistic, people imagined the dynamics of the devices closing or rotating into position. In contrast, when the display of the mechanisms was a diagram, people did not imagine the movement of the systems. Instead, they extracted the metric properties of the display (e.g., distances and angles) which they then compared or used in thumbnail derivations. Similarly, in the domain of mathematical reasoning, Moore (1993) presented children with pairs of actual winches or computerized diagrams of winches. Their task was to determine the relative behavior of blocks attached to the end of each winch. Moore found that children reasoned numerically earlier and more frequently with the diagrams than with the actual winches. Given these results, we thought a realistic presentation of the juice scenario might lead students towards a "hotter," more experiential perspective (cf. Mischel, Shoda, & Rodriguez, 1989). They might think of juice as a single flavorful entity. In contrast, a diagram might invite a "cooler," more analytic perspective. The children might emphasize quantitatively analyzable relationships.

Quantities

In addition to physical realism, we manipulated the problem quantities. There were non-numerical versions, versions with easily divided numbers, and versions with numbers that were more difficult to divide. Non-numerical versions of the problem stated that one glass was large and the other small. Because this information underspecifies the quantities, we thought that children would not be inclined to model the quantitative properties of the prob-

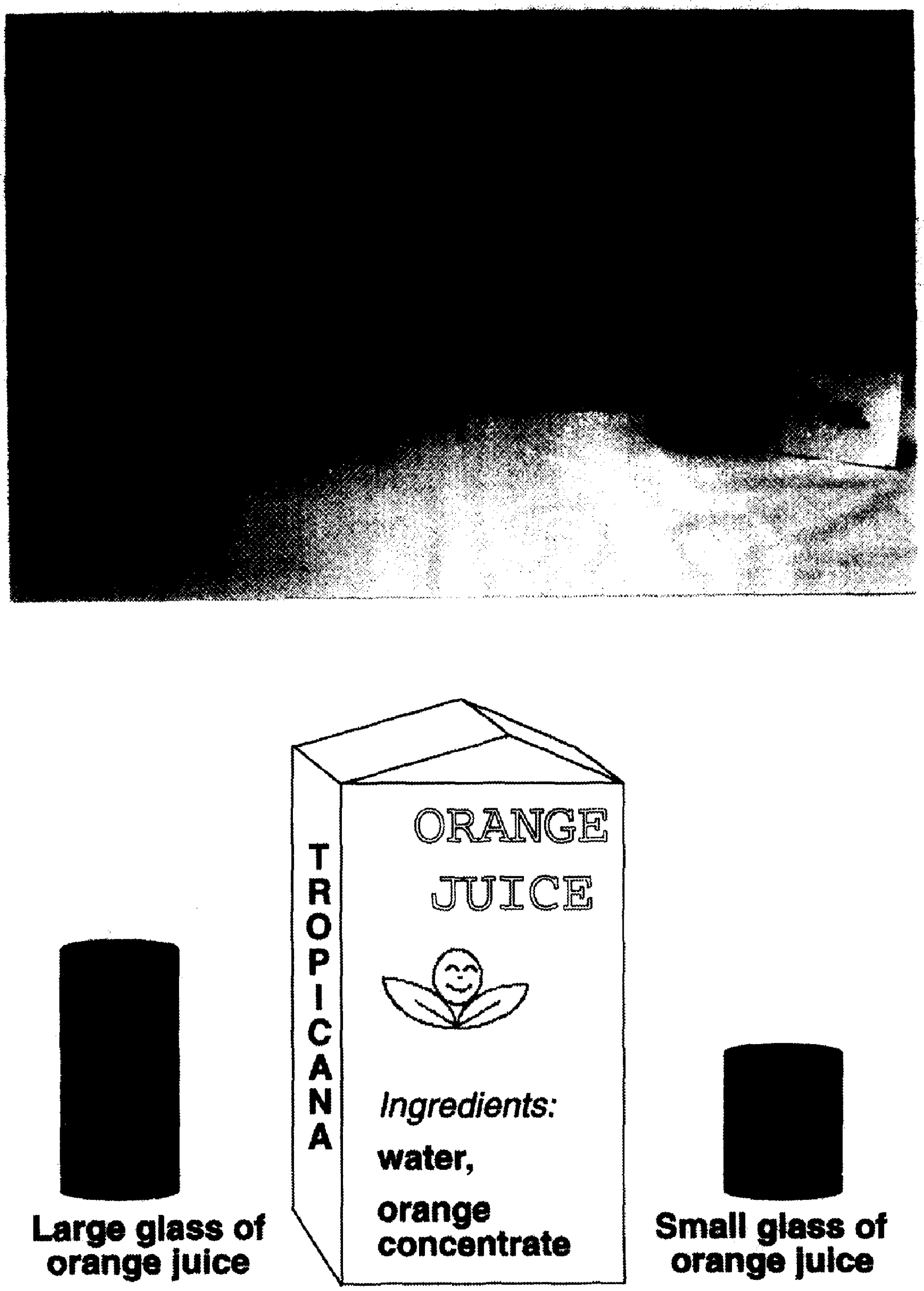

Figure 1. Two Examples of the Orange Juice Scenario Seen by the Children

lem. Numerical versions, on the other hand, do provide specific quantities. This may help children construct models that include quantitative information (cf. Greeno, 1991). As an analogy from a spatial domain, people are more likely to model "Mike is to the *left of* Bob," than "Mike is *next to* Bob," because the former specifies the nature of the spatial relationship (Mani & Johnson-Laird, 1981). The complexity of the specific numerical relationships afforded by the "easy" and "hard" problem versions may also affect model construction (Lovell & Butterworth, 1966; Lunzer & Pumfrey, 1966). Several researchers have documented that the numerical structure of a proportion problem influences performance (Hart, 1981; Karplus, Pulos, & Stage, 1983; Noelting, 1980b). So, for example, a carton with 20 oz. of concentrate and 20 oz. of water may afford a model of a one-to-one relationship between the two ingredients. In contrast, a carton with 24 oz. of concentrate and 40 oz. of water may not suggest an easily modeled relationship between the ingredients. Consequently, children may not construct a model that includes a structure that can capture the relationship between the ingredients.

IV. EXPERIMENT 1

In the first study, students in the 6th grade attempted the orange juice problem in one of four conditions. The conditions resulted from crossing the factors of display format and quantities. The children saw either a physical or diagram version of the juice problem presented at the front of a classroom, and the problem either indicated numerical or non-numerical quantities. The top of Figure 1 shows a photograph of the physical display that was used in the physical-numerical condition, and the bottom of Figure 1 shows the diagram used for the diagram-non-numerical condition. In each condition, the students' task was to decide whether the two glasses of juice tasted the same or different and to explain why.

Our primary interest was whether these non-structural problem variations would influence student understanding. One hypothesis was that the increased perceptual information of a physical display, relative to a diagram, would invite students to think about the problem experientially (e.g., "the juice comes from the same carton"). A second hypothesis was that the increased quantitative information of a numerical display, relative to a non-numerical display, would invite students to model a quantitative structure for the problem (e.g., "one glass is larger than the other"). We did not have *a priori* predictions for what would happen when cues to experiential and quantitative understanding were crossed, as in the cases of the physical display with numbers and the diagram display without numbers. This is the mixing ground of empirical and mathematical knowledge that we were exploring.

Method

Participants

The participants were ninety-eight students comprising four mathematics classrooms at an urban elementary school. The students were in the first 10 weeks of the 6th grade. For the experiment, the students left their original classrooms and were randomly mixed with one another.

Design and Materials

A 2 x 2 between-subjects design crossed the factors of *format* and *quantities*. The *physical* format used a carton of orange juice, two glasses, and three placards. The placards were placed next to the carton and glasses to indicate the relevant quantities. The *diagram* format used a poster board that showed line drawings of the containers and their quantities. The *numerical* quantities indicated that the carton of juice was composed of 40 oz. of water and 24 oz. of orange concentrate, and that the glasses were 7 oz. and 4 oz. The *non-numerical* quantities indicated that the carton had water and orange concentrate, and that one glass was larger and the other glass smaller. There were two dependent measures; the students' judgments about taste equality and their explanations. Students in all conditions received a form that asked, "Would the large glass taste the same as the small glass? If they would not taste the same, which one would taste more orangy?" The students circled their response. The bottom of the worksheet asked students to explain their answers.

Procedure

The students within each classroom were randomly reassigned to four new classrooms corresponding to the four conditions. Each room (condition) had a separate experimenter who read a script describing the problem setup. For the numerical conditions the script was, "A carton of orange juice was made from 40 oz. of water and 24 oz. of orange concentrate. A 7 oz. glass has juice from the carton, and a 4 oz. glass has juice from the carton." For the non-numerical conditions the script was, "A carton of orange juice was made from water and orange concentrate. A large glass has juice from the carton, and a small glass has juice from the carton." In the physical conditions the experimenters held up the placards as the corresponding information was spoken. In the diagram conditions the experimenters pointed to the information on the poster. In the physical conditions the experimenters poured the juice from the carton into the glasses as they noted each glass. After the experimenters completed their brief scripts they passed out the response forms. The forms were collected student by student as they were completed.

Results

Accuracy by Condition

A circled response was correct if it indicated that the glasses would taste the same, or if it indicated different taste with an explanation that relied on the carton of juice not having been shaken. Figure 2 indicates that over 50% of the students gave the correct answer in the physical display conditions regardless of quantitative information. In the diagram conditions, students only achieved this level of accuracy when there were non-numerical quantities. A logistic analysis crossed the factors of format and quantities with accuracy as the dependent measure (correct versus incorrect). A main effect of quantities indicates that the

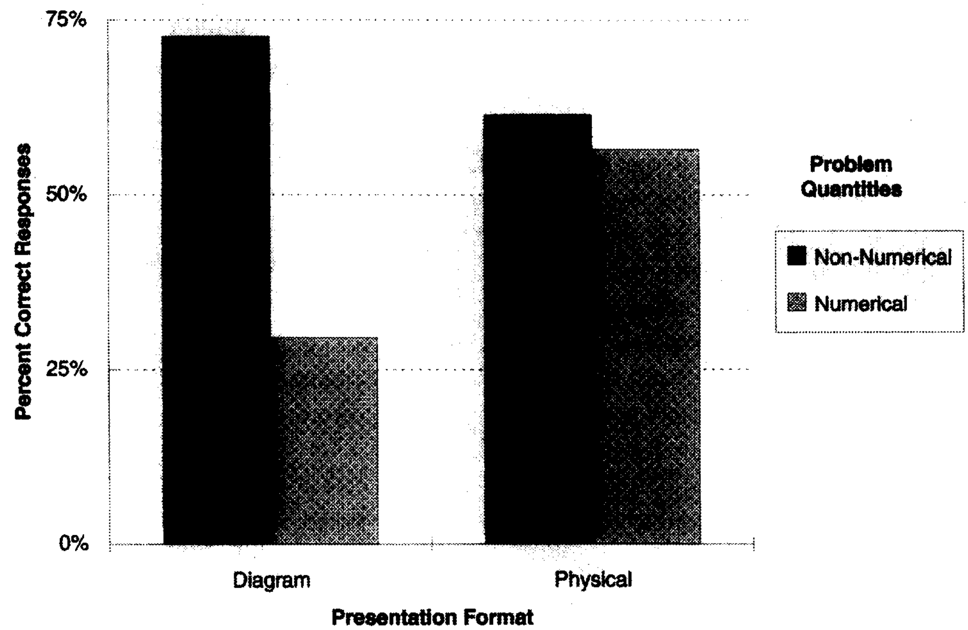

Figure 2. Percent Correct Responses by Condition (Exp. 1)

numerical versions led to poorer accuracy than non-numerical versions; $Z = 2.33$, $SE = .104$, $p < .05$. There was no main effect of format; $Z = .72$, $p > .45$; however, a marginal quantities by format interaction indicates that the numbers had less of an effect with the physical presentation than with the diagram presentation; $Z = 1.85$, $p < .07$.

Types of Explanations

We identified the explanations shown in plain type in Table 1. In this and the following experiments, a primary coder categorized each student's explanation. A second individual coded a random, one-fifth sample and made 93% identical codings (91% and 94% for Experiments 2 and 3, respectively). The primary coder's categorizations were used for analyses. The first column of Table 1 provides the distribution of explanations for Experiment 1.

For a representation to support proportional reasoning about this situation, it must represent the separate ingredients and their quantities. Upon review, we decided that student explanations could be factored according to whether they were *Quantified* and whether the concentrate and water were *Partitioned* into separate ingredients. There were four main categories of explanations: Quantified and Partitioned (QP), Quantified and Non-Partitioned (Q~P), Non-Quantified and Partitioned (~QP), and Non-Quantified and Non-Partitioned (~Q~P). (There was also a fifth category for Non-Explanations.) The explanations in Table 1 are organized by these factors and then further separated into same- and different-taste explanations.

TABLE 1
Explanation Frequencies Broken Out by Quantification (Q) and Partitioning (P) for the Three Experiments

	Percent of Explanations		
	Exp 1	Exp 2	Exp 3
QP Models			
Articulate Proportional			
Same Relative Numerical Amounts of Ingredients	0.0	4.0	5.6
Proto-Proportional			
Same Relative Amounts of Ingredients	0.0	5.3	3.4
Same Numerical Amounts of Ingredients as Carton	0.0	2.7	0.0
Same Amount of Ingredients	0.0	5.3	5.6
Same-Taste Totals	**0.0**	**17.3**	**14.6**
Two Ingredients without Ratio			
Bigger Glass has more Water and Concentrate	0.0	4.0	5.6
Smaller Glass has less Water and Concentrate	0.0	2.7	2.2
Glasses have different Amounts of Water and Concentrate	0.0	5.3	2.2
A Glass has more(less) Concentrate and less(more) Water	4.1	4.0	0.0
Centrate on One Ingredient			
Bigger Glass has more Water	3.1	0.0	2.2
Bigger Glass has more Concentrate	2.0	4.0	2.2
Bigger Glass has less Concentrate	1.0	0.0	0.0
Smaller Glass has less Water	1.0	0.0	0.0
Smaller Glass has less Concentrate	1.0	0.0	0.0
Different-Taste Totals	**12.2**	**20.0**	**14.6**
Q~P Models			
Quantity Irrelevant			
Amount of Juice does not Affect Taste	9.2	30.7	12.4
Same-Taste Totals	**9.2**	**30.7**	**12.4**
Quantity = Quality			
Larger Glass has more Juice	29.6	1.3	10.1
Less Juice means more Flavor	2.0	0.0	0.0
Different Sizes = Different Taste	0.0	0.0	1.1
Different-Taste Totals	**31.6**	**1.3**	**11.2**
~QP Models			
Constituent Identity			
Same Ingredients	6.1	5.3	2.2
Same-Taste Totals	**6.1**	**5.3**	**2.2**
Unequal Mixing			
Not Shaken, First Glass gets Pulp	4.1	5.3	1.1
Different-Taste Totals	**4.1**	**5.3**	**1.1**
~Q~P Models			
Causal/Temporal			
Common Source	13.3	10.7	21.3
Whole Identity			
Orange Juice is Orange Juice	17.3	9.3	9.0
Same-Taste Totals	**30.6**	**20.0**	**30.3**
Different-Taste Totals	**0.0**	**0.0**	**0.0**
Non-Explanations			
Reassertion of Answer	3.0	1.3	7.9
No Explanation	1.0	0.0	2.2
Uninterpretable	2.0	0.0	3.4
Non-Explanation Totals	**6.0**	**1.3**	**13.5**

The top entries of Table 1 show QP explanations. A QP explanation of the orange juice task is a necessary, but insufficient, condition for indicating proportional reasoning; students also need to consider the relationship between the constituents within each glass to support a proportional inference. In Experiment 1 there were no proportional or same-taste QP explanations, and we defer discussion of these explanations until Experiment 2. There were two main types of different-taste explanations. *Two Ingredients without Ratio*: Students explicitly compared the constituent quantities across glasses but did not consider the constituent relationships within a glass. For example, one student wrote that the larger glass would taste stronger because, "The larger glass has more concentrate and water." Another student explained, "The 7 oz. would taste more orangy than 4 oz. because it would have more orange concentrate than the 4 oz. and the 4 oz. would have more water than the 7 oz." *Centration on One Ingredient*: One ingredient was compared between glasses. One student wrote, "There is more space in the 7 oz. so there could be more room for water." This is a Q explanation because the student considered the amount of water. It is also a P explanation because the student identified the separability of the constituents, although he subsequently only focused on one of the constituents.

The Q~P explanations led to same- and different-taste responses. *Quantity Irrelevant*: Students made explicit statements that the quantity of juice is irrelevant to determining the taste. For example, "It will taste the same because orange juice is orange juice no matter how big the cup is." In this case, the student acknowledged the different quantities in each glass, but did not consider the ingredients. *Quantity=Quality*: Students reasoned additively that an increase in the measure of a quantity yields an increase in an associated quality, much as an increase in the quantity of clay yields more weight (Harel et al., 1994). For example, one student wrote that the large glass would taste stronger, "because there's more orange juice in the large glass." A possible interpretation of the Quantity=Quality mistake is that students were thinking "longer lasting taste" rather than "stronger orange flavor." Consequently, more juice would lead to more orange taste. Debriefing did not support this possibility, and Experiment 3 formally rejects this interpretation.

The ~QP explanations led to same- and different-taste responses. *Constituent Identity*: Students acknowledged the separable constituents within the glasses but made no reference to quantity. For example, "The glasses have the same ingredients." Although it is possible that this student was thinking that the glasses have the same quantities or ratios of ingredients, this would be a far inference based on what was written. *Unequal Mixing*: Students who gave different-taste responses relied on the physical properties of the orange juice mixture. One student reasoned, "probably most of the pulp was not at the top of the carton," hence the first glass filled would get less pulp. In this explanation, pouring and carton shaking play a causal role in determining the quality of oranginess, while at the same time the constituent of orange pulp is distinguished. Although this student thought in terms of the amount of pulp, we do not consider this a Q explanation because it does not quantify the concentrate or water.

The ~Q~P explanations always pointed to same-taste responses. *Causal/Temporal*: Students noted that the two glasses of juice came from the same carton, and there was no reason to suspect a change. As one student wrote, "They both have the same orange juice and they can't change by themselves." *Whole Identity*: Students reasoned that both glasses have the same kind of orange juice, focusing on the identity, rather than causal, relationship between the three quantities of orange juice. As one student wrote, "Orange juice is orange juice and it will taste the same." Using either of these models, students have not made a quantitative comparison, but instead they have reasoned about the physical properties of the situation.

The Relationships between Explanation Type, Accuracy, and Condition

We next consider whether the students' use of quantification and partitioning helps explain the circled taste responses and the condition effects. Although the same-different responses preceded the students' explanations, the explanations tended to reflect the responses. In several cases, however, students who gave same-taste responses wrote QP explanations that stated why the juice would taste different. These mismatches were infrequent and do not bias the analyses.

Table 2 indicates the percentage of correct responses associated with the four explanation types. The marginals show that Q explanations were associated with different-taste responses. Students with QP models were accurate only 8% of the time, and students with Q~P models were accurate only 27% of the time. In contrast, students who considered neither the quantities nor the separate ingredients (~Q~P) were accurate 100% of the time. A logistic analysis statistically tested the relationship between explanation type and solution accuracy. Quantification and partitioning were crossed factors with accuracy as the dependent measure. Quantifying the problem led to significantly poorer performance; $Z = 4.68$, $SE = .24$, $p < .01$. Partitioning the problem led to marginally poorer performance, $Z = 1.73$, $p < .09$. There was no interaction of partitioning and quantification; $Z = .63$, $p > .5$. In sum, students who made the correct response were not doing so on the basis of proportional reasoning. In fact, those students who considered properties that could have entered into proportional relations were the least accurate, whereas the students who did not consider those properties were the most accurate.

The relationship between the quantified explanations and the incorrect responses helps explain the poor performance in the diagram-numerical condition. It appears that the diagram-numerical condition caused the greatest inaccuracy because the students most frequently reasoned quantitatively in that condition. The frequencies in Table 2 indicate that the diagram-numerical condition led to the most Q explanations (74%) and the fewest ~Q~P explanations (17%). A logistic analysis formally shows that the format and quantity factors influenced the use of Q and P explanations. The two numerical conditions led to more quantified explanations; $Z = 2.08$, $SE = .14$, $p < .05$. A marginal format by quantities interaction shows that the numbers led to more quantified explanations in the diagram condition than the physical condition; $Z = 1.95$, $p < .06$, all other effects, $p > .3$.

Discussion

Even though the basic orange juice problem did not change across conditions, the children exhibited systematically different understandings. This indicates that qualitative domain

TABLE 2
Number of Explanations and Percent of Correct Answers by Condition (Exp. 1)

	Numerical				Non-Numerical					
	Diagram $n = 27$		Physical $n = 23$		Diagram $n = 22$		Physical $n = 26$		Totals $N = 98$	
Model Type	*n*	Correct	*n*	Correct	*n*	Correct	*n*	Correct	*n*	Correct
Q & P	5	0%	4	0%	0	n/a	3	33%	12	8%
Q & ~P	15	20%	8	38%	8	38%	10	20%	41	27%
~Q & P	1	100%	3	67%	3	100%	3	100%	10	90%
~Q & ~P	4	100%	8	100%	9	100%	9	100%	30	100%
None	2	0%	0	n/a	2	50%	1	100%	5	40%

Note. The Q and ~Q indicate whether or not the explanation considered quantities. The P and ~P indicate whether or not the explanation partitioned the constituents of concentrate and water.

knowledge is not solely responsible for how the children came to understand the problem. In particular, the numbers affected the students' understanding of the problem for the diagram format but not the physical format. Students in the diagram-numerical condition tended to state that the juice in the two glasses would taste different, whereas students in the other three cells of the design did not. A second finding is that the students did not appear to have a proportional understanding of the problem. There were no cases where a student provided a same-taste explanation based on a quantitative equivalence between the two glasses (i.e., equal ratios). Because many of the children did write quantified explanations, it seems unlikely that the children had a proportional understanding but just preferred not to bother with quantitative explanations. Our following explanation of these two findings emphasizes whether and how the children tried to bring together empirical and mathematical knowledge. First, we interpret the lack of proportional understanding, then we interpret the format by quantity interaction.

The Lack of Proportional Understanding

We propose that the children did not reason proportionally because their existing mathematical tools were not sophisticated enough to use the available quantitative information to help construct a ratio-based representation. As we argue in the *General Discussion*, quantitative mental models are built around specific quantitative relationships. Students of this age, however, have had little exposure to tools, such as percent, that can help them determine a specific multiplicative relationship between number pairs such as 40::24 and 7::4. Consequently, students in the number conditions did not have the mathematical wherewithal to convert the available numbers into modelable, multiplicative forms (cf. Karplus, Pulos, & Stage, 1983). Students in the non-numerical conditions were in a somewhat similar situation. The non-specific quantities (i.e., larger and smaller) did not offer a specific multiplicative relationship, and therefore did not support the construction of ratios. In either case, without a model of a specific ratio of the ingredients, the children could not develop a proportional understanding of the situation.

One alternative interpretation for the lack of proportional reasoning might be that the children had not developed conceptual structures for comparing ratios. A second alternative is that the children did not use their empirical knowledge to help constrain their understanding of the quantitative relationships. A result that makes both alternatives seem unlikely has to do with the nature of the students' errors. Only 8% of all the students centered on one constituent (e.g., the larger glass has more concentrate). Centering on one dimension of information is the common error made in proportion problems. The lack of centration suggests that the children tacitly understood that the ratio between the concentrate and water stayed the same when it was poured from the carton. Thus, the students seemed to have the conceptual wherewithal to understand that the constituent relationships in two glasses were comparable. Moreover, they used the empirical knowledge that the juice came from the same carton to inform this understanding.

One method for further weighing the preceding alternatives is to present the orange juice problem with easier numbers that are presumably within the purview of the children's mathematical tools. If the students reason proportionally, this would demonstrate that the challenge in bringing mathematics to the world is not always due to a lack of conceptual structure or due to a lack of empirical knowledge. Rather, it is sometimes due to a lack of mathematical knowledge that can help structure an understanding of a complex empirical situation.

The Effects of the Display Format and Quantities

The interaction of the format and quantity factors may be understood through the students' use of quantitative (Q) and partitioned (P) explanations. In the diagram-numerical condition 74% of the students gave Q explanations compared to 36% of the students in the diagram-non-numerical condition. Moreover, only 15% of the diagram-numerical students used ~Q~P explanations, compared to 41% of the diagram-non-numerical students. The numbers invited the children to reason quantitatively in the context of the diagram. What makes this an interesting result is that in the context of the physical display the numbers did not have an effect. Both physical conditions exhibited approximately 50% Q explanations and 35% ~Q~P explanations.

This effect may be explained by considering the influences of empirical and quantitative information in the four cells of the design. Let us suppose that the physical display provided sufficient empirical information (e.g., pouring the juice from the carton into the glasses) that the students had a "good understanding" of the problem without modeling the quantities. Given that the quantitative information was difficult to handle, they may have been disinclined to use it, and instead, relied solely on their empirical understanding. In contrast, the diagram conditions did not strongly afford reasoning about the perceptual aspects of the situation and instead pulled for a more analytic stance. In the diagram-numerical condition the children tried to use the numbers to analyze the situation (and failed). In the non-numerical diagram condition, the children were not given specific quantitative information that they could use in their analyses, so they relied primarily on their empirical knowledge, however weakly cued by the diagram.

An alternative class of interpretations that helps to clarify our position involves two forms of "seduction": perceptual and social. First, perhaps the physicality was such a powerful cue that the students were perceptually seduced to disregard the quantities altogether (e.g., Bruner, Oliver, & Greenfield, 1966; Bryant & Kopytynska, 1976). This interpretation differs from ours in a subtle way. Unlike the perceptual seduction account, we do not assume that children in the physical conditions disregarded quantitative information. Rather, we claim that the quantities were too difficult to help them construct a quantified model of the empirical situation (cf. Carpenter, 1975).

A second source of seduction may have been the social context (e.g., McGarrigle & Donaldson, 1975; Light & Gilmour, 1983). The simplicity of the correct answer may have caused the students to outsmart themselves. Maybe they assumed that there must be some difference if the experimenters were asking such an obvious question (e.g., Samuel & Bryant, 1984). This interpretation, however, does not explain why this did not occur equally in all the conditions. A more powerful alternative is that the students in the diagram-numerical condition perceived a "mathematical task demand" because the diagram looked something like a textbook problem. Consequently, they tried to "push symbols" in disregard of their empirical knowledge. Again, this interpretation differs from ours in a subtle way. The seduction interpretation claims that the diagram led the children to disregard empirical knowledge. We claim that diagrams, in general, lead to more analytic problem-solving approaches, not that they cause people to disregard empirical knowledge (Schwartz, 1995).

We can test the seduction interpretations with the same manipulation designed to test our account of why the children did not reason proportionally; namely, by using numerical relations within the children's mathematical purview. If the easier numbers lead children in the physical condition to use quantitative information, this would show that the physical presentation does not seduce the children to disregard quantitative information. And, if the students who receive easier numbers in the diagram format show correct understanding, then this shows that the diagram and numbers do not cause the students to disregard their empirical knowledge in favor of blind symbol pushing.

V. EXPERIMENT 2

The quantities factor of Experiment 2 included the levels of hard and easy numbers. The hard numbers replicated Experiment 1. For the easy number conditions, the glasses were 2 oz. and 4 oz. (2::4), and the carton had 20 oz. of water and 20 oz. of concentrate (20::20). Greeno (1991) has claimed that models involving specific multiplicative factors or divisors are easier to construct than those that do not involve specific values, and that models involving doubling and halving are probably available to many people. Supporting this claim, Spinillo and Bryant (1991) showed that young children can make proportion judgments involving the "half" boundary. The 20::20 carton may help children construct a half-half relationship. Moreover, it should be relatively easy to map the half-half relationship into the 2::4 glasses. So, on the one hand, these easier numbers may help children develop a proportional understanding of the juice scenario. On the other hand, the 2::4 glasses are in a specific doubling relationship. Consequently, the students may emphasize the larger quantity of juice in the larger glass. Based on Experiment 1, this should often result in a different-taste response. We did not have *a priori* predictions about how format would interact with easy numbers, so we develop our explanations in the *Results* and *Discussion*.

A complement to the proposal that mathematics can help people construct a model-based, situational understanding is the proposal that the subsequent use of mathematics can also modify one's model or one's faith in a model. Siegler (1981), for example, explained the initial appearance of conservation in small number tasks by noting that children use mathematical strategies on small numbers that test and support their initial qualitative hunches. "As soon as children suspect that adding objects to a collection or taking them away may affect their number, they can simply count or pair the collections before and after the transformation. Eventually, they come to realize that these tests always indicate that when something has been subtracted there are fewer" (p. 54).

To examine the effect of explicit mathematical work on a previously constructed understanding, we asked the children to attempt a mathematical justification after their judgments and explanations, and then afterwards, to make a judgment about the problem one more time. Per the EQM frame, we expected the students' initial understanding to affect their subsequent mathematical work. We also expected, however, that the explicit mathematical work would change the understanding of some students. In particular, students who originally judge same-taste may subsequently change their beliefs, if their initial models are not ratio-based. If their models are not ratio-based, then they cannot support proportional mathematics, and without proportional mathematics, the students should create an unsuccessful, explicit "test" of their same-taste beliefs. This negative test should undermine their same-taste confidence. In contrast, students who initially constructed ratio-based models should not revise their judgments, because their models should yield computational methods that are successful.

Method

Participants

Seventy-five students from an urban middle-school were randomly assigned to condition. The students were in the final 10 weeks of the 6th grade. The students had above average mathematical ability. For the 58 students for whom we could obtain results on the Tennessee Comprehensive Assessment Program, the average percentile ranking was 83.16 (SD = 18.67).

Design, Materials, & Procedure

The *quantities* and *format* factors were crossed to make four between-subjects conditions that were run in separate rooms. The levels of format were *diagram* and *physical*. The quantities factor had the levels of *easy* and *hard* numbers. The hard numbers replicated Experiment 1. The easy numbers used a carton with 20 oz. of water and 20 oz. of concentrate, and the glasses were 4 oz. and 2 oz. A within-subject factor, called *mathematizing*, was added to the current experiment. Students reported taste judgments both *before* and *after* they tried to construct a mathematical justification for their response. The mathematical justifications also served as a dependent measure. There were four phases to the experiment. In phase 1, an experimenter read the script. In phase 2, the students filled in a response form similar to Experiment 1 but that included a confidence scale ranging from 1 to 5. Students circled the number that indicated how sure they were of their same-different answers with 1 being "very unsure" and 5 being "very sure." In phase 3, students received a sheet that requested a mathematical justification for their response. Students were encouraged to write something and were given as much time as needed. In phase 4, the students circled their taste judgment and indicated their confidence on a scale from 1 to 5 on a new sheet.

Results

Accuracy Prior to Mathematizing

Figure 3 shows that, as before, students were more frequently correct in the physical-hard condition than the diagram-hard condition. Compared to Experiment 1, there were more correct responses for the hard number conditions. This could be a result of the higher math ability, the extra months of maturation, or instruction. More importantly, for the easy num-

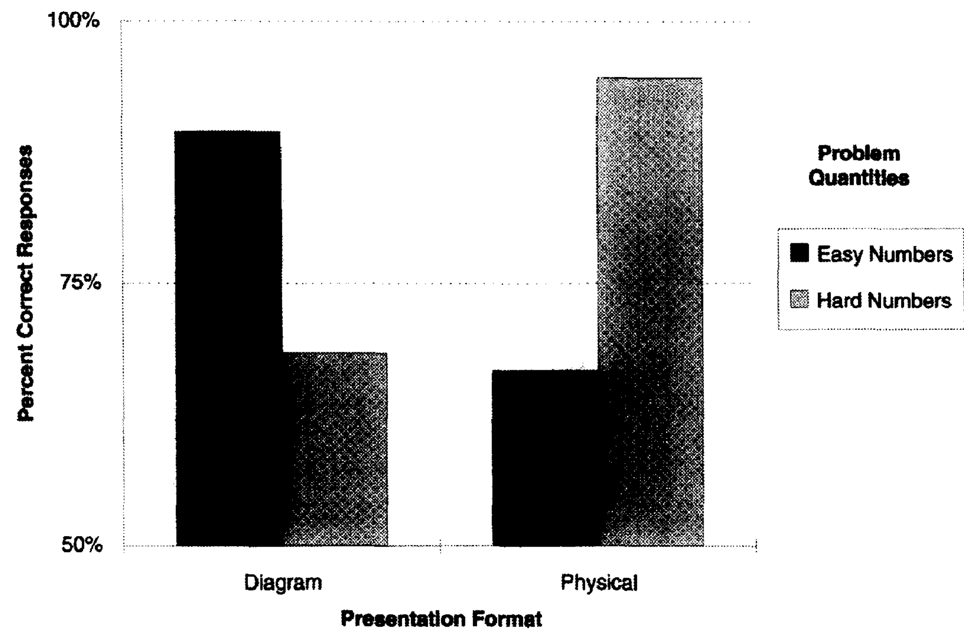

Figure 3. Percent Correct Responses by Condition (Exp. 2)

bers, students were more frequently correct in the diagram condition than in the physical condition. Thus, the effect of format was inverted for the easy and hard numbers. The format by quantities interaction was significant; $Z = 2.40$, $SE = .16$, $p < .05$, with no evidence of main effects; both Z's < 1. The success of the diagram-easy students shows that the combined effect of the diagram and numbers is not simply to lead the children into blind number pushing.

Table 3 shows the distribution of explanation types and their accuracy across the four conditions. Only 37% of the physical-hard students used Q models, whereas 83% of the physical-easy students used Q models. This latter result shows that physicality per se does not seduce children to disregard quantities. Overall, students used Q models more frequently in the easy-number conditions than in the hard-number conditions; $Z = 2.4$, $SE = .13$, $p < .05$. This supports our proposal that simpler numbers, more in the purview of the children's mathematical tools, would influence their understanding of the situation. There were no other significant effects; Z's < 1.5. The question at hand, then, is why the easy numbers led to inaccurate Q models in the physical condition but not in the diagram condition. The following analyses attempt to explain this result specifically and the crossover interaction more generally. We begin by considering the types of QP models found in the different conditions.

The Effect of the QP Models on Accuracy

This experiment generated proportional reasoning. Table 1 shows that 17% of the students gave correct QP explanations. *Articulate proportional:* Students explained that the ratio of

TABLE 3
Number of Explanations and Percent of Correct Answers by Condition (Exp. 2)

	Easy Numbers				Hard Numbers					
	Diagram $n = 19$		Physical $n = 18$		Diagram $n = 19$		Physical $n = 19$		Totals $N = 75$	
Model Type	*n*	Correct	*n*	Correct	*n*	Correct	*n*	Correct	*n*	Correct
Q & P	7	86%	7	14%	8	25%	5	80%	27	48%
Q & ~P	9	89%	8	100%	5	100%	2	100%	24	96%
~Q & P	2	100%	1	100%	2	100%	3	100%	8	100%
~Q & ~P	1	100%	2	100%	3	100%	9	100%	15	100%
None	0	n/a	0	n/a	1	100%	0	n/a	1	100%

Note. The Q and ~Q indicate whether or not the explanation considered quantities. The P and ~P indicate whether or not the explanation partitioned the constituents of concentrate and water.

concentrate and water was the same across the carton and two glasses. For example, one student wrote that the concentrate and water relationship "was 50-50." *Proto-proportional:* These explanations also led to a same-taste conclusion, but the use of language was less explicit. One student wrote, "The carton and glasses have the same amount of concentrate and water." Because the carton and glasses manifestly did not have similar amounts, we conclude that this child was trying to articulate that the containers had the same "relative" amounts of constituents but simply did not have the appropriate language to do so (e.g., "ratio" or "percent").

Not all the QP explanations generated in this experiment were proportional, however. In fact, excepting one student, only the QP explanations led to inaccurate responses. To demonstrate this effect statistically, the quantification and partitioning of the explanations were crossed factors with confidence score as the dependent measure. The continuous confidence scores allow us to use parametric statistics. Confidence scores were signed so that a different-taste response received a negative value (e.g., a same-taste confidence of 4 was scored as +4, and a different-taste confidence of 4 became -4). The average confidence scores, broken out by explanation type, were: QP = 0.44, Q~P = 4.02, ~QP = 3.25, and ~Q~P = 3.93. There was a main effect of partitioning on confidence; $F(1, 68) = 21.2$, $MSE = 86.5$, $p < .01$, and there was a marginal main effect of quantifying; $F(1,68) = 2.8$, $p < .1$. Both of these main effects, however, were compromised with a quantifying by partitioning interaction; $F(1,68) = 4.9$, $p < .05$. This interaction indicates that students with QP explanations had less same-taste confidence than students who used the other three models. These data complement the aforementioned fact that, excepting one student, only the QP explanations led to different-taste answers.

The confidence that students had in their QP models varied systematically by condition. Considering only those 24 students who used QP models, students were confident in a same-taste response in the diagram-easy ($M = 2.86$, $n = 7$) and physical-hard ($M = 3.0$, $n = 5$) conditions, but they were confident in a different-taste response in the diagram-hard ($M = -1.38$, $n = 8$) and physical-easy ($M = -1.71$, $n = 7$) conditions. The interaction of format and quantities on the confidence of students who used QP models is significant; $F(1,23) = 13.8$, $MSe = .72$, $p < .01$.

Why Easy Numbers Interacted with Format

Given the foregoing results, we are in a position to offer an explanation for why the easy numbers interacted with format. It appears that the diagram-easy students successfully constructed QP models with ratio structures, and this explains why they had high confidence in their same-taste responses. The easy numbers enabled them to map the carton's half-concentrate and half-water structure into the glasses. In contrast, in the physical condition, the easy numbers led to non-proportional QP models, and this explains why they had low-confidence in a same-taste response. Our interpretation of this latter effect is that the physical format led the students to model the juice primarily as a single material entity (just as this format had done in the physical-hard condition). At the same time, the 2::4 doubling relationship between the glasses invited the students to model the unequal juice amounts. As a result, the physical-easy students reasoned that different quantities of juice taste different.

A weakness of this explanation is that the students who made errors in the physical-easy condition were identified as having partitioned the constituents. Consequently, it is hard to claim that they were reasoning about juice as a unitary entity. Our Q and P factoring of the explanations, however, may be too coarse. The QP students who made mistakes tended to emphasize that the larger glass had more concentrate and water. Although they acknowledged the separability of the ingredients, as captured by our factoring scheme, they were still reasoning primarily with respect to the whole quantities of juice. Further evidence supporting the interpretation that the physical-easy students were reasoning with whole quantities may be had from their mathematical justifications.

Types of Mathematical Solutions

As we examined the mathematizing results, two coherent mathematical approaches emerged. The *half-half* solution used a numerical relationship between the constituents in each glass. For example, one student wrote, "20 + 20 = 40, 1 + 1 = 2, 2 + 2 = 4," implying two equal parts in each container. In contrast, the *double* solution used a numerical relationship between the amounts of juice in each glass. For example, one student wrote, "2 + 2 = 4 glass, 2 + 4 = 6 glass." Evidently, the student added two 2 oz. glasses of juice to make the 4 oz. glass, and then continued on to add the 2 oz. and 4 oz. glass to make a 6 oz. glass. A primary coder, blind to the format condition, categorized the mathematical work either into the *half-half* or *double* categories, and if neither, into an *other* category. There were no cases where a student tried both half-half and double solutions. The half-half category also includes more complex ratio attempts for the hard numbers. Similarly, attempts to add or subtract the glass quantities for the hard numbers were coded in the double category. Neither of these attempts was frequent, so we retain the more descriptive labels. The "other" category folds together a variety of mathematical attempts. The intent of these attempts were not easily evaluated from the written work, nor were they as relevant for understanding the contrast between the two easy conditions. A secondary coder categorized 40% of the mathematical efforts with 92% agreement.

Figure 4 shows that the two easy conditions led to most of the half-half and double explanations; $Z = 2.4$, $SE = .24$, $p < .05$. It also shows that 53% of the students in the diagram-easy condition used the half-half solution, compared to 11% in the physical-easy condition. Inversely, 44% of the students in the physical-easy condition used the doubling solution, compared to 21% in the diagram-easy condition; $Z = 2.4$, $SE = .27$, $p < .05$. These results suggest that the physical-easy students were modeling the juice as a unitary entity, because they used whole quantities in their mathematical justifications. In contrast, the diagram-easy students modeled the ingredients, as indicated by their use of half-half relationships.

Effects of Mathematizing on Same-Different Confidence

After mathematizing, there was a drop in the frequency of correct answers in all conditions except the diagram-easy condition (change in percent of correct answers: diagram-easy 0%, diagram-hard -25%, physical-easy -30%, physical-hard -29%). A logistic analysis

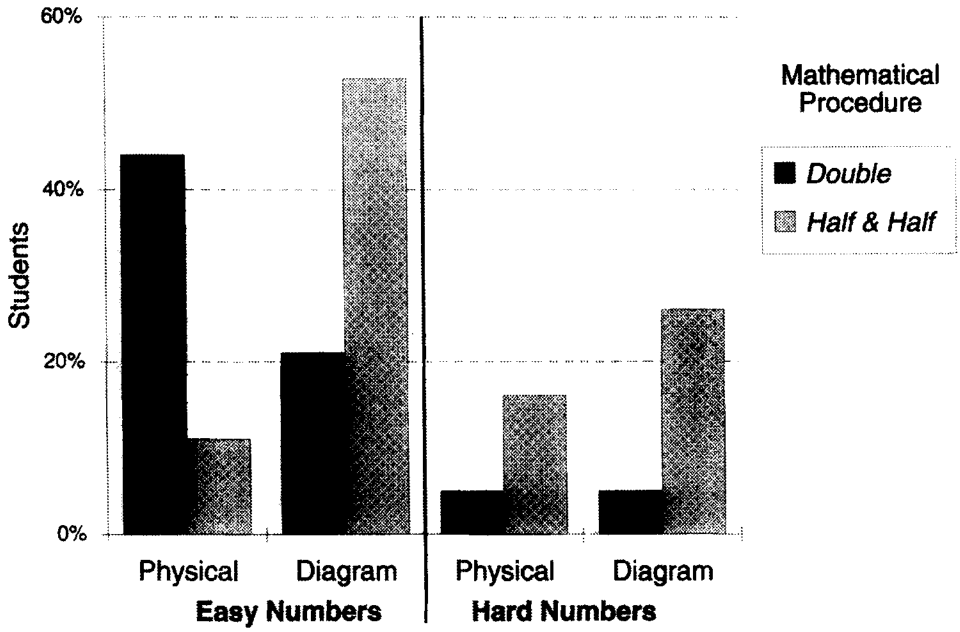

Figure 4. Distribution of Mathematical Justifications by Condition (Exp. 2)

indicated that students in the physical conditions were the most likely to change their answers; $Z = 2.5$, $SE = .20$, $p < .05$, no other effects. Figure 5 shows confidence ratings before and after students mathematized their answers. Except for the diagram-easy condition, in which we had found the majority of the successful half-half solutions, mathematizing caused decreases in same-taste confidence. This makes sense if one supposes that mathematics can be used retrospectively to evaluate one's understanding of a situation. The non-ratio models could not shape a computation that resulted in an equality between the glasses, and the students lost confidence in their same-taste answers.

To demonstrate the effect statistically, the confidence ratings before and after mathematizing were within-subject measures, with format and quantities as crossed between-subjects factors. The results yielded a complex set of two-way interactions. There was a main effect of mathematizing; $F(1,71) = 21.86$, $MSe = 5.72$, $p < .01$, whereby same-taste confidence diminished after mathematizing. However, this effect was compromised with a format by mathematizing interaction; $F(1,71) = 7.97$, $p < .01$, and a quantities by mathematizing interaction; $F(1,71) = 3.72$, $p < .06$. The interactions indicate that confidence dropped the most for the physical and hard quantity conditions. There was no format by quantities by mathematizing interaction; $F(1,71) = .07$, $p > .75$. The lack of the three-way interaction is surprising given the evident drop in confidence ratings in all but the diagram-easy cell. The three-way interaction did not reach significance because the diagram-easy condition confidences were higher both before and after mathematizing. As a result, the format by quantities interaction could be captured by a strong two-way interaction that averaged the before and after confidence ratings; $F(1,71) = 8.76$, $MSe = 15.55$, $p < .01$.

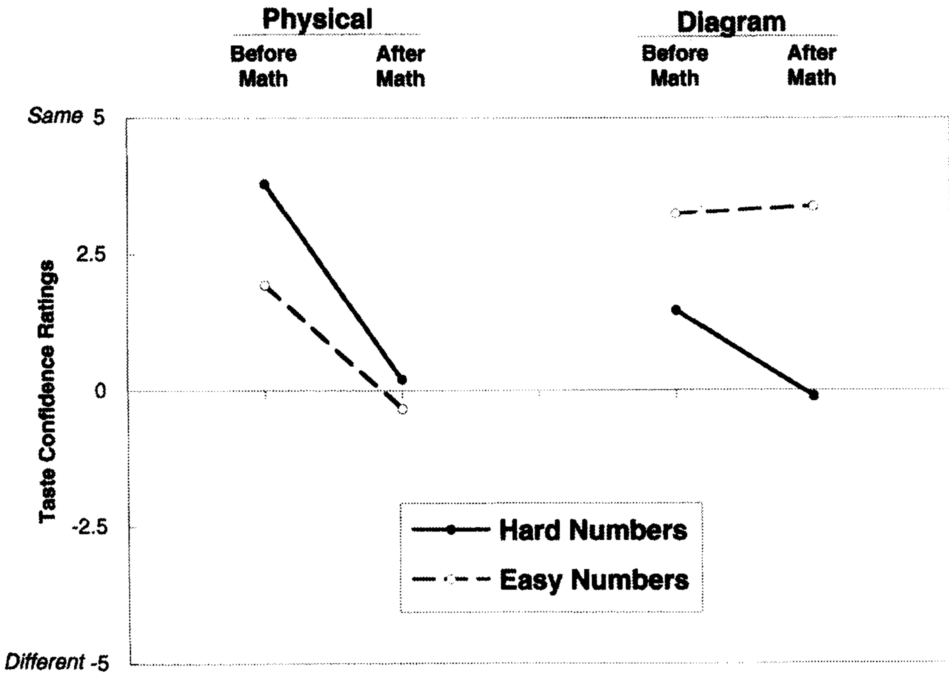

Figure 5. Changes in Confidence as a Consequence of Making Mathematical Justifications (Exp. 2)

Averaged over the before/after ratings, there were no main effects of quantities; $F(1,71) = 1.22, p > .25$, or format; $F(1,71) = .84, p > .35$.

Discussion

The current experiment demonstrated that if quantitative information is within the purview of children's mathematical knowledge it can have a profound effect on their understanding of and judgments about an empirical problem. It was mainly through the easier numbers that the children developed a proportional understanding of the juice problem. At the same time, because the quantities and format factors produced a cross-over interaction, it is clear that quantitative information and mathematical knowledge alone did not cause different types of understanding. When easy numbers were coupled with a diagram, the students correctly determined that the glasses would taste the same. But, when the easy numbers were coupled with a physical display, many students reasoned that the glasses would taste different. This latter result is intriguing because the students gave same-taste responses when the physical display included hard numbers.

One reason this finding may be surprising is that there is a tendency to think of mathematics as being applied via associative rules (e.g., if X then do Y). So, one might assume that if easy numbers are associated with a same response, and if physicality is also linked to a same response, then their combination should lead to a same response. This should occur whether the two rules strengthen one another, or whether one rule is dominant due to stronger cue-activation (Halford, 1993; Klahr & Wallace, 1976). The results, however, do not easily support a "rule-competition" explanation. One fix is to propose that there is associative interference between rules (Briars & Siegler, 1984). To our knowledge, however, this only occurs when associative rules have competing outcomes. Alternatively, one might postulate an additional rule. Siegler (1981), for example, added a rule to account for the fact that young children give correct responses to conservation problems with small numbers, but not large numbers. Although adding a rule does account for the data, we believe there are other considerations that can explain such interactive effects. For example, in Siegler's study, the children may have had the mathematical tools (e.g., counting) to handle small numbers but not large ones.

We use Figure 6 to help explain our interpretation. In the figure, we use a large font to indicate information that was in the mental foreground during model construction, whereas the small font represents less salient information. For example, one may see that for both physical conditions, the separate ingredients in the carton were not salient because the children were thinking of juice as a single, unpartitioned entity. One may also note that because of this they did not model an ingredient relationship in the glasses. For these students, the glasses were filled with juice, as indicated by the gray shading, not concentrate and water. In contrast, in the diagram conditions, students treated the ingredients as separate constituents in the carton. The jagged line of the upper-right panel indicates that in the diagram-hard condition, the students could not put the ingredients into a specific quantitative relationship. Consequently, these children could not carry the difficult ratio of 24::40 to the equally difficult glass sizes, and therefore they thought primarily of juice in the glasses (shown by the gray shading). In contrast, the students had enough mathematical sophistication to handle the diagram-easy numbers. They could carry the half-half ratio to all the containers, and therefore, they were able to keep the ingredients separate in their model of the glasses.

The easier numbers made it possible for students to use their mathematical knowledge to help structure their understanding of the problem. This was why there were more Q models in the easy conditions than the hard conditions. The use of mathematics, however, did not guarantee correct understanding. In the case of the physical format, for which the students thought of juice as a single entity, the 2::4 glasses led the students to model the doubling relationship, and they drew their inferences on the basis of unequal juice quantities. In the case of the diagram format, the simple numbers helped the students to model a half-half relationship in the carton, and the evenness of the 2::4 glasses allowed them to carry this relationship to each glass. As a result, they developed a model of equal constituent ratios on which to base their taste judgment. For both the diagram and physical presentations, we can see how empirical knowledge (e.g., juice as comprised of two ingredients, juice coming from the same container) and mathematical knowledge (e.g., 2+2=4) interacted to influence how the students understood the world.

The data suggest that the current examples of proportional reasoning resulted from models constructed in working memory during problem solving and did not result from the retrieval of an intact, long-term memory structure like a proportional reasoning schema.

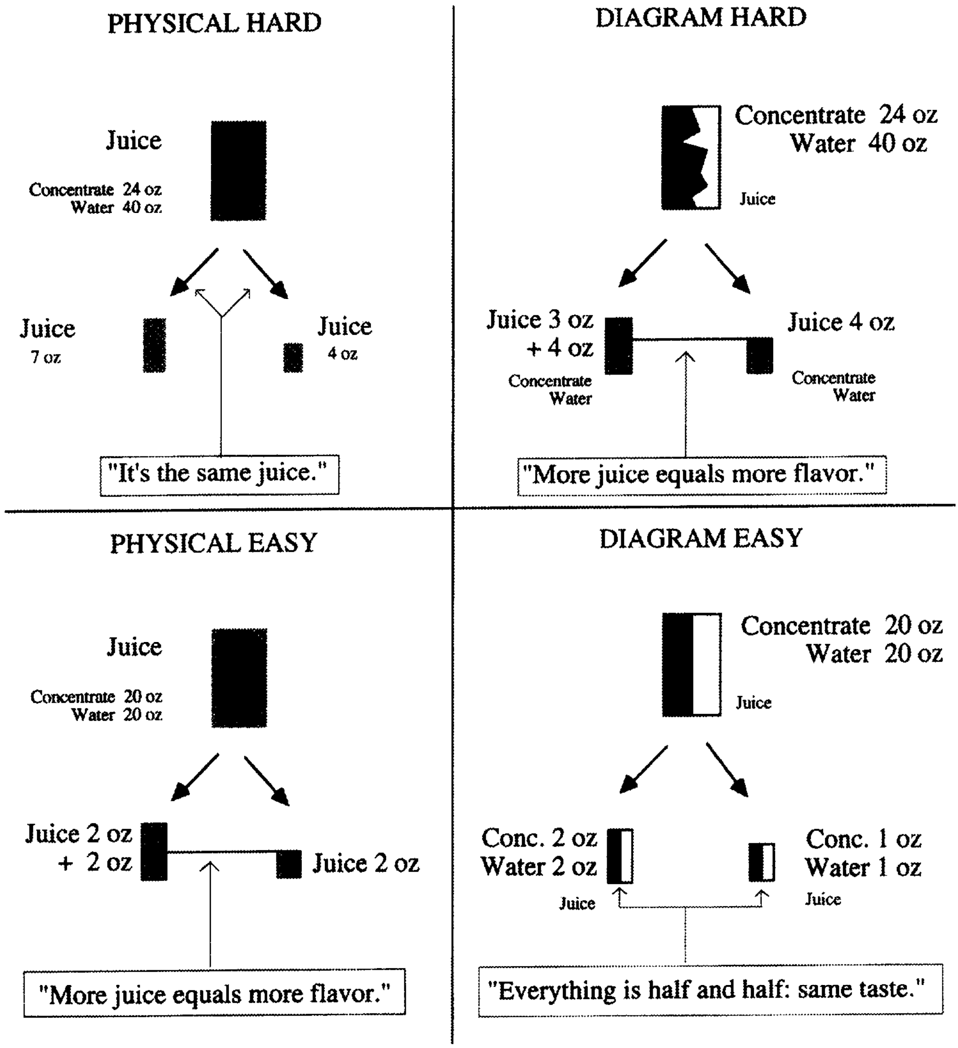

Figure 6. A Schematic of Children's Reasoning in Each Condition.
The large type indicates information that was in the foreground of the children's thinking, whereas the small type represents less salient information.

Consider that the students appeared to have tacit knowledge that the ratios would not change when poured from the carton; only 4% of the children argued for a different taste using the reason that one glass had more concentrate or water than the other. They also appeared to have retrievable knowledge that ratio determines flavor (as demonstrated by the diagram-easy condition). Yet, despite "possessing" the key information, the students did not reason proportionally when the numbers were difficult. We believe this is because the children needed to put the "tacit" information together into a coherent model before it constituted understanding. This is why the mathematical tools for handling easy numbers played such an important role. They unveiled relationships, like 1-to-1 correspondences, that the children could use to pull their models together.

Given our interpretations, we can make predictions about an experiment that makes even more subtle changes between the conditions. As before, the hard conditions use the 4::7 glasses, and the easy conditions use the 2::4 glasses. But now, both the hard and easy conditions use the 20::20 carton. The easy numbers should, of course, produce the same outcomes as before. But, according to our account, the new hard conditions should also produce the same outcomes as the previous hard number conditions. In the physical condition, the students should be inclined to think of juice experientially. Therefore, they will not model the partitioned 20::20 relationship in the carton. And, as before, the 4::7 glasses do not offer an easily modeled quantitative relationship, so the students will not incorporate the glass quantities into their model either. Consequently, they will tend to answer same taste based on empirical experience. In the diagram condition, the students should analyze the 20::20 relationship, but they should not have the mathematical tools to map this relationship easily into the 4::7 glasses. Consequently, they will not construct a model in which there is a specific ratio in the glasses, and they will tend to answer different taste because they will try to accommodate the 4 and 7 quantities additively by noting that one glass is "3 more" than the other. As an alternative hypothesis, perhaps the effect of the 20::20 carton in the diagram condition of Experiment 2 was to cue the retrieval of a proportional reasoning schema. If this is the case, then the 4::7 quantities in the following experiment should make little difference because the 20::20 carton will have already cued a proportional reasoning schema to organize the children's thoughts.

VI. EXPERIMENT 3

In this experiment, both the hard and easy number conditions used a 20::20 carton, with the only difference being the size of the glasses (4::7 and 2::4, respectively). Each child received a packet of materials that showed the problem as either a photograph or a diagram. Presenting the problem in packet form removes possible experimenter effects due to a group administration of the conditions. It also determines whether a "realism effect" extends to photographic displays as has been found elsewhere (DeLoache & Marzolf, 1992; Schwartz, 1995). If so, the results may have implications for the design of textbooks and the decision whether to use photographs or diagrams.

Experiment 3 included two further modifications to address lingering concerns. Experiment 2 recruited high-achieving students. To increase generality, the current experiment used an average student population similar to Experiment 1. A second concern was whether there was something odd about the details of the orange juice task that led to different-taste responses. To address this concern, a subset of students completed a set of secondary tasks in an interview setting. The tasks and their specific purposes are described below. If the secondary tasks show the same distribution of same and different responses, then this suggests that the results from the original scenario were not due to factors that were incidental to our main claims.

Method

Eighty-nine 6th-graders from four classrooms participated in the study. The students were in the first 10 weeks of the 6th grade and were of average ability. The students separated their desks from one another and were randomly handed one of four problem packets. For the *format* factor students either received packets with *diagrams* or color copies of *photographs* . For the crossed *quantities* factor the glasses were either in a *2::4* or *4::7* pairing. In either case, the carton was labeled with 20 oz. of water and 20 oz. of concentrate. Each packet included a cover sheet, the juice scenario, a sheet requesting a taste judgment and confidence rating (1 to 5), and a final explanation page. The students did not turn a page until they had completed their current page.

Based on their packet responses, two same-taste and two different-taste students from each condition were randomly selected for interview the next day. One interview was not completed due to experimenter error. The students solved four problems in order. (1) They re-solved the problem from the previous day. (2) They solved a *Paint* problem that was similar to the orange juice problem. This ensured that the previous results were not exclusive to non-visual judgments. Students saw a 10 oz. container of red paint and a 10 oz. container of white paint. They were told that the paint from another set of identical containers had been mixed to make a container of pink paint, and they were shown a pink paint container. They then saw two opaque containers. They were told that the two containers had been filled with the pink paint. Students from the 2::4 conditions were told that the target containers were 4 oz. and 8 oz. Students from the 4::7 conditions were told the containers were 5 oz. and 8 oz. The students were asked if the pink paint inside the opaque containers was the same color. (3) The students solved a *Spoons* problem from their original juice scenario. They decided whether a spoonful of juice taken from each glass would taste the same or different. If students believed that the spoonfuls tasted different, then they were not interpreting "more orangy taste" as "longer lasting taste." (4) In the *2 -> 1* problem, students were told that there were two glasses holding equivalent amounts of identical juice. They were told that one glass was poured into the other. They were asked whether this glass of juice would taste stronger. Because there were no numbers, this problem allowed us to examine whether the different-taste results were simply due to mindless number pushing.

Results and Discussion

Table 4 shows the frequency of correct response and Figure 7 shows the confidence scores. The experiment yielded the predicted cross-over interaction of format and quantities. The diagram-2::4 and photo-4::7 conditions led to more correct responses than the diagram-4::7 and photo-2::4 conditions; $Z = 2.15$, $SE = .11$, $p < .05$, and a greater average confidence in a same-taste response; $F(1,85) = 5.58$, $MSe = 13.7$, $p < .05$. This replication of Experiment 2 is striking in that the students were less advanced mathematically, a photograph was used rather than a physical set up, and only the sizes of the glasses differed between the numerical conditions. The interpretations of these results are the same as proposed in the *Discussion* of Experiment 2.

TABLE 4
Number of Explanations and Percent of Correct Answers by Condition (Exp. 3)

	2 oz. & 4 oz. Glass				4 oz. & 7 oz. Glass					
	Diagram $n = 22$		Physical $n = 22$		Diagram $n = 23$		Physical $n = 22$		Totals $N = 89$	
Model Type	n	Correct	n	Correct	n	Correct	n	Correct	n	Correct
Q & P	8	75%	6	33%	6	17%	6	67%	26	50%
Q & ~P	3	100%	5	40%	4	50%	9	44%	21	52%
~Q & P	1	100%	0	n/a	2	100%	0	n/a	3	100%
~Q & ~P	7	100%	6	83%	8	100%	6	100%	27	96%
None	3	67%	5	40%	3	33%	1	0%	12	42%

Note. The Q and ~Q indicate whether or not the explanation considered quantities. The P and ~P indicate whether or not the explanation partitioned the constituents of concentrate and water.

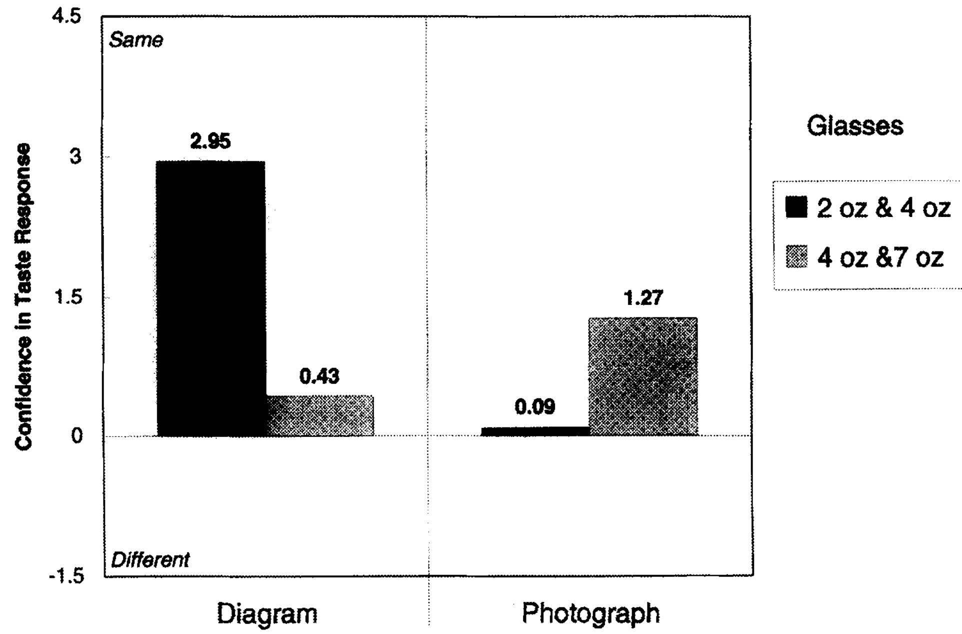

Figure 7. Average Confidence in Same-taste Responses by Condition (Exp. 3)

Table 4 shows that many students did not provide written explanations with the packet-based administration of the task. This makes a statistical analysis of the written explanations problematic due to self-selection; students who did not provide explanations were more likely to have given an incorrect response (see "none" rows of Table 4). Qualitatively, the patterns replicate the previous experiments in that the Q models generally led to incorrect responses in all but the diagram-2::4 condition. As in Experiment 2, students in the diagram-2::4 condition had a high rate of QP models, and these models were more likely to be correct than the answers based on QP models in the other conditions. Students in the photo-2::4 and diagram-4::7 conditions who used QP models were the least frequently correct of all. As before, the errors were primarily due to the belief that more juice (or more concentrate and water) would lead to more taste.

Figure 8 displays the responses to the interview questions. Each row corresponds to one student, rank ordered by accuracy. Three students consistently reasoned that the target containers would have the same quality (i.e., taste or color), and two students consistently reasoned that the containers would be different. The remaining ten students changed their answers across the questions. Although these changes might be taken as a sign of problems with task-reliability, they were expected because the previous results showed that subtle task variations affected student responses. Moreover, the variability makes sense if we assume that the students did not have a well-worn, higher-order schema (i.e., proportionality) that could determine the similarity among problems.

The summary frequencies at the bottom of Figure 8 indicate that the follow-up questions had nearly identical distributions of same-different responses as the original problem. Each

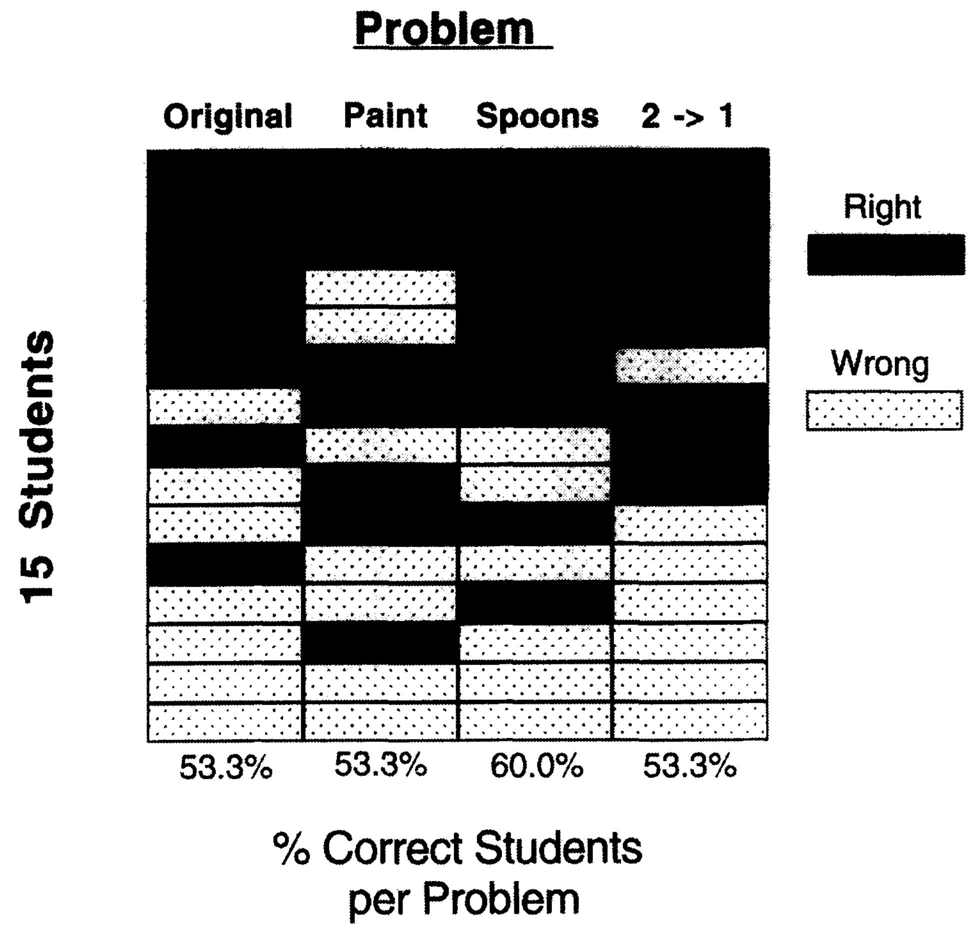

Figure 8. Distribution of Responses to the Four Interview Problems (Exp. 3)

problem counters a different alternative interpretation. The Paint problem indicates that the previous results were not peculiar to taste. The Spoons problem, in which the final quantity of juice in each spoon was equivalent, indicates that students were not generally thinking that stronger taste means longer lasting taste. Finally, responses to the 2->1 problem, in which two identical glasses were poured into one glass, show that different-taste answers were not the result of students being tricked into mindless computation. This is because the problem did not have numbers.

VII. GENERAL DISCUSSION

Three studies produced evidence supporting the proposition that mathematics provides leverage on the material world because it is used as a tool that makes a complex situation cognitively tractable, and thereby helps people construct mental models of that situation. The studies were conducted in the context of proportional situations because nascent pro-

TABLE 5
Overview of Average Taste Judgments by Experimental Condition

	Quantities of Carton & Glasses			
Format	Non-Numerical[a]	24::40 & 4::7[ab]	20::20 & 2::4[bc]	20::20 & 4::7[c]
Physical/Photo	Same-Taste	Same-Taste	Different-Taste	Same-Taste
Diagram	Same-Taste	Different-Taste	Same-Taste	Different-Taste

Note. [a]Experiment 1. [b]Experiment 2. [c]Experiment 3.

portional reasoning is an excellent example of where mathematics might provide new insight into the world. As children's knowledge of proportional relations evolves, so do the ways they can understand the empirical world. To develop evidence on how empirical and mathematical knowledge come to interact successfully, we manipulated children's use of empirical and mathematical knowledge by changing the numerical complexity and physical presentation of the orange juice scenario. The effects of these manipulations, summarized in Table 5, put us in a position to evaluate several explanations for the effectiveness of mathematics. We do this next and then sketch our account—quantitative mental models—more fully. We conclude with a discussion of the potential of our account with respect to cognitive development and instruction.

Understanding and Mathematical Application

One explanation for why mathematics effectively explains the material world is that people see the world through a cognitive filter of mathematical structure. Such a proposal would be consistent with modular nativism (e.g., Fodor, 1983) and has been used, most notably, to explain the universality of linguistic structure (Chomsky, 1966). Gopnik (1996) summarizes the position: "According to modularity theories, representations of the world are not constructed from evidence in the course of development. Instead representations are produced by innate structures, modules, or constraints. These structures may need to be triggered, but once they are triggered, they create mandatory representations of input" (p. 222). In quantitative domains, this claim may hold for some forms of topology because people invariably experience particular spatial structures, and it may hold at a "bootstrapping" level such that people necessarily see the world in terms of identities, aggregates, and symmetries. Nonetheless, as Piaget (1972) pointed out, such a claim does not explain why people can often find different mathematical structures in the same situation, why they apply a mathematical structure to one situation but not to its isomorph, and why they sometimes do not see any mathematical structure at all. Consequently, a cognitive filter explanation has difficulty explaining the decalage found in the current experiments. In proportional reasoning, at least, the effective application of mathematics is not a *necessary* consequence of people's cognitive constitution.

A second type of explanation, more suited to handling applications and misapplications of mathematics, is that there is an historical (empirically found) association between situations and mathematical knowledge. People apply mathematics to a given situation because they have learned which situations call for which mathematics. Such an associative account has great flexibility because the relationship between the empirical and the mathematical is structurally arbitrary. There are no constraints, other than historical success, for determining what associations should develop. The present results are problematic for straight-forward associative explanations. First, the results show that the application of mathematical knowledge cannot be solely ascribed to an historical association with a specific domain of situations; for the same task from the same domain children developed different understandings and different answers. Second, we explored a situation (physical-easy conditions) where two purported associative bonds, both calling for a same-taste response, would be activated. According to an associative account, one would predict a same-taste response, regardless of whether one association was stronger or whether the two associations co-activated a same-taste response. The children, however, tended to give a different-taste response.

A third class of explanation is that there are structural correspondences between situations and mathematical knowledge. In this explanation, people apply mathematics by finding the isomorphism between situations and mathematics. The most prevalent version of the isomorphism explanation adopts the EQM frame, whereby the structure of one's qualitative understanding of a situation serves as the basis for understanding or for selecting structurally appropriate mathematical operations. Such an "EQM-isomorphism" could explain the effect of the diagram with the 20::20 carton and the 2::4 glasses. Perhaps the half-half structure of the carton helped the children to think in terms of a qualitative proportional schema, which in turn made them aware that they could compare ratios for the 2::4 glasses. If this explanation were correct then we should have also expected the children to reason proportionally for the diagrams with the 20::20 carton and the 4::7 glasses. The half-half structure was present here as well. Experiment 3 showed, however, that the children did not reason proportionally for the 4::7 problem. Evidently, the understanding of proportional structure came through children's ability to model the half-half ratio in the 2::4 glasses, rather than preceding it in a qualitative form.

There is another version of the isomorphism explanation that does not entail the EQM assumption. In this version, people map their experience of the world directly into mathematical structures. For example, Bassok and Holyoak (1989) demonstrated that when solving a physics problem people will map a formula learned in the context of algebra, but when solving an algebra problem they will not map a formula learned in the context of physics. According to theories of analogical process, people might understand a novel empirical situation by mapping it into a mathematical representation with a known structure (e.g., Gentner, 1983). In the current experiments, however, the children did not appear to have robust mathematical structures that could serve as knowledge sources. Instead, they appeared to combine partial empirical and mathematical knowledge into novel working memory structures. This raises the question of how the children determined which "parts" to include in their final knowledge structure, given that neither partial knowledge source can determine this alone. In the situation of mapping between partial knowledge, it is important to specify higher-order constraints that determine which mappings are acceptable. There have been several proposals for these constraints (e.g., Gentner & Toupin, 1986; Spellman & Holyoak, 1996). Halford, Wilson, and Phillips (in press), for example, argue that people choose relationships that reduce working memory demands and that this tendency can explain which relationships people will map through analogy. In the following section, we also describe one possible higher-order constraint that stems from the nature of working memory, although we do not emphasize analogical processes or isomorphisms.

One reason that we do not emphasize the use of analogy to map between empirical and mathematical knowledge is that it often comes with two assumptions. The first common assumption is that understanding in one domain is "borrowed" from one's understanding in another domain. The mapping process between domains is what "carries" this understanding across source and target. In the current case, however, the children appeared to develop new understandings by integrating empirical and mathematical knowledge into a single structure. The second common assumption is that there is a pre-existing similarity or isomorphism between the source and target domains, in this case mathematical and empirical. This similarity assumption *presupposes* why mathematics can so effectively explain the world (cf. Medin, Goldstone, & Gentner, 1993), but this is what we are trying to investigate. To avoid each of these assumptions, we do not try to explain the effectiveness of mathematics from a structuralist perspective in which the emphasis is on structural correspondences. Instead, we explore a functionalist perspective in which mathematics helps cognitive operations. So, instead of viewing mathematical and empirical knowledge as parallel but isomorphic streams across which people create bridges, we treat them as sources of information that feed into and co-constrain a process in which people construct a single representation. This representation is consistent with, but not equal to, the original streams of knowledge. The completed representation, which we call a quantitative mental model, may in turn yield new insights about the world.

Quantitative Models

The preceding explanations incorporate important general mechanisms including analogical mapping, retrieval, associative strengthening, and bootstrapping architectures. These general mechanisms should be included in a full explanation of the current data. Our goal, however, is somewhat more domain specific. We are trying to explain why a formal system of mathematics, as opposed to something like a formal system of logic, provides such powerful leverage on the world. Our speculation is that mathematics provides a set of cultural tools that function to simplify complex situations into cognitively tractable structures. An average and standard deviation, for example, provide a way to reduce the complexity of a distribution into a form that is more amenable to internal reflection. In contrast, logical formalisms seem to increase complexity (in our experience at least), which may explain why logic instruction has dubious lasting effects compared to mathematics instruction (e.g., Nisbett, Fong, Lehman, & Cheng, 1987).

In addition to the current data, prior evidence indicates that children can demonstrate numerical conservation (Siegler, 1981; Winer, 1974; Zimiles, 1966) and proportional reasoning (Hart, 1981; Karplus, Pulos, & Stage, 1983; Lave, 1988; Noelting, 1980b) with "easy" numbers even though they cannot do so with hard numbers. Our explanation is that the hard numbers exceed the mathematical tools available to the children, and consequently, they cannot transform the difficult numerical relationships into modelable forms. If we entertain this functional role for formal mathematical knowledge, the critical question concerns the nature of mental models that makes them amenable to mathematics and to the synthesis of empirical and mathematical information.

While there are numerous proposed properties that distinguish mental models from other types of representational structures (e.g., Holland, Holyoak, Nisbett, & Thagard, 1986; Johnson-Laird, 1983), we emphasize specificity here (Schwartz & Black, 1996a). A mental model is constructed "on-line" to represent a specific state of affairs. The property of specificity is why, in the domain of deduction for example, it is often necessary to entertain multiple models to draw an inference; one needs a separate model for each specific possibility (Johnson-Laird, 1983). In the domain of quantity, specificity implies that mental models should not be characterized as relational structures in which specific numerical attributes only play an incidental role. Although models hold relational information, they also require specific instantiations. One does not model a triangle in general, one models a triangle with specific angles. This is not to claim that one cannot model a specific triangle to aid in drawing general conclusions, or that one cannot develop non-model representations for things like triangles-in-general. Schwartz and Black (1996b), for example, asked people to reason about the motions of connected gears. When the problems were novel, people modeled each gear's specific motion with their hands. However, over a period of time, people induced a more abstract rule (e.g., all odd gears turn the same direction) that did not require a model representation of each specific gear and its motion. Perhaps in the domain of proportional reasoning, the construction of specific models provides the basis for the induction of more abstract schemata that are stored in long-term memory.

One reason that we suggest characterizing early proportional reasoning in terms of mental models is that the children's understanding was highly influenced by the specific quantities of the problems. Moreover, the children did not exhibit proportional understanding when there were no specific quantities. Similarly, Bowers, Cobb, and McClain (in press) found that young children's understanding of the conservation of material quantity was intertwined with their ability to manipulate specific numerical quantities. The children apparently needed specific values for their initial quantitative understandings to take form. One can highlight the significance of these findings by considering approaches to understanding in which relationships, rather than values, play the primary role in the construction of representations. As a specific example, Halford, Wilson, and Phillips (in press) propose that entertaining multiple binary relationships entails a high processing load, and consequently, people will try to embed (chunk) those relationships in a higher-order relationship (or they will treat them sequentially). This type of relational approach is primarily syntactic in that what matters is the order of the relationships (e.g., unary, binary, ternary), not their content. According to this approach, two binary relationships like *double (4, 2)* and *more-than (4, 2)* are equivalent with respect to their working memory demands regardless of their specific content. The current results show that this type of domain-independent, relational approach is insufficient for modelling quantitative understanding; it is also necessary to consider the semantics of the specific quantities. So, for example, even though they are both binary relationships, children in the photograph conditions systematically modelled the double relationship between the 2::4 glasses, but they ignored the more-than relationship between the 4::7 glasses. We propose that the key to understanding this effect involves a consideration of the role of specificity in quantitative mental models.

To make specificity a generative construct, it is necessary to outline those characteristics that support model construction. For example, one might compute π to a thousand digits, but this kind of specificity would not make π any easier to model. Our suggestion is that mental models, in general, have a commensurability constraint. This constraint is captured by the expression, "You can't compare apples and oranges." Individuals prefer mental models in which the represented elements can be described in terms of one another. With respect to quantitative models of ratios, the commensurability constraint manifests itself as a preference for multiplicative specificity. Mental models best represent entities that constitute integer multiples because the multiples may be represented in terms of one another (e.g., 2::4, 20::20, a symmetrically folded paper).

One reason that models prefer multiplicative specificity is that it provides informational redundancy and a correlated decrease in working memory demands. One entity helps to specify the other. This co-specification might be thought of as a form of chunking that reduces the number of free parameters or independent dimensions of information in working memory (cf. English & Halford, 1993). One might speculate that working memory is constituted to capitalize on natural symmetries where one part "reflects" another. So, for example, in the current research, a child in the physical-easy conditions may have applied two distinct mathematical operations to the 2::4 glasses. She may have figured that 4 is double 2, and she may have figured that 4 plus 2 equals 6. Because the former relies on a "doubling," it presumably demands less of working memory, and therefore the child may have been more inclined to incorporate the double structure into her developing representation of the problem.

In the following paragraphs, we consider different levels of co-specification to motivate our hypothesis. A 1::1 relation should be the easiest to model because the quantity of one term completely specifies the quantity of the second term. The relative ease of representing a one-to-one correspondence may explain why the children in the current studies modeled the 20::20 carton in a part-to-part, 1::1 relation but never in a part-to-whole, 20 out of 40 (or, 1 out of 2) relation. The part-to-part relation only requires a simple reflection.

At the next level of co-specification is a 1::2 relation, because one needs an additional term to indicate the "times 2" relationship. Still, the terms do co-specify one another, and therefore, we should expect that children would be able to reason proporitionally about a carton with 10 oz. of concentrate and 20 oz. of water before they could handle the hard numbers used here (e.g., 24::40).

A much harder type of relation is 4::7 because the terms cannot be easily described by one another; the 4 cannot serve as an indivisible base unit for specifying the 7. Tools such as percent help people model non-integer relations like 4::7 in a relatively simple form (i.e., 57%). But if one does not possess the appropriate tools, multiplicatively non-specific quantities may be too complex to model in a multiplicative relationship. In this case, if the perceived need to consider quantities is high, as it apparently was in the diagram conditions, children may instead construct an additively specific model like "3 more" for the 4 and 7 glasses. Or, they might rely on a more abstract, qualitative schema that they had induced from prior occasions. Assuming that the children in question do not have a qualitative, multiplicative schema, they would probably rely on an additive one (e.g., *more-than (large(glass(A)), small(glass(B))))*.

Finally, and somewhat counter-intuitively, a qualitative small::large relation should be the most difficult to model. Sometimes, qualitative relations are more difficult than numerical ones (e.g., Heller et al., 1990). Imagine, for example, that x is larger than y, and one must decide whether the fraction x/y necessarily gets larger when the numerator and denominator both increase. An informal survey indicated that many people postulate specific numerical values and increments to infer the answer. One way to see why qualitative values may be more difficult to model is to compare a ratio representation and a qualitative representation of the value 4 when using a binary array with 10 cells. A ratio representation, where x means on and o means off, would be *xxxxoooooo*. In contrast, a qualitative version could be *xoxxoooxoo*, or *oooxxxooox*, or *oxoxxooxoo*, and so forth. Because the qualitative version does not necessarily specify ordinal, interval or ratio information, its representation has many different instantiations and much less available structure for building a coherent understanding. Consequently, modeling the relations between two different qualitative values may be very difficult. Detecting a lack of one-to-one correspondence, for example, would require more search of two qualitative arrays than two ratio arrays.

Sometimes people can compare qualitative values based on a perceptual experience of a manifest difference (e.g., dim, bright). But, as we developed in the Introduction, a perceptual comparison is not the same thing as a proportional understanding of a situation. Alternatively, one might rely on an additive schema to relate qualitative values. Or, an individual might postulate specific quantities to facilitate model construction, as in the case of reasoning about the fraction x/y above, or as in the case of saying that four-sevenths is close to one-half. But these are sophisticated strategies and would only result from a strong motivation to construct a quantitative mental model. More likely, without the pull of specific quantitaties, one would just rely on empirical knowledge, if it were readily available, to understand qualitative relationships.

Empirical knowledge joins into quantitative models by also providing information and constraints on mental model construction. One way this occurs is that empirical knowledge can ensure that some quantitative relationships remain invariant. For example, in the current studies, the children appeared to have tacit knowledge that the relationship between the concentrate and water would remain invariant through the pouring transformation. Empirical knowledge can also influence which quantities one chooses to measure or model. For example, the realistic displays led the children to think of juice primarily as a whole quantity rather than as a relationship between two sub-quantities. And, of course, empirical knowledge of the answer can constrain or usurp model construction altogether. All of these constraints are weak in the sense that a situation has many different empirical properties (as well as quantitative ones) that individuals may choose from as they attempt to develop a coherent model of a situation. Mathematical and empirical knowledge interact to determine which constraints are put into play in a quantitative mental model.

To develop our proposal with respect to the current data, we begin with the physical/photo conditions and first consider the students' model of the carton. According to the empirical constraint that juice be represented as a single entity, the students should not have constructed a quantitative model of the carton regardless of the available quantities; they did not have any ingredient quantities to relate. Combining this expectation with the proposed preference for quantitative specificity, we get a reasonable match to the data. For the non-numeric condition, the students avoided quantitative models and relied on "juice is juice" knowledge. For the hard glasses (4::7), the numerical relationship is multiplicatively difficult, so we would expect the children to either avoid making a model or rely on an additive schema. Our proposal does not provide clear guidance on which should have occurred. The results, however, showed that the children avoided making a model and answered same-taste, presumably because the physical/photo conditions called forth sufficient empirical knowledge that the students were not compelled to incorporate the hard numbers. Finally, for the multiplicatively specific 2::4 numbers, we should expect that students would model one glass as having twice as much juice as the other, because multiplicative specificity is easy (and inviting) to model.

In the diagram conditions, the students were less likely to think of juice as a single experiential entity. Nonetheless, they did heed the empirical constraint that the ratio of ingredients was invariant among the containers. As before, the non-numeric condition would not be expected to lead to a quantitative model, and it did not. In the presence of numbers, due to the analytic pull of the diagrams, we should expect the children to make more efforts to accommodate quantitative information than in the physical/photo conditions. With the hard numbers, we should expect the children to attempt additive relationships and answer different-taste. This occurred with the 24::40 carton and 4::7 glasses. Interestingly, it also occurred with the 20::20 carton and the 4::7 glasses. This latter result fits our account assuming that the children did not have adequate tools to easily model the half-half relationship in the 7 oz. glass, but did have tools to handle an additive, "3 more/less" relationship. Further research that manipulates or documents the children's level of mathematical competence would be useful in this regard. Finally, for the 20::20 carton and 2::4 glasses, the children should have been able to model multiplicative relationships all around and reach a proportional understanding. It is interesting to note that in this condition, the children modeled the most information: the ingredient quantities in the carton, the ingredient quantities in each glass, the invariance of the ingredient ratio through the pouring transformation, and the effect of ratio on taste. We can compare this to the physical 20::20 and 2::4 condition where the students thought of the quantity of juice in the two glasses and its effect on taste, or the photograph 20::20 and 4::7 condition where the students only thought about the fact that the juice came from the same carton. Evidently, mathematical operations can help the students to use relationships that increase the amount of information they can model.

Our description of quantitative mental models requires more evidence before it is useful to construct a predictive process model. One place to begin is with three of the central elements of our overall proposal: (a) models are constructed in working memory, (b) there are specificity constraints that help alleviate working memory demands, and (c) mathematical tools can aid in model construction by determining specific relationships and thereby reduce working memory demands. One way to evaluate the role of specificty is to provide people with a variety of scenarios that include specific and non-specific quantities. On a later memory test, people should recall the gist of multiplicatively specific relationships because they can construct meaningful models (Bransford, Barclay, & Franks, 1972), whereas they should recall non-specific quantitative relationships verbatim, if at all (Mani & Johnson-Laird, 1982). This type of experiment could include various types of specific relationships (e.g., 3::3, 3::9, 3::12) to see whether the specificity claim can extend beyond the halving and doubling relations used in the current studies. Using the same paradigm, one might explore the role of mathematical knowledge by using more sophisticated relationships and manipulating prior knowledge. For example, imagine adults read a story about population growth. If they know about different types of exponential functions, then the use of specific and non-specific quantities should have more of an effect on their recall than on the recall of adults who do not know anything about exponential growth. With respect to working memory, it is important to get a sense of the number of models that people are constructing. For example, is it better to conceive of the students as making one model for the juice problem, or two models—one for the carton and one for the glasses. In our proposal, working memory demand decreases as one discovers specific symmetries (i.e., co-specifications), whereas in the mental model theory of deduction, working memory demand decreases with fewer models (Johnson-Laird, 1983). It is not clear whether specificity and the number of models are separable dimensions of complexity. One way to examine this question is to provide a series of problems that systematically vary the number and specificity of possible models. A nice candidate for use with adults is a problem involving two containers of paint. In phase one, 1 oz. from a 10 oz. container of red paint is thoroughly mixed into a 10 oz. container of white paint. In phase two, 1 oz. is taken from the modified white container and mixed into the red container. Will the two containers have inverse ratios of red to white paint? Informal observations indicate that many adults incorrectly believe that the two containers are not inverse ratios. They do not bother to construct a multiplicative model of the 1 oz. of paint that is moved during phase two. Yet, at the same time, it seems possible that solving this problem, or even understanding the explanation, would be more difficult if the quantities were non-numerical (non-specific).

Developmental and Instructional Possibilities

The sketch of quantitative mental models suggests an explanation for the development of proportional reasoning that does not assume a developmental increase in working memory or the growth of fundamentally new conceptual structures. Perhaps mathematical experiences lead to the development of mathematical tools. These tools enable children to simplify phenomena so that they can encode them within the working memory constraints of basic model structures. The challenge of proportional reasoning may not be found so much in a lack of conceptual structure or an immature working memory as it is found in the empirical complexity of proportional situations. Mathematical tools help simplify this empirical complexity so that it may be understood within basic model structures. From our perspective, when children learn about percents, they do not develop a fundamentally new part-whole competence. Instead, they learn a cultural tool that can help them turn a complex situation into a multiplicatively specific one. Fractions, decimals, division, estimation, LCM, and so forth, give children a repertoire of mathematical tools that enables them to model the world into basic commensurable structures.

The overall sketch of quantitative models generates some suggestions for improving instruction. Consider the familiar lament that children and adults, particularly in the context of word problems, disregard empirical constraints and resort to symbol pushing (e.g., Reusser, 1988; Silver, 1986). How might we overcome this problem in educational settings?

First, given students' limited repertoire of formal mathematical tools and the automaticity with which they can apply those tools, the numbers in problems are often too difficult to incorporate into a quantitative model. As a result, the student can either rely on remembered associations between mathematical operations and problem situations (cf. Moore, Dixon, & Haines, 1991) or haphazardly guess which symbol manipulation is most appropriate. The numbers and the problem situation do not constrain one another at a structural level because the quantities are too complex to incorporate into a model that includes empirical information as well. Thus, our first suggestion is that instruction might rely more on simple numbers, or might continually help students map complex numbers and operations into simpler ones.

Second, instruction often treats mathematics as a computational recipe for "solving" an already understood situation. Statistical instruction is a notorious example. Students are often taught statistics as a general way to test hypotheses about the world, and rarely are they taught statistics as a way to understand and model structures in the world. Given the typical emphasis on the computational side of mathematics, it is little wonder that students do not always use mathematics to make sense of a situation. Instruction may be more effective if students also learn about mathematics as a tool for structuring specific models of the world (e.g., Bransford et al., in press). Students should get a chance to see how such tools help solve problems of cognitive complexity and how they allow them to understand new aspects of their world.

For example, imagine that we give college students two sets of numbers (e.g., 2-4-6-8 vs. 4-5-5-6). We point out that the two sets have a similarity. We ask the students to notice that there is a single number for each set that helps determine this similarity—the average. This single number is easier to keep in mind and communicate than the total distribution. We then ask the students to come up with a method for determining a single number for each set that can capture what is different (i.e., the variances). After they invent their own methods, often a range formula like subtract the smallest number from the largest , we provide new numbers that highlight other properties of distributions against which they can test their inventions (e.g., 0-2-4-6-8 versus 0-8-8-8-8, or 0-2-4-6-8 versus 2-6). After several cycles of invention, testing, and revision, we ultimately provide the students with the conventional approach (e.g., standard deviation). Lessons like this help students "see the point" of the variance formula while they also help students differentiate the components of the formula and its quantitative referents (cf. Schwartz & Bransford, in press). Seeing the power of a good formalism for aiding with complexity issues makes students more likely to use that representational tool in the future (Schwartz, 1993). Moreover, such lessons help students notice the empirical properties that the tools were designed to help model. Preliminary results indicate that students who learn variance through this "semiotic invention to cultural convention" cycle are more likely to notice variance-related issues in real-world tasks like evaluating potential employees. If they hold up, these results would fit our overall story. Although mathematics is general and indifferent to empirical situations, it is the ally of the sciences because it provides powerful tools for finding, modeling, and understanding relationships in the material world.

Acknowledgments: We would like to thank the teachers, students, and parents who permitted us to conduct this research. We would also like to thank Graeme Halford and an anonymous reviewer for helpful suggestions and pointers to the literature. Finally, we would like to thank Lisa Oakes for comments on an earlier draft and Howard Sandler for being such a structuralist.

REFERENCES

Ahl, V. A., Moore, C. F., & Dixon, J. A. (1992). Development of intuitive and numerical proportional reasoning. *Cognitive Development, 7*, 81-108.

Baranes, R., Perry, M., & Stigler, J. W. (1989). Activation of real-world knowledge in the solution of word problems. *Cognition and Instruction, 6*, 287-318.

Bassok, M., & Holyoak, K. J. (1989). Interdomain transfer between isomorphic topics in algebra and physics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 153-166.

Bassok, M., & Olseth, K. L. (1995). Object-based representations: Transfer between cases of continuous and discrete models of change. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1522-1538.

Ben-Zeev, T. (1995). The nature and origin of *Rational Errors* in arithmetic thinking: Induction from examples and prior examples. *Cognitive Science, 19*, 341-376.

Bowers, J., Cobb, P., & McClain, K. (in press). The evolution of mathematical practices. *Cognition & Instruction.*

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology, 3*, 193-209.

Bransford, J. D., Zech, L., Schwartz, D. L., Barron, B. J., Vye, N. J., & the Cognition and Technology Group at Vanderbilt. (in press). Design principles for the development of environments that invite and sustain mathematical thinking. To appear in P. Cobb (Ed.), *Symbolizing, Communicating, and Mathematizing: Pespectives on Discourse, Tools, and Instructional Design.* Mahwah, NJ: Erlbaum.

Briars, D. J., & Siegler, R. S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology, 20*, 607-618.

Bruner, J. S., Oliver, R. R., & Greenfield, P. M. (1966) *Studies in cognitive growth.* New York: John Wiley.

Bryant, P. E., & Kopytynska, H. (1976). Spontaneous measurement by young children. *Nature, 260*, 773.

Carpenter, T. P. (1975). Measurement concepts of first- and second-grade students. *Journal of Research in Mathematics Education, 6*, 3-14.

Carpenter, T. P., Fennema, E., & Romberg, T. A. (Eds.) (1993). *Rational numbers: An integration of research. Studies in mathematical thinking and learning.* Hillsdale, NJ: Erlbaum.

Case, R. (1978). Intellectual development from birth to adulthood: A neo-Piagetian interpretation. In R. S. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, NJ: Erlbaum.

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, Serial No. 246).

Changeux, J. -P., & Connes, A. (1995). *Conversations on mind, matter, and mathematics.* (M. B. DeBevoise, Trans.). Princeton, NJ: Princeton University Press. (Original work published 1989.)

Chomsky, N. (1980). *Cartesian linguistics: A chapter in the history of rationalist thought.* New York: Harper & Row.

Cobb, P., Yackel, E., & Wood, T. (1992). Interaction and learning in mathematics classroom situations. *Educational Studies in Mathematics, 23,* 99-122.

Cognition and Technology Group at Vanderbilt. (1997). *The Jasper project: Lessons in curriculum, instruction, assessment, and professional development.* Mahwah, NJ: Erlbaum.

Cramer, K., Post, T., & Currier, S. (1993). Learning and teaching ratio and proportion: Research implications. In D. T. Owens (Ed.), *Research ideas for the classroom: Middle grades mathematics.* New York: Macmillan.

Cummins, D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology, 20,* 405-438.

Davies, P. (1992). *The mind of God.* New York, NY: Simon & Schuster.

DeLoache, J. S., & Marzolf, D. P. (1992). When a picture is not worth a thousand words: Young children's understanding of pictures and models. *Cognitive Development, 7,* 317-329.

Dixon, J. A., & Moore, C. F. (1996). The developmental role of intuitive principles in choosing mathematical strategies. *Developmental Psychology, 32,* 241-253.

English, L. D., & Halford, G. S. (1995). *Mathematics education: Models and processes.* Mahwah, NJ: Erlbaum.

Flavell, J. H. (1972). An analysis of cognitive developmental sequences. *Genetic Psychology Monographs, 86,* 279-350.

Feynman, R. (1965). *The character of physical law.* Cambridge, MA: MIT Press.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology.* Cambridge, MA: MIT Press.

Gelman, R. (1982). Accessing one-to-one correspondence: still another paper about conservation. *British Journal of Psychology, 73,* 209-220.

Gelman, R., & Brenneman, K. (1994). First principles that can support both universal and culture-specific learning about number and music. In L. A. Hirschfeld & R. A. Gelman (Eds.), *Mapping the mind.* Cambridge, MA: Cambridge University Press.

Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. *Cognitive Science, 7,* 155-170.

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

Greeno, J. G. (1989). Situation models, mental models, and generative knowledge. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert Simon.* Hillsdale, NJ: Erlbaum.

Greeno, J. G. (1991). Number sense as situated knowing in a conceptual domain. *Journal for Research in Mathematics Education, 22,* 170-218.

Greeno, J. G., Smith, D. R., & Moore, J. L. (1993). Transfer of situated learning. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, Cognition, and Instruction.* Norwood, NJ: Ablex.

Gopnik, A. (1996). The post-Piagetian era. *Psychological Science, 7,* 221-225.

Halford, G. S. (1993). *Children's understanding: The development of mental models.* Hillsdale, NJ: Erlbaum.

Halford, G. S., Wilson, W. H., & Phillips, S. (in press). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences.*

Hardiman, P. T., Well, A. D., & Pollastek, A. (1984). Usefulness of a balance model in understanding the mean. *Journal of Educational Psychology, 5,* 792-801.

Harel, G., Behr, M., Lesh, R., & Post, T. (1994). Invariance of ratio: The case of children's anticipatory scheme of constancy of taste. *Journal for Research in Mathematics Education, 25,* 324-345.

Hart, K. (Ed.). (1981). *Children's understanding of mathematics.* London: John Murray.

Hatano, G., & Osawa, K. (1983). Digit memory of grand experts in abacus-driven mental calculation. *Cognition, 15,* 95-110.

Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology, 87,* 18-32.

Heller, P. M., Post, T. R., Behr, M., & Lesh, R. (1990). Qualitative and numerical reasoning about fractions and rates by seventh-and eighth-grade students. *Journal for Research in Mathematics Education, 21,* 388-402.

Hinsley, D., Hayes, J. R., & Simon, H. A. (1977). From words to equations. In P. Carpenter & M. Just (Eds.), *Cognitive processes in comprehension.* Hillsdale, NJ: Erlbaum.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning and discovery.* Cambridge, MA: The MIT Press.

Huttenlocher, J., Jordan, N. C., & Levine, S. C. (1994). A mental model for early arithmetic. *Journal of Experimental Psychology: General, 123*, 284-296.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness.* Cambridge, MA: Cambridge University Press.

Karmiloff-Smith, A. (1979). Micro- and macro-developmental changes in language acquisition and other representational systems. *Cognitive Science, 3*, 91-118.

Karplus, R. (1981). Education and formal thought: a modest proposal. In I. E. Siegel, D. M. Brodzinsky, & R. M. Golinkoff (Eds.), *New directions in Piagetian Theory and Practice.* Hillsdale, NJ: Erlbaum.

Karplus, R., Pulos, S., & Stage, E. (1983). Proportional reasoning of early adolescents. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes.* New York: Academic Press.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*, 109-129.

Klahr, D., & Wallace, J. G. (1976). *Cognitive development: An information processing view.* Hillsdale, NJ: Erlbaum.

Lave, J. (1988). *Cognition in practice.* New York: Cambridge University Press.

Lesh, R., & Lamon, S. (1992). Assessing authentic mathematical performance. In R. Lesh & S. Lamon (Eds.), *Assessment of authentic performance in school mathematics.* Washington, DC: American Association for the Advancement of Science.

Lesh, R., Post, T., & Behr, M. (1988). Proportional reasoning. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in the middle grades.* Hillsdale, NJ: Erlbaum.

Light, P., & Gilmour, A. (1983). Conservation or conversation? Contextual facilitation of inappropriate conservation judgments. *Journal of Experimental Child Psychology, 36*, 356-363.

Lovell, K., & Butterworth, I. B. (1966). Abilities underlying the understanding of proportionality. *Mathematics Teaching, 37*, 5-9.

Lunzer, E. A., & Pumfrey, P. D. (1966). Understanding proportionality. *Mathematics Teaching, 34*, 7-12.

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition, 10*, 181-187.

McGarrigle, J., & Donaldson, M. (1975). Conservation accidents. *Cognition, 3*, 341-350.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9*, 320-334.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*, 254-278.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244*, 933-938.

Moore, C. F., Dixon, J. A., & Haines, B. A. (1991). Components of understanding in proportional reasoning: A fuzzy set representation of developmental progressions. *Child Development, 62*, 441-459.

Moore, J. L. (1993). *Comparisons of a physical model and computer representation in reasoning and learning about learning functions.* Unpublished doctoral dissertation. Stanford University, Palo Alto.

Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science, 238*, 625-630.

Noelting, G. (1980a). The development of proportional reasoning and the ratio concept: I. Differentiation of stages. *Educational Studies in Mathematics, 11*, 217-253.

Noelting, G. (1980b). The development of proportional reasoning and the ratio concept: Part II—Problem structure at successive stages; problem solving strategies and the mechanism of adaptive restructuring. *Educational Studies in Mathematics, 11*, 331-363.

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

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. *Human Development, 15*, 1-12.

Piaget, J., Grize, J. Szeminska, A., & Bang, V. (1977). *The psychology and epistemology of functions.* (First published in French, 1968.)

Putnam, R. T., Lampert, M., & Peterson, P. L. (1990). Alternative perspectives on knowing mathematics in the elementary schools. *Review of Research in Education, 16*, 57-150.

Reed, S. K. (1993). A schema-based theory of transfer. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction.* Norwood, NJ: Ablex.

Resnick, L. B. (1992). From protoquantities to operators: Building mathematical competence on a foundation of everyday knowledge. In G. Leinhardt, R. T. Putnam, & R. Hattrup (Eds.), *Analysis of arithmetic for mathematics teaching*. Hillsdale, NJ: Erlbaum.

Reusser, K. (1988). Problem solving beyond the logic of things: Contextual effects on understanding and solving word problems. *Instructional Science, 17*, 309-338.

Roazzi, A., & Bryant, P. (1993). Social class, context and cognitive development. In P. Light & G. Butterworth (Eds.), *Context and Cognition*. Hillsdale, NJ: Erlbaum.

Ross, B. H. (1989). Remindings in learning and instruction. In S. Vosniadou & A.Ortony (Eds.), *Similarity and analogical reasoning*. New York: Cambridge University Press.

Samuel, J., & Bryant, P. (1984). Asking only one question in the conservation experiment. *Journal of Child Psychology and Psychiatry, 25*, 315-318.

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology, 17*, 36-71.

Schwartz, D. L. (1993). The construction and analogical transfer of symbolic visualizations. *Journal of Research in Science Teaching, 30*, 1309-1325.

Schwartz, D. L. (1995). Reasoning about the referent of a picture versus reasoning about the picture as the referent: An effect of visual realism. *Memory & Cognition, 23*, 709-722.

Schwartz, D. L., & Black, J. B. (1996a). Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology, 30*, 154-219.

Schwartz, D. L., & Black, J. B. (1996b). Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science, 20*, 457-497.

Schwartz, D. L., & Bransford, J. D. (in press). A time for telling. *Cognition & Instruction.*

Schwartz, D. L., & Hegarty, M. (1996). Coordinating multiple representations for reasoning about mechanical devices. In P. L. Olivier (Ed.), *Cognitive and computational models of spatial representations*. Stanford, CA: AAAI Press.

Sherin, B. (1996). *The symbolic basis of physical intuition: A study of two symbol systems in physics instruction.* Unpublished doctoral dissertation. University of California, Berkeley.

Siegler, R. S. (1981). Developmental sequences within and between concepts. *Monographs of the Society for Research in Child Development, 46* (2, Serial No. 189).

Silver, E. A. (1986). Using conceptual and procedural knowledge: A focus on relationships. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics*. Hillsdale, NJ: Erlbaum.

Spellman, B. A., & Holyoak, K. J. (1996). Pragmatics in analogical mapping. *Cognitive Psychology, 31*, 307-346.

Spinillo, A. G., & Bryant, P. (1991). Children's proportional judgments: The importance of "Half." *Child Development, 62*, 427-440.

Vosniadou, S., & Brewer, W. F. (1993). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24*, 535-585.

Winer, G. A. (1974). Conservation of different quantities among preschool children. *Child Development, 45*, 839-842.

Wynn, K. (1992). Addition and subtraction by human infants. Evidence against empirical accounts of the origins of numerical knowledge. *Mind and Language, 7*, 315-332.

Zimiles, H. (1966). The development of conservation and differentiation of number. *Monographs of the Society for Research in Child Development, 31* (Whole No. 108).