

a. *Learning of labels.* Children in the training group did not require many trial blocks to learn the labels. The mean number of trial blocks before the block of last error was 2.2 (excluding one child who never met the mastery criterion and was replaced by another child). Among the 10 children in the training group, 9 did not make an error after the third trial block.

b. *Magnitude comparisons.* The training in labeling greatly reduced the number of errors that the children made. Children who were trained to apply the labels were correct on 96% of the magnitude comparisons, versus 79% correct answers among children in the control group.

The labeling training also changed the distribution of children's errors. Children in the trained group made 65% of their errors on the 9 items involving within-cluster comparisons. In contrast, children in the control group made only 29% of their errors on the same 9 items.

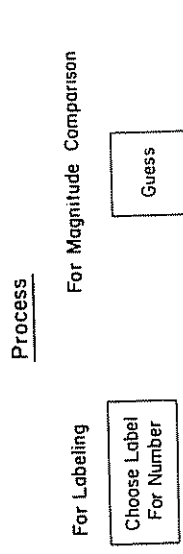
E. THREE MODELS OF NUMERICAL MAGNITUDE COMPARISON

The three experiments described above provided converging evidence that numerical categories play an important role in preschoolers' numerical magnitude comparisons. Multidimensional scaling and hierarchical clustering analyses of preschoolers' untutored magnitude comparisons revealed clusters of numbers rather than a regular distribution. The children's verbal labeling of numbers predicted both their level of skill in performing the magnitude comparisons and the relative difficulty that each magnitude comparison problem posed for them. Teaching the children an overall categorical organization reduced their number of errors on the numerical comparison task and changed their distribution of errors on it as well. Thus, it appeared that any model of preschoolers' numerical magnitude comparisons should assign an important place to the children's categorizations of numbers.

Figure 5 depicts four such categorical models that would generate magnitude comparison performance of varying degrees of proficiency. First, let us consider Model I (Fig. 5A), which would produce the approximately chance level performance that we observed in most of the 3-year-olds. Although these children's magnitude comparisons did not indicate any knowledge of numerical magnitudes, other data indicated that they did possess some relevant information. In particular, their classifications of numbers as little, medium, and big in the

A. Representation
(Empirically derived probabilities of each number being assigned to each category)

Category	1	2	3	4	5	6	7	8	9
Smallest	.6								
Small	.2	.25	.42	.08	.25	.08	.17		.08
Medium	.2	.58	.42	.92	.58	.67	.50	.67	.67
Big	.17	.17	.17	.17	.17	.25	.33	.33	.25



B. Representation

Category	1	2	3	4	5	6	7	8	9
Smallest	.7								
Small	.2	.72	.61	.22	.06	.06			.06
Medium	.1	.28	.33	.67	.72	.67	.56	.27	.44
Big		.06	.11	.22	.33	.44	.67	.56	.56

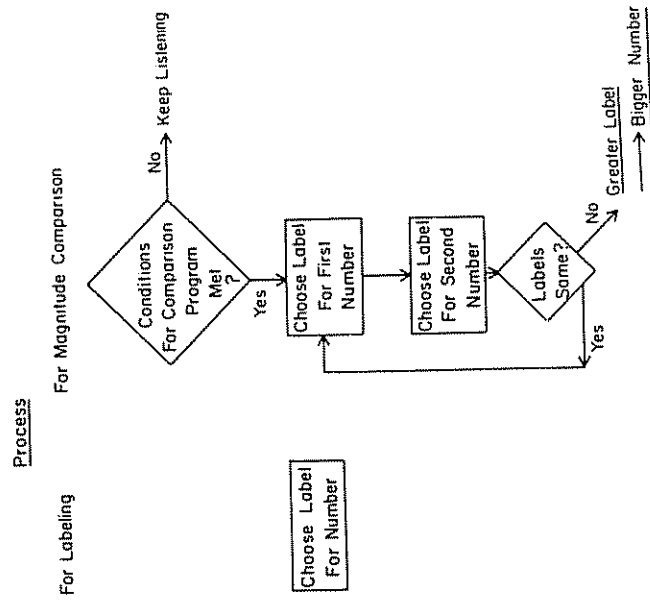


Fig. 5. Models of magnitude comparison. (A) Model I. (Note: Here and in Fig. 5B, the probabilities for the numbers 2-9 are the group level probabilities that children in the labeling experiment (3-year-olds in Fig. 5A, 4-year-olds in Fig. 5B) assigned each label to each number. The Fig. 4 scaling and clustering results suggested that the number 1 was in a separate "smallest" category not tapped by the labeling procedure; therefore, hypothetical probabilities have been assigned to the number 1.) (B) Model II. See pp. 282 and 283 for Fig. 5C and D.

C. Representation

Categorical Organization	Probability	Number
A	.5	Smallest Small Medium Big
B	.2	Smallest Small Medium Big
C	.1	Smallest Small Medium Big
D	.1	Smallest Small Medium Big
E	.1	Smallest Small Medium Big

D. Representation

Categorical Organization	Probability	Number
A	1.0	1 2 3 4 5 6 7 8 9
Smallest	.7	.72 .61 .22 .06
Small	.2	.28 .33 .67 .56 .27 .44
Medium	.06	.11 .22 .33 .44 .67 .56
Big		

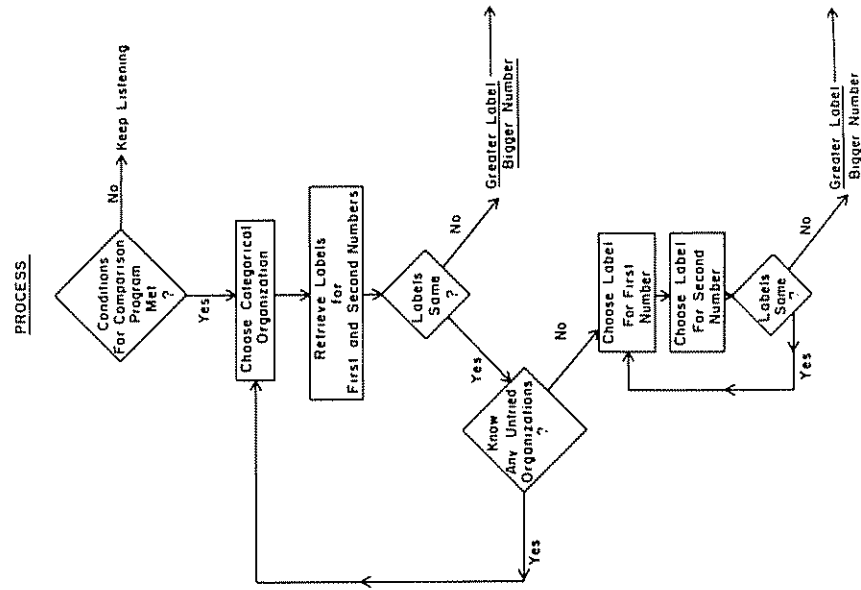
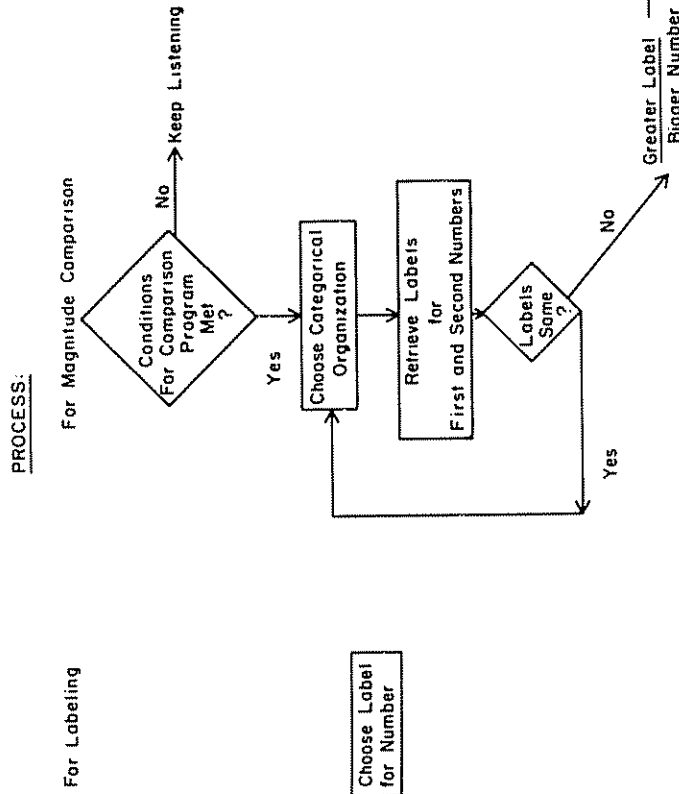


Fig. 5. (continued) (C) Model III. (Note: The probabilities of each number being assigned to each label were chosen to correspond roughly to the labeling experiment probabilities that the 5-year-olds assigned each label to each number. An exact matching was impossible, because in the labeling experiment, the children were asked to assign labels to numbers independently, but within this model, category boundaries are dependent upon each other.)

Fig. 5. (continued). (D) Model II-III.

labeling experiment corresponded reasonably well to the relative magnitudes of the numbers.

The representation of the 3-year-olds' knowledge (Fig. 5A) is based directly on their performance in the labeling experiment. Viewed vertically, this matrix indicates the probability that each number will be assigned each categorical label (in the context of the numbers 1-9). Viewed horizontally, the matrix indicates each category's range and the probability that each category will include each number. When asked whether each number is small, medium, or large, children simply retrieve categorical labels with the probabilities indicated in the representation. When they are asked to compare magnitudes, however, children do not utilize any of the knowledge from the representation; they simply guess, thus producing chance-level magnitude comparison performance.

Next consider Model II, hypothesized to underlie the moderately expert performance (60-95% correct) of the majority of 4-year-olds (Fig. 5B). Model II incorporates one major change from Model I: the Model II magnitude comparison process makes use of the categorical information. As shown in Fig. 5B, when children are presented a magnitude comparison problem, they generate a label for each number. The choices of labels are independent for the two numbers, and occur for each number with the probabilities shown by looking down the columns of the representation. If the labels differ, children choose as bigger the number associated with the larger label. If the labels are identical, children regenerate labels for each number until the labels discriminate between them. Note that, regardless of the particular probabilities assigned to the labels, this type of model almost inevitably leads to errors. Unless the child creates at least ($N - 1$) distinct categories for the N numbers in the comparison set, he or she will always have some probability of assigning the larger label to the smaller number.

Model III, intended to characterize the near-perfect performance of most of the 5-year-olds, is quite similar to Banks, Fujii, and Kayra-Stuart's (1976) model of adult performance. It also represents a logical extension of the Fig. 5B model of moderately skilled performance. The only large change from that model is that the choice of labels on the magnitude comparison task is not at the level of the individual numbers, but rather at the level of the overall categorical organization. As shown in Fig. 5C, the child's representation includes several alternative divisions of the 9 numbers into categories. The process indicates that each time a magnitude comparison problem is presented, the child chooses one of the categorical organizations with the probability shown in the representation. Once a particular organization is chosen, the assignment of labels to numbers is determined. If the labels differ, the child chooses the number attached to the larger label as being bigger. If the labels are identical, the child chooses a new categorical organization. The process continues until a categorical organization is chosen that assigns different labels to the two numbers being compared.

Finally, Model II-III, a hybrid of Models II and III, is hypothesized to produce the performance of the trained children in the training experiment. This model is diagramed in Fig. 5D. The representation includes the Fig. 5B model probabilities of each number being assigned each label, and also the one categorical organization that the children were taught in the training procedure. The process starts with the Model III approach in which a categorical organization is selected. Because the only categorical organization that the children know is the one they were taught, they use it to assign labels to the numbers. If this procedure assigns different labels to the numbers, the children answer correctly. If it assigns the same labels, however, they would not know any alternative organizations, and therefore would revert to the Model II process of assigning labels to each number independently until the labels differed.

These models give rise to the types of magnitude comparison performance, labeling, and learning that were observed among each age group in each of the three above-described experiments. Model I prescribes the 3-year-olds' behavior in a rather direct way. As implied by the probabilities given in Fig. 5A, these children can assign labels to the numbers that have some correspondence to their magnitudes. However, they do not use this knowledge in comparing the magnitudes; instead they simply guess, producing on all 36 problems the roughly chance performance that was observed.

Model II predicts a considerably lower error rate and a distribution of errors that is linked to the characteristics of the problems. Averaged across all 36 problems, the model generates 84% rather than 50% correct answers. The 84% correct figure is quite close to the 4-year-olds' observed percentage correct, 81%. The symbolic distance effect emerges because numbers that are farther apart are less likely to be assigned incorrectly ordered labels. The main effect emerges for much the same reason; the distribution of labels changes more rapidly between numbers at the small end of the scale than between numbers at the large end. The labeling experiment results of accurate labeling of numbers as small, medium, and large, would arise from children choosing labels for the numbers in accord with the probabilities given in Fig. 5B.

Model III predicts that no errors should occur, because within each categorical organization, the category boundaries are correctly ordered and nonoverlapping. The effects of categorization will continue to be seen in solution time patterns, however. Between-category comparisons will on average arise earlier in the comparison process when the digits are farther apart and when the magnitude of the minimum number is small. Sekuler and Mierkiewicz's (1977) data on 6-year-olds and older children are in accord with these predictions.

Finally, Model II-III, the model of what the trained children learned, would produce performance similar to Model III on the between-category comparisons and performance similar to Model II on the within-category ones. Overall, it generates 91% correct performance, an accuracy level quite close to the 96% that

was observed among the trained children. The superior performance predicted by this model on, and only on, the between-category problems explains both why trained children made fewer errors overall than untrained ones, and why the bulk of the few errors they made were within-category errors.

In addition to accounting for our data and providing a plausible outline of the development of numerical comparison skills, these models raise two more general questions. First, is the categorization approach generally used in comparing magnitudes across different age groups and across different types of materials, or is it limited to young children's numerical magnitude comparisons? Second, what are the developmental implications of the parallels, observed here and in many previous studies, between early error patterns and later solution times?

The first question concerns the generality of the models across age groups and across types of stimulus materials. Several recent experiments have demonstrated categorical effects in populations other than preschoolers and in domains other than numbers. Kosslyn, Murphy, Bemesderfer, and Feinstein (1977) obtained categorical effects with adults for highly overlearned, artificially imposed categories in the context of length judgments. Pliske and Smith (1979) found that adults spontaneously used gender as a basis of categorization in making distance judgments in a situation in which gender divided the lengths into nonoverlapping categories. Maki (1981) provided a similar demonstration of adults using state boundaries as a basis for categorizing cities as being farther east or west. Thus, the division of stimuli into categories for purposes of comparing magnitudes is far from unique to preschoolers as a population or to numbers as a type of stimulus material. It does seem likely, though, that children (and probably adults) who are in the process of learning new dimensional orderings will be especially likely to categorize stimuli as a means of reducing imposing learning tasks to manageable dimensions (as in the adage, "Divide and conquer"). Under such circumstances, differing degrees of information about the individual stimuli might be expected to dictate the form of the categories. Stimuli associated with numerous referents (e.g., the small numbers in the experiment presented here) will be placed in single member or small categories, but stimuli about which little is known (e.g., the larger numbers here) will be lumped into multimember units.

The second question concerns the parallels that arose between young children's error patterns and the previously reported solution times of older children and adults on the same numerical comparison task. This parallel is far from unique to the research presented here; data on such diverse cognitive skills as analogical reasoning, transitive inference, quantification, and attention have shown similar parallels (Chi & Klahr, 1975; Manis, Keating, & Morrison, 1980; Sternberg, 1977; Trabasso, Riley, & Wilson, 1975). The present Models II and III account for the similarities in the numerical comparison context and by analogy suggest why they may appear in others. Examination of Fig. 5B shows that the representation of the problem dictates the distribution of early errors. With learn-

ing and development, increasingly task-appropriate processes are adopted and greatly reduce or eliminate errors (Fig. 5C). However, the representation that these processes operate upon remains relatively unchanged in the sense that the same numbers usually receive the same labels. This leads to relatively long solution times on those problems that earlier produced errors. We do not know at present whether representations of stimuli generally remain more stable over development than the processes that operate upon them; Newell (1972) pointed out difficulties in even posing the question. Nonetheless, relatively stable representations and relatively rapidly changing processes provide at least one explanation for the parallels between early errors and later solution times.

V. Preschoolers' Knowledge of Addition

A. EXISTING RESEARCH ON CHILDREN

Young children's knowledge of addition has been the subject of far more research than their skill in counting or in comparing numbers. No doubt, this large body of work is due in large part to the traditional centrality of addition in the elementary school curriculum and to its importance in everyday life. Two primary questions have emerged within the research: what addition tasks can young children perform, and what processes do they use to perform them?

1. *What Young Children Know about Addition*

Much early work on addition was devoted to determining which problems young children find easy to solve and which they find difficult. One of the most elaborately reported examples of such a study was that of Knight and Behrens (1928). These researchers presented second graders the 100 addition problems formed by the factorial combination of augend (0-9) plus addend (0-9). They reported that the greater the sum, the more difficult the problem, and also that ties were easier than would have been expected from considering their sums alone. When we reanalyzed the rankings of Knight and Behrens using regression analyses, we found that the size of the smaller number was also predictive of problem difficulty. It accounted for exactly as much of the variance as the sum when all problems were considered (62%) and more of the variance when only the nontie problems were included (81% vs 67%).

More recently, interest has shifted to very young children's knowledge of the principles underlying addition. Smedslund (1966) demonstrated that by 5 or 6 years of age, children possess the most basic knowledge about addition: that adding increases quantity. Subsequent investigations (e.g., Gelman, 1972; Gelman & Starkey, 1981; Siegler, 1981a) indicated that children as young as 2 and 3 years possess similar knowledge, at least when the augend and addend sizes are

small. Some understanding of the principles of inversion and compensation as they apply to addition and subtraction also seems to be present even before children enter school (Cooper, Starkey, Blevins, Goth, & Leitner, 1978; Gelman & Starkey, 1981).

2. *How Children Solve Addition Problems*

Another issue that has long been of interest is how young children solve addition problems. For many years, lack of revealing methodologies limited insights into these processes. However, the development of chronometric methods proved to be an important breakthrough.

Perhaps the best-known model of the addition process, the min model of Groen and Parkman (1972), is based upon a chronometric approach. Groen and Parkman started with the question of whether material as overlearned as addition facts was retrieved from memory directly or whether answers were reconstructed each time problems were presented. To address this issue, they examined first graders' patterns of solution times for all integer addition problems with sums ≤ 9 , and also adults' patterns of solution times for all integer addition problems with augends and addends ≤ 9 . In each case, they found that solution times were directly proportional to the size of the minimum number.

On the basis of this evidence they formulated the min model. Within this model, the adder chooses the larger of the two numbers and then increments it by one a number of times equal to the smaller number. The amount of time required to choose the larger number is assumed to be constant for all problems, as is the time per increment. Therefore, the only factor contributing to differences among problems in solution times is the number of increments dictated by the minimum number.

The min model fit Groen and Parkman's (1972) data quite well except on ties, where both children and adults were much faster than would have been predicted. Also unexpectedly, the solution times on ties were nearly constant over all of the items that were tested. These data led Groen and Parkman to amend their model so that ties were reproduced directly from "fast-access memory," while other problems were reconstructed by the incrementing process.

Subsequent experiments have provided a mixed record of support and non-support for Groen and Parkman's model as it applies to children. On the positive side, other investigators have replicated the findings that the size of the minimum number is the best predictor of young children's solution times and that ties are solved uniformly faster than might be expected from the sizes of their minima (Svenson, 1975; Svenson & Broquist, 1975). Verbal explanations have provided converging evidence: when Svenson (1975) asked children how they solved specific problems, they often explained that they chose the larger number and counted upward from it (Svenson, 1975). Even preschoolers who were taught addition by a different method often later produced solution times proportional to the minimum number (Groen & Resnick, 1977).

Not all of the evidence with children has been consistent with the min model, however. The size of the last number seems to play a role above and beyond that of the size of the minimum number (see Groen & Parkman, 1972, Fig. 2; Svenson, 1975). The same children whose verbal statements indicated awareness of the min strategy also indicated reliance on a variety of more specific pieces of knowledge; they said they used ties and the number 10 as reference points from which to solve other problems (Svenson, 1975). Finally, close scrutiny of Groen and Parkman's (1972) data indicates that their model fit much better at low values of minima than at higher values; the correlation between the size of the minimum and the mean solution time was $r = .82$ for minimum numbers of sizes 0-2, while it was only $r = .11$ for minimum numbers of sizes 2-4 (our reanalysis of all cases; if ties are excluded, the correlations are $r = .87$ and $r = .48$, respectively). These findings suggest that attributes of particular problems, other than their minimum numbers, influence children's addition.

Because the subsequent data have been only partially favorable to Groen and Parkman's model, several investigators have proposed alternative models. Svenson (1975) proposed a model basically similar to that of Groen and Parkman except that children would take time to reorder the numbers if the larger number were second; this reordering operation accounted for the effects exercised by the second number independent of those of the minimum number. The model of Ashcraft and Battaglia (1978) represented a more radical departure. It accounted for differences among the solution times of various problems in terms of search and retrieval processes rather than in terms of any reconstructive process.

All of these models of how addition is done are based primarily upon chronometric data; another source of evidence about the addition process comes from clinical descriptions. A study of Ilg and Ames (1951) is perhaps the most comprehensive of these clinical studies. On the basis of observations of very large numbers of beginning adders, Ilg and Ames hypothesized a four-stage sequence for the development of addition skills. First, children were said to perform all addition problems by counting on from 1. Slightly later, they were said to know a few addition facts "by heart" and to count on from the smaller number to compute the others. Still later, they were said to count on from the larger number (as in the min model) to solve those problems they had not memorized. Finally, they were said to have memorized many problems and to use a variety of specific strategies to solve the others. For example, they might break up $14 + 3$ into $4 + 3 = 7$; $7 + 10 = 17$.

This last observation of specific strategy use has been echoed in numerous other clinical descriptions of young children's addition. Several approaches have been noted: children have been said to rely on 5s, 10s, and ties as reference points from which to calculate answers, to move their feet rhythmically to help them count, to count on from the larger number, and to count on from 1 (Fuson & Richards, 1982; Hebbeler, 1976; Yoshimura, 1974). Although these descriptions of addition strategies have been largely anecdotal, the reports have been persis-

tent enough to leave little doubt that young children employ a variety of methods in adding.

Considered as a group, these studies reveal a cleavage between the types of models that have been formulated and many of the phenomena that have been observed. On the one hand, the models of addition—Groen and Parkman (1972), Ashcraft and Battaglia (1978), and Svenson (1975)—have all been designed to depict *the* strategy people use to solve addition problems. On the other hand, the detailed observations of addition have revealed a large number of distinct strategies that are used on particular problems. The discrepancy suggests that the chromometric data on children's addition may reflect an averaging over different strategies rather than a consistent adherence to any one strategy; this would call into question any single model that purported to be *the* way in which children solve all addition problems (cf. Estes, 1956; Newell, 1972). The existence of numerous distinct addition procedures would also suggest that, above and beyond determining how children execute any particular addition strategy, we would also need to learn how they choose among alternative approaches.

B. PRESCHOOLERS' ADDITION STRATEGIES

Considering the possibility that children use different addition strategies on different problems raises a large number of questions that have not been addressed previously. Which strategies are used most frequently? What are their accuracy and temporal characteristics? Are choices of strategies systematically related to the particular numbers involved in each problem? Does variable strategy use help children add more accurately and/or more quickly than using the same strategy at all times? If so, why?

Videotaping children in the process of adding seemed the ideal way to learn more about their strategies and thus to address these questions. Therefore, this was the approach taken in the next experiment.

1. Method

The children who participated in this experiment were the same 3-, 4-, and 5-year-olds who participated in the counting and magnitude comparison studies. Each child was brought individually to a videotaping laboratory within the preschool and seated at a table. The room contained a camera mounted in a corner; aside from the camera, all of the videotaping equipment was kept in an adjacent room. Children were given the following instructions:

Today we're going to play another number game. I want you to imagine that you have a pile of oranges. I'll give you more oranges to add to your pile; then you need to tell me how many oranges you have altogether. Okay? You have *m* oranges, and I'm going to give you *n* to add to your pile. How many do you have altogether?

Children were presented the 25 problems formed by the factorial combinations of augend (1–5) and addend (1–5). Each problem was presented twice. The 25 problems were arbitrarily divided into groups of 9, 8, and 8, with each group being presented to the child in a different session. When all 25 problems had been presented, the cycle was repeated; thus, each child received 50 items spread over six sessions. Each session lasted approximately 15 minutes. No feedback was given at any time, except for periodic assurances that the child was doing well.

2. Results

The analyses of addition performance that will be reported differ in two ways from the analyses of counting and magnitude comparison. First, except for preliminary, aggregate level analyses, the data of the 4- and 5-year-olds are analyzed together because children in the two age groups were quite similar in their absolute level of performance, the predictors of their performance, and the strategies that they used. Second, again except for the preliminary analyses, only cursory descriptions of the 3-year-olds' performance will be reported, and no model will be postulated. The absolute level of performance among children in this age group was very low, none of the variables that were examined predicted the relative difficulty of the problems for them, and visible strategies were rarely observed. Thus, we simply do not possess rich enough data to infer much about the 3-year-olds' knowledge of addition, beyond the facts that they quite often answered correctly $1 + 1$, $2 + 1$, and $1 + 2$ and that they rarely used visible strategies to do so.

a. Overall level of performance. As in the previous experiments on counting and magnitude comparison, the period from 3 to 5 years proved to be one of substantial development. Three-year-olds were correct on 20% of the addition problems, 4-year-olds were correct on 66%, and 5-year-olds were correct on 79%.

b. Predictors of error patterns. Stepwise regression analyses were used to examine the distribution of errors on the 25 problems. The initial analysis included five predictor variables: size of the augend, size of the addend, size of the minimum number, size of the sum, and size of the square of the sum.

None of these predictors was significantly correlated with the 3-year-olds' error patterns, each of them failing to account for even 15% of the variance. In contrast, the 4- and 5-year-olds' error patterns were much more orderly. The sum was found to be the best single predictor, accounting for 49% of the variance.

Careful examination of the 4- and 5-year-olds' data suggested that three tendencies that were not captured in the above list of predictors also accounted for substantial amounts of variance. First, problems involving the number 5 as either augend or addend were much easier than would have been expected from their

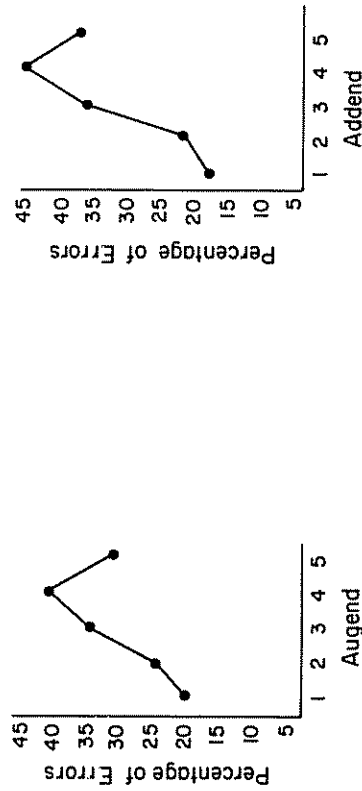


Fig. 6. Percentage of errors on arithmetic problems as a function of augend and addend size.

absolute magnitudes. As shown in Fig. 6, problem difficulty increased monotonically with augend and addend sizes for the range 1-4, but this trend was broken at the number 5. Second, as Svenson (1975) found with third graders, the size of the last number exercised an effect independent of the sum. Third, as Groen and Parkman (1972) and Knight and Behrens (1928) found with first and second graders, ties were easier than would have been anticipated. A new regression analysis indicated that these four variables accounted for 86% of the variance in the distribution of 4- and 5-year-olds' errors: sum accounted for 49%; sum and the presence of a five accounted for 74%; sum, fives, and the size of the last number accounted for 80%; and sum, fives, last, and ties accounted for 86%.

c. Consistency of performance. Consistency of addition performance was similar to consistency of magnitude comparison performance. Children's total number of correct answers showed substantial consistency; over the two occasions, the correlations were $r = .90$ for the 5-year-olds, $r = .71$ for the 4-year-olds, and $r = .77$ for the two groups combined. Performance on individual problems, however, showed little consistency; 5-year-olds advanced the same incorrect answer on 3% of trials, and 4-year-olds did so on 4%. As with the magnitude comparisons, these data seemed to demand that in any model of addition performance, errors would be at least in part the products of probabilistic and/or reconstructive processes rather than solely of determinate, reproductive ones.

d. Strategy use. Scrutiny of the videotapes revealed that children adopted at least four different approaches to solving addition problems: the counting fingers strategy, the fingers strategy, the counting strategy, and (for want of a better name) the no visible strategy approach. These strategies could be identified very reliably; two independent raters agreed on 98 of 100 ratings. The four

strategies that were identified differed in their visible and audible manifestations, in their accuracy, in their temporal characteristics, and in the types of errors with which they were associated.

In the counting fingers strategy, children put up the fingers on one hand, then put up the fingers on the other hand, and then counted the two sets of fingers. The counts generally began with the leftmost finger on the left hand, the hand that was usually put up first, and ended on the rightmost finger on the right hand, the hand that was usually put up second. As shown in Table IV, the strategy was used moderately often, was very slow, and produced accurate performance. When errors occurred, they were most often close to the correct solution.

In the fingers strategy, children put up their fingers as in the counting fingers strategy but showed no evidence of counting them. The fingers and the counting fingers strategies produced similarly accurate performance, but the time needed to execute the fingers approach was much shorter. This tendency was very consistent over individuals; 20 of the 21 children who used both strategies at least twice had faster mean solution times on the fingers than on the counting fingers trials. Errors on the fingers trials, like those on the counting fingers trials, were usually close misses (Table IV).

In the counting strategy, children counted aloud but without any visible referent. In all cases these counts started from 1. The strategy was associated with moderately long solution times and was the least accurate and the least often used of the four strategies. Errors were often quite distant from the correct answer; fewer than half were within one of it.

The fourth category was a catchall for those trials on which the children did not engage in any visible or audible behaviors that seemed related to the addition process; these trials were grouped together as "no visible strategy." This approach was the most frequently used, the second least accurate, and the most rapidly executed of the four strategies. The accuracy data fell into a distinctly bimodal distribution. Sixteen of the 30 children performed at high accuracy rates on no visible strategy trials; they were correct on 86-100% of trials. The other 14

TABLE IV
Characteristics of Arithmetic Strategies

Strategy	Trials on which strategy used (%)	Mean solution time (sec)	Correct answers (%)	Errors on which answer was within one of correct sum (%)
Counting fingers	15	14.0	87	70
Fingers	13	6.6	89	80
Counting	8	9.0	54	44
No visible strategy	64	4.0	66	41

children were much less accurate, ranging from 13 to 67% correct. This bimodality was not an artifact of a small number of observations producing high variability, as the approach was used on an average of 32 trials per child. Both the high-accuracy and the low-accuracy children executed the approach rapidly, and both tended to make errors that were far from the correct answer.

e. Strategy use and problem difficulty. The most striking finding of the experiment involved the connection between strategy use and problem difficulty. As shown in Table V, the percentage of errors that each problem elicited was very closely related to the percentage of trials for that problem on which one of the three visible strategies was used ($r = .91$). The relationship was actually attenuated slightly by the tendency of children more frequently to solve problems on which they used a visible strategy; if only the percentage of errors on no visible strategy trials is used to estimate difficulty, the correlation between frequency of visible strategy use and problem difficulty increases slightly to $r = .92$.

Using the visible strategies more often on the more difficult problems proved to be highly adaptive. Performance was more accurate on 24 of the 25 problems when visible strategies were used than when they were not. Moreover, the frequency of use of visible strategies on the 25 problems correlated $r = .73$ with the gain in accuracy from using a visible strategy (the difference on each problem between error rates when children did and did not use a visible strategy).

These results answered two of our original questions. First, use of visible strategies was indeed related to the characteristics of the addition problems, with the more difficult problems eliciting a greater frequency of strategy use. Second, employing such visible strategies aided children's performance, with the greatest benefits accruing on the most difficult problems. Left unanswered, however, was the question of how children chose among the available strategies so as to produce these relationships.

Two types of decision processes seemed the most likely. One possibility was that children were consciously aware of problem difficulty and used their assessment of it as the decision criterion, adopting a visible strategy when they judged the problem too difficult to solve without one. This might be labeled the metacognitive interpretation. The second possibility was that the decision to use a visible strategy arose as a by-product of other solution processes rather than through any process involving explicit judgments of difficulty. This might be labeled the by-product of solution processes interpretation.

The metacognitive hypothesis is based on the following model of the strategy choice process:

problem difficulty \rightarrow judgments of problem difficulty \rightarrow use of visible strategies

TABLE V
Errors and Strategy Use on Arithmetic Problems

Problem	Errors (%)	Trials on which visible strategy used (%)
1+1	3	10
1+2	7	15
2+1	10	22
2+2	12	18
5+1	14	25
3+1, 5+2	19	24, 32
4+1, 2+3	20	24, 39
1+5	25	34
3+2	27	39
1+4, 5+5, 3+3	31	32, 34, 34
4+2, 5+3	32	37, 39
1+3	34	31
2+4	36	47
2+5	37	41
3+5, 4+4	41	56, 34
4+5, 5+4	47	49, 46
3+4	53	54
4+3	59	58

This model implies that the actual difficulty of problems and children's judgments of their difficulty should be highly correlated (especially given the observed high correlation between problem difficulty and visible strategy use). To test this prediction, we asked a group of 12 5-year-olds, students at a nursery school very similar to the one at which the original experiment had been run, to label each of the 25 problems as easy, hard, or in between. "Hard" ratings were quantified as 2, "easy" ratings as 0, and "in-between" ratings as 1. These ratings were correlated with problem difficulty as estimated from three error and solution time data sets: ours, Knight and Behrens' (1928), and Groen and Parkman's (1972). Regression analyses of the difficulty ratings were also run in order to determine whether the same variables predicted the ratings as predicted actual difficulty.

The difficulty ratings correlated only moderately with the actual difficulty of the problems as estimated by the errors of our sample ($r = .47$), the errors of the Knight and Behrens sample ($r = .50$), and the solution times of the Groen and Parkman sample ($r = .31$). It might be argued that these relatively low correlations could have resulted from our obtaining the addition performance and the

difficulty ratings from different samples of children. However, the correlations were much higher on other intersample comparisons: between the percentages of errors on the 25 problems in our samples and those of Knight and Behrens ($r = .77$), between the percentage of errors on these problems in our sample and the solution times in the Groen and Parkman sample ($r = .72$), and between the percentage of errors in the Knight and Behrens sample and the solution times in the Groen and Parkman sample ($r = .89$). The predictors of the difficulty ratings also differed from the predictors of the errors and solution times; the size of the larger number was the best predictor of the ratings, but, as reported above, the sizes of the minimum number and of the sum were the best predictors of actual difficulty. Thus, the data lent little support to the view that the children's judgments of problem difficulty were responsible for the relationship between strategy use and problem difficulty.

C. A MODEL OF STRATEGY CHOICE IN ADDITION

The second possibility was that the correlation between visible strategy use and problem difficulty was a byproduct of the solution process, rather than resulting from any knowledge of problem difficulty per se. One model that would give rise to such a correlation is displayed in Fig. 7. The representation, shown on the top of Fig. 7, includes the probabilities that the answer to each addition problem can be recalled correctly without recourse to any external strategy and with a confidence that exceeds some arbitrarily set criterion level. These probabilities will differ for each child and for each situation in which a given child finds himself or herself. The ones shown in Fig. 7 were calculated for the group level probabilities within the experiment presented here that on each problem, children used no visible strategy and generated the correct answer.

The process that is applied to this representation, shown on the bottom of Fig. 7, begins with the setting of a confidence criterion by which children decide whether they are sure enough of an answer to give it. Then they try to recall the answer to the problem.⁹ If their confidence in the recalled answer exceeds the

⁹"Recall" is used here in quite a loose sense. It is entirely possible that on some of the trials where children did not use a visible strategy and answered correctly, they used the min approach described by Groen and Parkman or some other strategy. The evidence was equivocal as to which particular approach they employed. Regression analyses of all trials on which the children used no visible strategy indicated that the sum of the numbers was the best predictor of errors, accounting for 47% of the variance. The size of the minimum number was the next best predictor, accounting for 41% of the variance. When only nontrivial cases were considered, the size of the minimum number was the best predictor, accounting for 66% of the variance; the sum of the two numbers was the next best, accounting for 52%. Regression analyses of solution times revealed a similar picture. When all correct trials were considered, sum was the best predictor, accounting for 33% of the variance; the minimum number accounted for 23%. However, when just nontrivial cases were considered, the size of the minimum number accounted for 62% of the variance in solution times, while the size of the sum

Representation

Sums and (Probability of Recalling Each Sum with Confidence Beyond Criterion)

Augend	Addend				
	1	2	3	4	5
1	2(.86)	3(.75)	4(.45)	5(.35)	6(.49)
2	3(.68)	4(.69)	5(.49)	6(.28)	7(.26)
3	4(.58)	5(.40)	6(.35)	7(.09)	8(.13)
4	5(.55)	6(.23)	7(.07)	8(.36)	9(.15)
5	6(.68)	7(.52)	8(.34)	9(.08)	10(.38)

PROCESS:

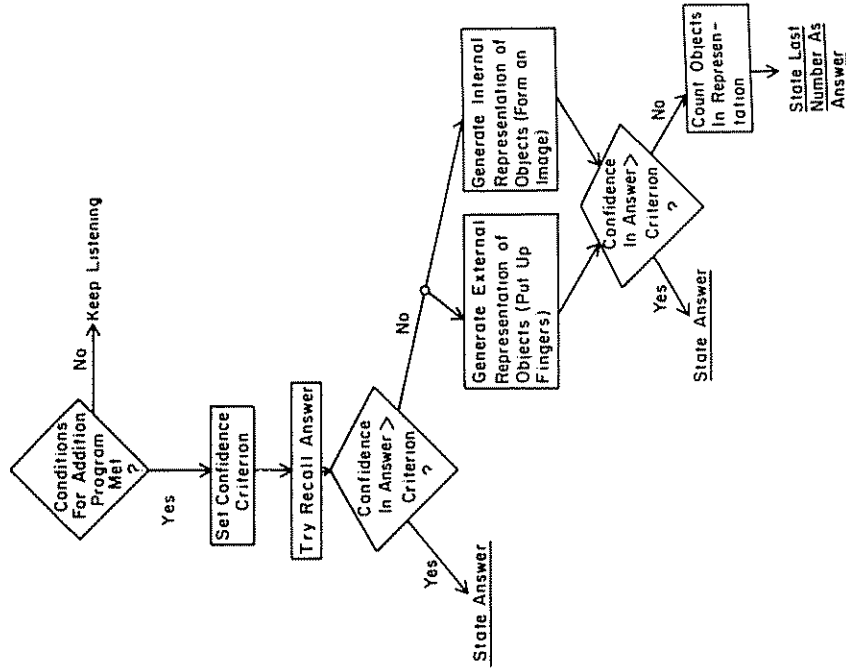


Fig. 7. Model of strategy choice in addition. (Note: The probabilities in parentheses represent the percentage of trials on which children used no visible strategy and answered correctly. They can be thought of as associative strengths.)

criterion, they give it as the answer. Otherwise, they augment their representations of the numbers indicated by *augend* and *addend*, either externally, by putting up their fingers, or internally, by forming some type of imaginal representation. If their confidence in an answer at this point exceeds the criterion, they give it; otherwise, they count their fingers or the imaged objects and state the last number of the count as their result.

This model gives rise to the four approaches that we observed. If children answer at the first "recall" point, they will not have used any visible strategy. If they put up their fingers but answer without counting them, they will have used the finger strategy. If they image objects corresponding to *augend* and *addend* and then count aloud, they would be classified as having used the counting strategy. Finally, if they put up their fingers, count them, and answer after counting, they would be classified as having used the counting fingers strategy.

The model also suggests explanations for the relative solution times and accuracy rates of the four strategies. It predicts straightforwardly that the no visible strategy approach should be the fastest, the fingers approach the next fastest, and the counting and counting fingers approaches the slowest of the four strategies; all of the processing steps necessary to execute each of the faster strategies are included within the steps necessary to execute the slower ones. Predictions of relative accuracy also can be derived, albeit not quite as directly. Both the overall low accuracy and the bimodal distribution of accuracies of the no visible strategy approach follow from the view that some children used this approach because they set loose criteria for deciding when they knew the answer, and others used it because they did not need any external aids to retrieve the correct answer. The high accuracy of the counting fingers strategy would have been expected since 4- and 5-year-olds have been shown to be very adept at counting the 2-9 objects required by the problems (Gelman & Gallistel, 1978). Several considerations may have contributed to the high accuracy of the fingers approach: pattern recognition of the number of fingers that were put up, kinesthetic cues associated with putting up particular sets of fingers, and longer search time than that typical of the no visible strategy approach.¹⁰ Finally, the relative inaccuracy of the

accounted for 50%. Another possibly relevant source of data was that on the counting and counting fingers strategy trials, on which the 4- and 5-year-olds were heard counting aloud, their counts always started from one. Thus, it was difficult to tell on the no visible strategy trials whether children generally used the *min* approach, whether they generally used the sum approach (like the *min* model but counts start at one), whether some children used one and some the other, or whether the same children sometimes used one and sometimes the other.

¹⁰To understand why kinesthetic cues associated with raising one's fingers might be helpful for recalling a sum, consider what would happen if your eyes were closed and someone lifted one or more fingers on each of your hands. With at least some combinations, the number of fingers raised might well "feel" like 2, 4, or 10 (for example). Whether children are helped by such information is an open question, but the kinesthetic cues offered at least one explanation for the superior accuracy of the fingers approach.

counting approach was predicted by Kosslyn's (1978) finding that children have difficulty maintaining a clear image for as long as it took to execute this strategy, 9 seconds.

Perhaps the most important feature of the model is that it allows us to account for the correlation between problem difficulty and strategy use. No explicit knowledge about problem difficulty is required to produce the relationship. Instead, at each step in the solution process, the child considers whether his or her confidence in an answer exceeds the level demanded by his or her criterion. The more difficult the problem, the less likely it is that this will occur. Thus, at least those children whose criteria are relatively high are led to take increasingly effortful steps to solve the more difficult problems. In a sense, these children use internalized strategies when they can and externalized ones when they must.

Such flexible strategy selection has obvious advantages. It minimizes effort while maximizing the probability of a correct answer. By simply adjusting the confidence criterion, children can adapt to situations in which accuracy is the critical consideration or to situations in which speed or lack of effort is. Adults have been shown to possess similar propensities to avoid cognitive effort and to adjust strategies to the difficulty of particular problems. Siegler and Atlas (reported in Siegler & Klahr, in press) found that adults computed quantitative solutions to balance scale problems only when the problems could not be solved by simpler qualitative comparisons. Glushko and Cooper (1978) found that even in simple sentence verification situations, adults varied their approach depending upon the task demands.

These and the findings presented here suggest that from early in childhood, two systemic principles may govern the construction of information-processing routines: minimize the effort needed to accomplish any particular goal, and maximize each routine's flexibility to adjust to different task environments. These principles would have obvious adaptive value and would be in keeping with the flexible strategy use that has been so frequently observed. The implication for future research is that examining the ways in which people choose among alternative strategies for solving problems may be at least as informative as focusing on how they execute any given strategy.

VI. Conclusions: The Development of Numerical Knowledge

At the outset of this article, we proposed to examine several aspects of young children's knowledge of numbers, to devise models of their knowledge within each task domain, and eventually to formulate one or more comprehensive models, including the information within each of the specific ones. This last goal, the formulation of models that stretch across task domains, has been given

considerable homage in the abstract by developmental psychologists, but few such accounts have been stated at a sufficiently precise level to be meaningfully evaluated. We believe that the formulation of detailed but encompassing models is crucial to understanding cognitive growth. Therefore, in this last section, we shall focus on the issue of what type of larger system might produce the numerical skills of preschoolers that we observed in each of the particular areas.

Newell (1973), in his article, "You Can't Play 20 Questions With Nature and Win," eloquently argued the case for building large-scale integrative models. He contended that although cognitive psychologists have succeeded in identifying robust phenomena and in accounting for performance in particular situations, the research has failed to cumulate. Among Newell's suggested means of escape from this dilemma was devising a single model capable of producing performance on many tasks. This suggestion has been influential in motivating large-scale computer simulations of thought, language, imagery, and memory (Anderson, 1976; Kintsch & van Dijk, 1978; Kosslyn, 1978; LNR, 1975). The models that we shall present of young children's knowledge of numbers differ from these others in being much less ambitious in scope, in not yet being specified at the level of running computer simulations (programs in OPS5 are currently being written), and in being primarily concerned with development. However, the motivation for building them was the same.

Figure 8 outlines our current understanding of preschoolers' knowledge of numbers. The three models within Fig. 8 are ordered from the least to the most advanced, and correspond to the knowledge that we hypothesize is most often possessed by 3-, 4-, and 5-year-olds, respectively.¹¹ A cursory examination of the models reveals two features: they are quite forbidding looking, and they appear rather similar to each other. Because of the models' forbidding appearance, we will describe one of them, Model II, at some length. Because of the similarities among the three models, we will characterize the depictions of the least and the

¹¹The models in Fig. 8 are intended to provide descriptions of the modal tendencies among 3-, 4-, and 5-year-olds, but should not be taken to imply a lockstep progression among the three skills or a perfect correlation between age and skill. Empirically, each skill possessed a moderately high correlation with age and also a moderately high correlation with the other two skills. Age correlated $r = .64$ with percentage of correct answers on the numerical comparison problems, $r = .68$ with percentage of correct answers on the addition problems, and $r = .49$ with how high children counted. Across the three age groups, the percentage of correct answers on the addition task correlated $r = .80$ with the percentage of correct answers on the numerical comparison task, the percentage of correct answers on the comparison task correlated $r = .60$ with the highest number counted on the counting task, and the percentage of correct answers on the addition task correlated $r = .65$ with the highest number counted on the counting task. The correlations between pairs of tasks were also fairly high within age groups, averaging $r = .54$ for the 9 within-age correlations. Overall, 27 of the 39 children (70%) who performed all three tasks would have been assigned to the same aggregate model by virtue of their performance on each of the three individual tasks.

most advanced knowledge in terms of differences between them and the formulation of intermediate level knowledge.

When we examine Model II, the model of 4-year-olds' knowledge, a basic hierarchical form becomes evident. Numbers as a class are at the top of the hierarchy, then categories of numbers (e.g., small numbers), and finally individual numbers (e.g., 6). There are connections both across and within levels of the hierarchy.

At the top, numbers as a class can be operated upon by a number of processes: they can be counted, their magnitudes can be compared, and they can be added and (presumably) subtracted. That children treat numbers as a class distinct from other classes was evident in what they did not do as much as in what they did. No child ever gave a nonnumeric answer to an addition problem or used any nonnumeric term in counting. Children at times did use nonstandard numbers in their counting strings, but the nonstandard numbers were always combinations of standard ones. Therefore, we believe that several processes that can operate upon numbers are attached to numbers as a class rather than to particular groups of numbers or to individual numbers. (The details of these processes are omitted from the diagrams in Fig. 8 only because of considerations of space; they are shown earlier in the article, in the figures indicated.)

At the next lower level of the hierarchy are categories of numbers, ordered in terms of magnitudes. Both the number conservation data reported by Siegler (1981a) and the magnitude comparison data reported in the present investigation suggest that these categories possess psychological reality for young children. Illustratively, the conservation operators that are applied to small numbers differ from those that are applied to large ones; as shown in Fig. 1B, children apply the correct transformational rules to small numbers of objects, but with larger groups judge the longer row to have more.

The numerical categories occupy an intermediate position within the hierarchy. Each category is linked both upward to the class of numbers and downward to individual numbers. The particular probabilities linking the categories to the individual numbers are based on those that appeared in the Fig. 5B representation—the empirically derived probabilities that 4-year-olds assigned each label to each number. (In order to make Fig. 8 relatively readable, all probabilities have been rounded to the nearest tenth and probabilities of 10% or less are not shown.)

The lowest level of the hierarchy involves individual numbers. In addition to being tied to the category labels with varying probabilities, the numbers are at times tied to each other by "next" connections. Some numbers are also labeled as members of the digit repetition and rule applicability lists. The smaller ones are involved in specific addition facts that the children are more or less confident of knowing. Although our experiments did not tap other information about indi-

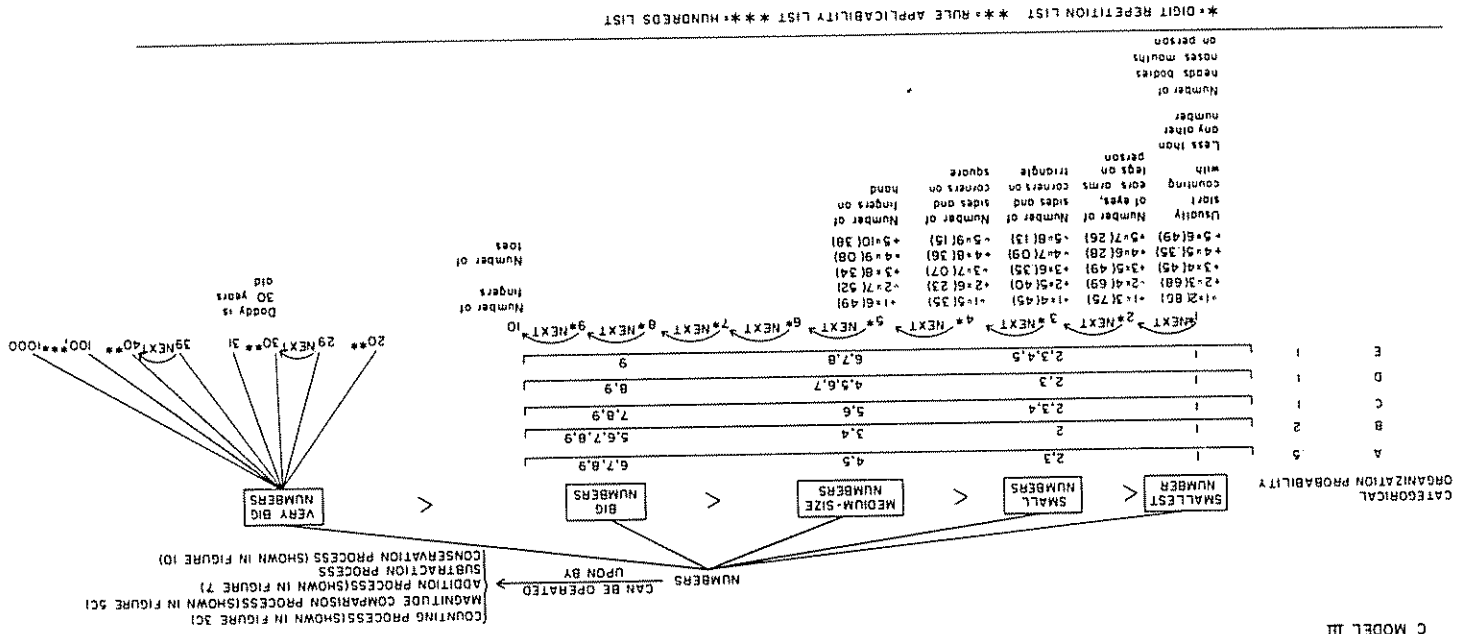
vidual numbers, informal discussions with preschoolers suggest that they know many other facts about them. For example, a 4-year-old told us that 1 is the number that she starts counting with, that it is the number of heads, bodies, noses, and mouths on a person, and that it is the smallest number.

The Model II depiction of moderately skilled performance provides a vantage point for considering the more advanced knowledge depicted in Model III and the less advanced knowledge depicted within Model I. First, let us consider some properties that are hypothesized not to change within this age and skill range. As mentioned above, the basic hierarchical form of the representation is constant across the three models. The children have knowledge about numbers in general, about categories of numbers, and about particular numbers. Also relatively constant across the models are many of the particular connections within and across levels of the hierarchy: even in Model I, the larger numbers are more often assigned to the larger categories; even in Model I, some of the "next" connections between digits are present; even in Model I, some facts linking individual numbers to other semantic properties (e.g., people have two hands) are known. Development in these (though not all) aspects of the representations appears to be a gradual, incremental process.

At the other extreme, development can be seen in sharpest relief in the processes that children apply to their representations. These processes change greatly in all three task domains between Models I and II, and the processes for counting and comparing undergo large changes between Models II and III as well. The counting process changes from using only "next" connections in Model I to also using rule applicability and digit repetition lists in Model II and III. The pattern is reminiscent of the often-expressed speculation that development entails at least as great a growth in what children can do with information as in the amount of information that they possess (Bruner, 1973; Piaget, 1972; Simon, 1972).

The issue of intertask relationships is addressed implicitly in all three models. The models suggest that preschoolers' understandings of counting, comparing, conserving, and adding are linked in some ways, but not in all of the ways that they could be. All of the processes operate upon a common representation, and this seems to produce some commonalities. Most dramatically, both conservation and magnitude comparison processes utilize the categorizations of numbers, and all processes except magnitude comparison make use of the links among the individual numbers. Other intertask connections that could have been present

Fig 8 (C) Model III



were not, however. Preschoolers could have used their knowledge of counting to compare numerical magnitudes but they did not seem to. They could have used their knowledge of comparing to add numbers, as in the Groen and Parkman min model, but again they did not seem to. This last finding especially suggests that early mathematical skills may develop in relative isolation from one another; once children are proficient in the individual skills, they may make greater use of the potential interconnections among them.

How can we evaluate the quality of these models? Empirically, they predict in detail preschoolers' counting, comparing, conserving, and adding. Model II can again be used to illustrate. When counting, the model will stop at "9s," will skip and repeat entire decades, will introduce nonstandard numbers if the boundary of its digit list is too high, and will count on at least to the next "9" from points within or beyond its spontaneous counting range. When comparing the magnitudes of numbers, it will err most often on problems with large minima and small splits, will assign labels that correspond reasonably well to the relative magnitudes of the numbers, and will learn from instruction that adds an overall categorical organization to the existing connections between individual numbers and categories. When performing number conservation problems, it will judge small number problems in terms of the type of transformation but will judge large number problems in terms of the relative lengths of the rows. When adding, it will usually recall the answers to the easiest problems without using visible strategies, and will more frequently use such visible strategies as the problem increases in difficulty. Thus, the model mimics a considerable range of preschoolers' behaviors in manipulating numbers.

A second virtue of the models is the quality that Klahr and Wallace (1976) termed developmental tractability. For most of the changes between models, we can easily imagine how the more advanced form could grow out of the less advanced one. In counting, children first learn the "next" connections that are the only relations that bind the first numbers they encounter; then they add to this knowledge information about the cyclical patterns inherent in the next higher numbers they learn; eventually they extend the list membership notion to include the much larger numbers that they encounter yet later. In learning about numerical magnitudes, children first obtain a rough sense of magnitudes that allows them to assign individual numbers to categories having some correspondence to the sizes of the numbers; then they learn how to use the categorical information to compare magnitudes; finally they impose overall categorical organizations to subsume the connections between individual numbers and categories but avert errors. In learning about number conservation, children first rely on the type of transformation only in limited situations, and gradually expand that reliance to encompass all three transformations and all set sizes. In addition, children first memorize solutions to specific problems, and then learn supplementary reconstructive strategies to use on problems where they cannot retrieve the answer.

Thus, development in this age range and content area involves few false starts; children build on what they already know to construct increasingly successful approaches.

The separation between representations and processes in the model proved useful for specifying the source of this developmental tractability. Modeling approaches that focus solely on processes, such as the rule assessment approach, might have revealed as much about developmental changes in the preschool period, but probably would not have revealed the developmental constancies that also were present. The representation-process distinction also provided a basis for hypothesizing why early error patterns often foreshadow later reaction time patterns: increasingly powerful processes operating upon fundamentally similar representations.

A final strength of the models is that they should be easy for us and other investigators to build upon. They can be expanded both outward, to encompass additional aspects of preschoolers' knowledge of numbers, and upward, to include the more advanced knowledge of school age children. One early sign of this intellectual "developmental tractability" was the ease with which we could integrate the new information that 4-year-olds in the magnitude comparison training experiment were taught with the model of their existing knowledge (i.e., the extension of the Fig. 5B model to produce the one represented in Fig. 5D). We anticipate that it also will be relatively straightforward to add to the present models information about preschoolers' ability to subtract, to count objects, to subitize, and to estimate the numerosity of large collections. Other reasonable goals include expanding the models upward so as to include more complex addition and subtraction skills, the relationships of addition and subtraction to multiplication and division, and the extension of arithmetic operations to the rational numbers. This is not to underestimate the difficulty of achieving these objectives, but rather to affirm that the present hierarchical models provide a base that is far from closed.

All models have weaknesses as well as strengths. The two greatest weaknesses of the present models seem to be a lack of detail concerning how children choose to use a particular process in a particular situation and a lack of flexibility to cope with novel situations. With regard to the first point, the model of each process is introduced by the rather opaque test, "conditions for process x met?" Even the 3-year-olds counted when we asked them to count, compared when we asked them to compare, and added when we asked them to add. We do not understand, however, how they knew to do so. Recent efforts to discover how children interpret arithmetic word problems (Carpenter & Moser, 1981; Greeno *et al.*, 1981; Neshor, 1981) represent a first step toward modeling how children understand instructions. Without further research on how language understanding occurs, however, this part of the model must remain a black box.

The second weakness of the models is their lack of flexibility for adapting to

novel situations. This weakness applies most directly to the portrayal of magnitude comparison. The links between categories and individual numbers are presented as fixed in all three models. This may be a realistic depiction of the long-term memory contents of preschoolers, but almost certainly would not continue to be a realistic depiction in older children and adults. The lack of even a poorly developed mechanism for taking into account the effects of context is a weak point in the general developmental tractability of the models. In addition, just as children may not apply the same process to solving each addition problem, they may not apply the same process to solving every magnitude comparison problem. They may directly retrieve some pairs, may judge others relative to some common reference point, and may in general use any number of idiosyncratic judgment techniques in the comparison process. Illustratively, when one of us recently asked his 5-year-old son whether 16 or 33 was the bigger number, the child counted out loud from 1 to 16 and then said that 33 was bigger. Our error data, Sekuler and Mierkiewicz's (1977) reaction time data, and previous observations of this child indicate that the counting strategy is far from the rule in this age range or even for this individual; nonetheless, children may use it, and probably many other approaches, sometimes.

Both the strengths and the weaknesses of our models converge on two final points. First, conceptual development is far too complex for us to assess children's understanding by examining performance on a single task. No one age is the age at which a concept is understood, and there is little meaning to saying that one concept is understood before, after, or simultaneously with another. Conceptual understanding has many facets, and only by investigating a concept both broadly and deeply do we have any hope of discovering what people know about it. Second, it is possible and desirable to build integrative models of children's knowledge across different tasks corresponding to a single concept. These models help us to realize which aspects of children's understandings we have accounted for, to notice the aspects that we have not yet addressed, and to face those implications that we did not intend and would like to change. In short, such models can help the work cumulate.

REFERENCES

- Aiken, L. R., & Williams, E. N. Three variables related to reaction time to compare digits. *Perceptual and Motor Skills*, 1968, 27, 199-206.
- Anderson, J. R. *Language, memory, and thought*. Hillsdale, N.J.: Erlbaum, 1976.
- Anderson, J. R. Arguments concerning representations for mental imagery. *Psychological Review*, 1978, 85, 249-277.
- Ashcraft, M. H., & Battaglia, J. Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, 4, 527-538.
- Banks, W. P. Encoding and processing of symbolic information in comparative judgments. *The Psychology of Learning and Motivation*, 1977, 11, 101-159.
- Banks, W. P., Fujii, M., & Kayra-Stuart, F. Semantic congruity effects in comparative judgments of magnitude of digits. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, 2, 435-447.
- Banks, W. P., & Hill, D. K. The apparent magnitude of number scaled by random production. *Journal of Experimental Psychology*, 1974, 102, 353-376. (Monograph)
- Bellin, H. Cognitive capacities of young children: A replication. *Science*, 1968, 162, 920-921.
- Bever, T. G., Mehler, J., & Epstein, J. What children do in spite of what they know. *Science*, 1968, 162, 921-924.
- Braune, M. D. S. The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods. *Psychological Monographs*, 1959, 73 (Whole No. 475).
- Brammer, C. J. *The origins of the number concept*. New York: Praeger, 1979.
- Brown, A. L. The construction of temporal succession by pre-operational children. In A. D. Pick (Ed.), *Minnesota Symposium on Child Psychology* (Vol. 10). Minneapolis: Univ. of Minnesota, 1976.
- Brown, A. L. Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology*. Hillsdale, N.J.: Erlbaum, 1978.
- Bruner, J. S. *Beyond the information given: Studies in the psychology of knowing*. New York: Norton, 1973.
- Bryant, P. E. *Perception and understanding in young children*. New York: Basic Books, 1974.
- Carpenter, T. P., & Moser, J. M. The development of addition and subtraction problem solving skills. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective*. Hillsdale, N.J.: Erlbaum, 1981, in press.
- Chi, M. T., & Klahr, D. Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology*, 1975, 19, 434-439.
- Cooper, R., Starkey, P., Blevins, B., Goth, P., & Lettner, E. *Number development: Addition and subtraction*. Paper presented at the meeting of the Jean Piaget Society, Philadelphia, May 1978.
- Estes, W. K. The problem of inference form curves based on group data. *Psychological Bulletin*, 1956, 53, 134-139.
- Fairbank, B. A. *Experiments on the temporal aspects of number perception*. Unpublished doctoral dissertation, University of Arizona, 1969.
- Flavell, J. H. Stage-related properties of cognitive development. *Cognitive Psychology*, 1971, 2, 421-453.
- Fuson, K. C., & Richards, J. Children's construction of the counting numbers: From a spew to a bidirectional chain. Paper presented at the meetings of the American Education Research Association, 1980, Boston.
- Fuson, K. C., & Richards, J. The acquisition and elaboration of the number sequence. In C. Brunner (Ed.), *Progress in cognitive development*. Vol. 1. Berlin and New York: Springer-Verlag, 1982, in press.
- Gelman, R. The nature and development of early number concepts. *Advances in Child Development and Behavior* 1972, 7, 115-167.
- Gelman, R. Cognitive development. *Annual Review of Psychology*, 1978, 29, 297-332.
- Gelman, R., & Gallistel, C. R. *The child's understanding of number*. Cambridge, Mass.: Harvard Univ. Press, 1978.
- Gelman, R., & Starkey, P. Development of addition and subtraction abilities prior to formal schooling in arithmetic. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective*. Hillsdale, N.J.: Erlbaum, 1981, in press.
- Ginsburg, H. *Children's arithmetic: The learning process*. New York: Van Nostrand, 1977.

- Glushko, R. J., & Cooper, L. A. Spatial comprehension and comparison processes in verification tasks. *Cognitive Psychology*, 1978, **10**, 391-421.
- Greeno, J. G., Riley, M. S., & Gelman, R. *Young children's counting and understanding of principles*. Unpublished manuscript.
- Greeno, G. J., & Parkman, J. M. A chronometric analysis of simple addition. *Psychological Review*, 1972, **79**, 329-343.
- Greeno, G. J., & Resnick, L. B. Can preschool children invent addition algorithms? *Journal of Educational Psychology*, 1977, **69**, 645-652.
- Hebber, K. *The development of children's problem-solving skills in addition*. Unpublished doctoral dissertation, Cornell University, 1976.
- Ilg, F., & Ames, L. B. Developmental trends in arithmetic. *Journal of Genetic Psychology*, 1951, **79**, 3-28.
- Johnson, S. C. Hierarchical clustering schemes. *Psychometrika*, 1967, **32**, 241-254.
- Kintsch, W., & van Dijk, T. A. Toward a model of text comprehension and production. *Psychological Review*, 1978, **85**, 363-394.
- Klahr, D., & Wallace, J. G. *Cognitive development: An information-processing view*. Hillsdale, N.J.: Erlbaum, 1976.
- Knight, F. B., & Behrens, M. S. *The learning of the 100 addition combinations and the 100 subtraction combinations*. New York: Longmans, Green, 1928.
- Kosslyn, S. M. Imagery and cognitive development: A teleological approach. In R. S. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, N.J.: Erlbaum, 1978.
- Kosslyn, S. M., Murphy, G. L., Bemesderfer, M. E., & Feinstein, J. J. Category and continuum in mental comparisons. *Journal of Experimental Psychology: General*, 1977, **106**, 341-375.
- Kruskal, J. B., & Wish, M. *Multidimensional scaling*. Beverly Hills, Calif.: Sage Univ. Press, 1978.
- Lehman, H. *Introduction to the philosophy of mathematics*. Totowa, N.J.: Rowman & Littlefield, 1979.
- LNR. *Explorations in cognition*. San Francisco: Freeman, 1975.
- Maki, R. H. Categorization and distance effects with spatial linear orders. *Journal of Experimental Psychology: Human Learning and Memory*, 1981, **7**, 15-32.
- Mans, F. R., Keating, D. P., & Morrison, F. J. Developmental differences in the allocation of processing capacity. *Journal of Experimental Child Psychology*, 1980, **29**, 156-169.
- Mehler, J., & Bever, T. G. Cognitive capacity of very young children. *Science*, 1967, **158**, 141-142.
- Miller, S. A. Nonverbal assessment of conservation of number. *Child Development*, 1976, **47**, 722-728.
- Moyer, R. S., & Bayer, R. H. Mental comparison and the symbolic distance effect. *Cognitive Psychology*, 1976, **8**, 228-246.
- Moyer, R. S., & Dumas, S. T. Mental comparison. *The psychology of learning and motivation*, 1978, **12**, 117-155.
- Moyer, R. S., & Landauer, T. K. The time required for judgments of numerical inequality. *Nature (London)*, 1967, **215**, 1519-1520.
- Neisser, V. General, academic, and artificial intelligence. In L. B. Resnick (Ed.), *The nature of intelligence*. Hillsdale, N.J.: Erlbaum, 1976.
- Nesher, P. Levels of description in the analysis of addition and subtraction. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective*. Hillsdale, N.J.: Erlbaum, 1981, in press.
- Newell, A. A note on process-structure distinctions in developmental psychology. In S. Farnham-Diggery (Ed.), *Information processing in children*. New York: Academic Press, 1972.

- Newell, A. You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973.
- Parkman, J. M. Temporal aspects of digit and letter inequality judgments. *Journal of Experimental Psychology*, 1971, **91**, 191-205.
- Peano, G. *Formule de mathematiques* (Vols. 1-5). Turin: Broca, 1894-1908.
- Piaget, J. *The child's concept of number*. New York: Norton, 1952.
- Piaget, J. Intellectual evolution from adolescence to adulthood. *Human Development*, 1972, **15**, 1-12.
- Pfiske, R. M., & Smith, K. H. Semantic categorization in a linear order problem. *Memory & Cognition*, 1979, **7**, 297-302.
- Pollio, H. R., & Reinhardt, D. Rules and counting behavior. *Cognitive Psychology*, 1970, **1**, 388-402.
- Pollio, H. R., & Whitacre, J. Some observations on the use of natural numbers by preschool children. *Perceptual and Motor Skills*, 1970, **30**, 167-174.
- Rothenberg, B. B., & Courtney, R. G. Conservation of number in very young children. *Developmental Psychology*, 1969, **1**, 493-502.
- Rule, S. J. Equal discriminability scale of number. *Journal of Experimental Psychology*, 1969, **79**, 35-39.
- Schaeffer, B., Eggleston, V. H., & Scott, J. L. Number development in young children. *Cognitive Psychology*, 1974, **6**, 357-379.
- Sekuler, R., Armstrong, R., & Rubin, E. Processing numerical information: A choice time analysis. *Journal of Experimental Psychology*, 1971, **90**, 75-80.
- Sekuler, R., & Mierkiewicz, D. Children's judgments of numerical inequality. *Child Development*, 1977, **48**, 630-633.
- Shepard, R. N., Kilpatrick, D. W., & Cunningham, J. P. The internal representation of numbers. *Cognitive Psychology*, 1975, **6**, 82-138.
- Shepard, R. N., & Podgorny, P. Cognitive processes that resemble perceptual processes. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes*. Hillsdale, N.J.: Erlbaum, 1978.
- Siegler, R. S. Defining the locus of developmental differences in children's causal reasoning. *Journal of Experimental Child Psychology*, 1975, **20**, 512-525.
- Siegler, R. S. Three aspects of cognitive development. *Cognitive Psychology*, 1976, **8**, 481-520.
- Siegler, R. S. The origins of scientific reasoning. In R. S. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, N.J.: Erlbaum, 1978.
- Siegler, R. S. Developmental sequences within and between concepts. *Monographs of the Society for Research in Child Development*, 1981, **46** (Whole No. 189), (a).
- Siegler, R. S. Information processing approaches to development. In W. Kessen (Ed.), *Manual of Child Psychology: History, theories and methods*. New York: Wiley, 1981, in press. (b)
- Siegler, R. S., & Klahr, D. When do children learn: The relationship between existing knowledge and the ability to acquire new knowledge. In R. Glaser (Ed.), *Advances in instructional psychology*. Hillsdale, N.J.: Erlbaum, 1981, in press.
- Siegler, R. S., & Robinson, M. *Preschoolers' knowledge of very large numerical magnitudes*. Manuscript in preparation, 1981.
- Simon, H. A. On the development of the processor. In S. Farnham-Diggery (Ed.), *Information processing in children*. New York: Academic Press, 1972.
- Smedskind, J. Microanalysis of concrete reasoning. I. The difficulty of some combinations of addition and subtraction of one unit. *Scandinavian Journal of Psychology*, 1966, **7**, 145-156.
- Sternberg, R. J. Component processes in analogical reasoning. *Psychological Review*, 1977, **84**, 353-378.

- Strauss, S., & Levin, I. A commentary on Siegler, R. S. Developmental sequences within and between concepts. *Monographs of the Society for Research in Child Development*, 1981, 46, (Whole No. 189).
- Svenson, O. Analysis of time required by children for simple additions. *Acta Psychologica*, 1975, 39, 289-302.
- Svenson, O., & Broquist, S. Strategies for solving simple addition problems. *Scandinavian Journal of Psychology*, 1975, 16, 143-151.
- Trabasso, T. The role of memory as a system in making transitive inferences. In R. V. Kail, Jr. & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition*. Hillsdale, N.J.: Erlbaum, 1977.
- Trabasso, T. How do children solve class inclusion problems? In R. S. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, N.J.: Erlbaum, 1978.
- Trabasso, T., Riley, C. A., & Wilson, E. G. The representation of linear order and spatial strategies in reasoning: A developmental study. In R. J. Falmagne (Ed.), *Reasoning: Representation and process*. Hillsdale, N.J.: Erlbaum, 1975.
- Yoshimura, T. *Strategies for addition among young children*. Paper presented at the 16th annual convention of the Japanese Association of Educational Psychology, 1974.