DOCUMENT RESUME

ED 259 835 PS 015 153

AUTHOR Baroody, Arthur J.
TITLE Early Addition Estimates: Retrieval or Problem Solving?
SPONS AGENCY National Inst. of Child Health and Human Development (NIH), Bethesda, Md.
PUB DATE Apr 85
GRANT NICHD-HD-16757-02; NICHD-HD-16757-03
NOTE 57p.; Paper presented at the Biennial Meeting of the Society for Research in Child Development (Toronto, Ontario, Canada, April 25-28, 1985).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Addition; Cognitive Processes; Elementary Education; Elementary School Mathematics; *Estimation (Mathematics); Kindergarten Children; *Mathematics Skills; *Memory; Mental Retardation; Models; Pretests Posttests; *Problem Solving; Screening Tests; Semantics
IDENTIFIERS Critical Analysis; *Retrieval (Memory)

ABSTRACT

Limitations of the retrieval strategy dimension of Siegler's (1982, 1984) distributions-of-associations model of young children's estimation of sums are delineated, an alternative model is described, and findings of two studies designed to test key assumptions of the models are reported. In Study 1, kindergarten children with normal IQ and no formal arithmetic training were screened on prearithmetic skills, pretested 20 times over the course of 8 weeks on a set of problems to estimate their distributions of associations, given intensive arithmetic training for 8 weeks, and posttested. Transfer of learning was also assessed. Results provided little evidence of learning over the 20 repetitions of pretest problems. Pretest distributions of associations did not provide a good indication of how the children responded on the posttest. In Study 2, 30 moderately or mildly retarded subjects between the approximate ages of 6 and 20 were screened and tested to estimate their distributions of associations with a set of 16 basic addition combinations administered 20 times in seven or eight sessions over a period of 1 month. Results were similar to those from the kindergarten children, but mentally handicapped children tended more than kindergarten children to use one or two strategies that accounted for a large proportion of their responses. It was concluded that children's earliest estimates are not drawn randomly from all known numbers and that computing experience is insufficient to account for the type of estimation errors occurring before mastery of basic number combinations. (RH)

Early Addition Estimates:

Retrieval or Problem Solving?

Arthur J. Baroody

University of Rochester

Preparation of this paper was supported by Grants Numbered HD 16757-02 and HD 16757-03 from NICHD (NIH). I am indebted to Mary Reed, Margaret Stiite, and the pupils of the Council Rock Elementary School, Brighton, NY and to Rena Gaspard, the teachers, and children of the Foreman Center, B.O.C.E.S. I, Monroe County, NY, whose cooperation made this study possible. Thanks to Cathleen A. Mason and Robert S. Siegler for their helpful comments in preparing this paper. Thanks are due to Brian Burley, Kathy Breen, and Laurie Donofrio for their help in collecting the data. The author, however, assumes sole responsibility for the content of the paper.

Paper presented at the biennial meeting of the Society for Research in Child Development. Toronto, April, 1985.

Early Addition Estimates:

Retrieval or Problem Solving?

How do young children estimate the sums of single-digit addition problems before they learn the "basic addition facts"? Do they resort to guessing, recalling a stored association, or problem solving? Siegler (Siegler & Robinson, 1982; Siegler & Shrager, 1984) has advanced a model that suggests that estimates are retrieved from previously learned associations. In this paper, I delineate limitations of this model, describe an alternative model, and report on two studies designed to test the conjectures of these models.

It should be noted that this paper will deal with only one aspect of Siegler's model: the retrieval strategy. That is, the paper will focus on Siegler's description of children's responses to addition problems when they are asked not to count or use their fingers but to state the first answer that comes to mind. The paper does not address another important aspect of Siegler's model: the issue of strategy choice. More specifically, the paper does not deal with what strategy ("counting fingers," "fingers," "counting," or "retrieval") the child would naturally select or why one strategy would be preferred over another.

The Distributions-of-Associations Model

According to Siegler's model, estimates are not random but are influenced by prior knowledge. The probability of retrieving a particular answer is proportional to the associative strength between that answer and the problem. The distributions of associations describe the associative strengths among problems and various potential answers (see Figure 1).

Insert Figure 1 about here

Even before they begin computing sums, children have some basis for making estimates. More specifically, the formation of a distribution of associations is influenced by prior knowledge of numbers and the counting string. One piece of prior knowledge is that numbers as a general class are appropriate answers to addition problems. A second kind of prior knowledge is number-after relationships in the counting string. Thus, when a problem such as 2 + 4 is first presented to a child, it triggers an association with numbers with which a child has had previous exposure. If a child was familiar with the numbers 1 to 10, the response to 2 + 4 would be selected from this set of numbers. A child would be especially likely to advance the number in the count string that comes after the last addend. For 2 + 4, the child would most frequently respond "5," because 4 and 5 already share a relatively high degree of association.

Siegler describes an important exception to the straightforward, associative-retrieval account. For "descending problems" such as 4 + 2, the most prevalent estimate was the number in the counting string after the <u>first</u> addend (i.e., "5" in the case of 4 + 2). To account for these results, Siegler and Shrager (1984) hypothesize the introduction of a reasoning process. "The last addend in an addition problem may always activate its immediate successor as a potential answer. However, other knowledge that preschoolers have, namely that answers to addition problems should be at least as great as the larger addend, may prevent them from stating counting-string associates as answers on descending-series problems" (p. 265). On 4 + 2, for instance, the child would not say 3 as an answer because semantic knowledge disqualifies numbers less than 4.

An important assumption of the distrbutions-of-associations model is that children learn the answers they state. Thus when a child responds with an estimate of 5 to 2 + 4, the association between 5 and 2 + 4 is strengthened. Moreover, children's initial computational efforts may be error prone and so associations between 2 + 4 and various incorrect responses such as 3 or 4 are strengthened. Because an off-by-one error is a common calculational error, the incorrect response 5 is especially likely to be strengthened. In cases such as 2 + 4, where the counting-string associate is also a frequent computational error, the number after the last addend is an extraordinary likely response.

At some point, though, children learn to compute efficiently. As the child computes the correct sum more and more often, the association between a problem and its correct sum is gradually strengthened. Eventually the correct answer is produced so frequently that the association between the problem and the correct sum becomes preemptively strong.

A Critique of Distributions-of-Associations Model

Theoretically and empirically, the distributions-of-associations model has a number of weaknesses. First, what little evidence there is of children's earliest estimates is not consistent with the model's assumptions concerning the initial state of the distributions of associations. To reflect the assumption that children know enough to respond with a number, the computer simulation of the model starts with a set of minimal (absolute) associations between each problem and each possible answer, which are arbitrarily defined as the whole numbers 1 to 12. Each possible answer to a problem is assigned an absolute associative strength of .01 (a relative associative strength of about 1/12 or .08). To reflect the assumption that counting-string associates are a factor with ascending problems and ties, the associative value of the number after the last addend is, in effect, increased. Thus, for ascending problems and ties, at least, the model implies that a child's initial estimates for any given problem will include the whole range of known numbers and that, except for the counting-string associates, all the known numbers are equally likely to be given as initial estimates.

Currently no data have been collected that directly test these assumptions. However, the cross-sectional data collected by Ilg and Ames (1951) suggest that initial estimates do not range more or less evenly over all known numbers. This research found that early estimation errors were "more an error of method than an error of answer" (p. 10). More specifically, younger but not older subjects simply stated one of the addends or added one to the larger addend. In other words, it appeared that the younger subjects were not recalling an incorrect sum but were manufacturing an answer based on how they interpreted the task.

Second, the model does not provide a fine-grained account of Siegler and Shrager's (1984) estimation data summarized in Figure 1. For example, the model specifies why the estimate of 5 for 2 + 4 increases in relative frequency (5 is not only a counting-string associate but a common calculation error for the problem). Unclear is why an estimate of 4 is so infrequent (.02) in comparison to responses such as 3 or 8 (both .07). Because 4 is closer to the sum than 3 and because children tend to undercompute rather than overcompute, it would seem that 4 would be a more likely error than either 3 or 8.

Consider other specific aspects of the Figure 1 data that are inconsistent with the predictions of the model. For 5 + 5, the common off-by-one computing error should make 9 (.04) a much more probable response than 0 (.04) or 5 (.07). Moreover, because computing errors tend to be asymmetrical around a sum (children tend to undercount rather than overcount), 9 should have a somewhat higher associative strength than 11 (.04). The same kinds of logic apply to the following cases:

3 + 5,	7(.14),	5(.13);
4 + 3,	6(.09),	8(.09);
4 + 4,	7(.07),	6(.07);
4 + 5,	8(.11),	10(.11);
5 + 3,	7(.16),	5(.18);
5 + 4,	8(.11),	4(.11).

The discrepancy between prediction and data in Figure 1 is most striking in several cases where the counting-string associate and off-by-one computing error should have produced a dramatic difference. For 2 + 2, 3 has an associative strength of only 0.5, while 2 has a nearly equal value of .04 and 7 has an equal strength of 0.5. For 3 + 2, 4 (.11) should have an associative strength considerably higher than 2 (0.9). (Because it is a counting-string associate and closer to the sum, 3's relative associative value of .11 should also have been rather greater than 2's.)

Third, the role of semantic knowledge is invoked inconsistently and in a manner that does not agree with a significant portion of the empirical data. In the case of descending problems, the counting-string associate is not usually the most common estimation error. To account for this anomaly, the distributions-of-associations model suggests that semantic knowledge of addition disallows estimates less than the larger addend. For example, for 5 + 3, knowledge that addition implies incrementing would override the choice of 4—the counting-string associate. The model does not explain why semantic knowledge is not also used to check the estimates to ascending problems and ties. Moreover, the hypothesis does not account for the fact that a significant proportion of the responses to descending problems were "impossible sums" (rectangle-enclosed responses in Figure 1). Note, for example, that responses to 5 + 1 of 2, 4, and 5 had a cumulative frequency of .15; responses to 5 + 2 of 2, 3, 4, and 5 had a cumulative frequency of .45; responses to 5 + 3 of 2, 3, 4, and 5 had a cumulative frequency of .40; and responses to 5 + 4 of 4 and 5 had a cumulative frequency of .32. Indeed, for the 10 descending problems, nearly a quarter of the responses were not greater than both addends (in Figure 1, the mean cumulative frequency of answers equal to or less than the larger addend for these 10 problems was .23).

Fourth, the distributions-of-associations model does not take into account possible qualitative differences in estimation ability or allow for qualitative changes in children's estimates. The errors of method described by Ilg and Ames (1951), for example, seem to be qualitatively different. A tendency to repeat an addend would seem to be a very early estimation method—a response bias used by children who know very little about arithmetic. More advanced children might rely on the qualitatively different strategy of adding one to the larger addend to honor their knowledge that addition involves incrementing.

Baroody (1983) found that children of different levels of addition ability appeared to make qualitatively different types of estimates. Children who had to be shown a concrete counting-all procedure repeatedly simply stated the counting-string associate of the larger addend. Children who knew one sums and mastered a concrete procedure after only one demonstration tended to use an add-several strategy (e.g., estimating that 3 + 5 would add up to 7). A few children in this longitudinal study appeared to develop qualitatively different estimation strategies during the course of the study. Moreover, Baroody (1985) reports a case of a kindergartner who did not know zero sums. After brief instruction, the girl was able to respond efficiently even to novel zero problems. This evidence suggests that the girl learned a general zero (identity) rule that could be used to generate sums to previously unpracticed combinations.

The empirical basis for the distributions-of-associations model has two weaknesses: Siegler and Shrager's (1984) study was not longitudinal in design and the distributions of associations were tallied across subjects. Thus the study was not designed to test the assumptions concerning children's earliest estimates (what Siegler and Shrager refer to as the learning phase of the computer simulation of the model). Moreover, the study may have lumped together children who were at developmentally different levels and who were responding in qualitatively different manners. This would help to account for the discrepencies between the predicted and the actual relative associative strengths of some estimation errors, why the counting-string associates were not always the most common type of estimation error for descending problems, and why so many responses to descending problems were less than or equal to the larger addend. The nature of Siegler's analysis can also account for the discrepancy between the model and existing longitudinal and case study data: Why initial estimation errors may consist of stating one of the addends, and why progress in estimation ability appears to involve qualitative changes.

An Alternative Model

The problems of the distributions-of-associations model are avoided by an alternative model that suggests that semantic and associative memory work together to account for the development of mental arithmetic, including early estimation performances with single-digit problems. The alternative model shares some of the same assumptions as the distributions-of-associations model, but also differs from it in some fundamental ways. The alternative model—like the distributions-of-associations model--posits that a protracted computing experience is an important stage for mastering the basic number combinations.

Computing experience may well help to build up associated responses to specific arithmetic problems. However, unlike the distributions-of-associations model, the alternative model suggests that the amount of practice may not be the determining factor in mastering the basic number combinations. It may be that a sum is computed repeatedly with little or no impact on long-term memory (LTM). Like an unimportant telephone number that is needed for only the moment, the problem and computed sum are held in working memory but not entered into LTM. Sometimes though, as with a telephone number that forms an interesting and easily definable pattern, a combination may make a mark on LTM without conscious effort. The ties may be particularly susceptible to incidental learning because of the abudance of readily recognizable patterns. Common models are fingers, dice, and symmetrical objects such as wheels on a car. As with important telephone numbers, children may consciously make an effort to memorize some combinations. Interest, which may be motivated by a need for approval or fear of disapproval, may provide a powerful incentive to learn combinations. Very few repetitions may be required if a child has a great interest in mastering a combination (cf. Thorndike, 1922).

The alternative model suggests, then, that practice does not automatically leave traces (build up associative strengths) in LTM. At some juncture, a specific association may be noted between a problem and a sum and entered into LTM. Thereafter, practice may serve to strengthen that association so that gradually that response becomes more and more automatic.

The alternative model posits another way in which computing may lead to mastery of basic number combinations. Computing experience may be an important vehicle for discovering arithmetic relationships such as commutativity, zero rules (e.g., N + 0/0 + N = N), or one rules (e.g., N + 1/1 + N = the number after N in the number sequence). These arithmetic relationships may not only serve to check retrieved responses (as is the case of descending problems in the distributions-of-associations model), they may play an integral part in constructing responses. For example, a child who has learned a zero rule can quickly reason out that 0 + 9 equals 9 and that 42 + 0 equals 42 without having previously practiced these combinations or having stored a specific fact in LTM. Repeated practice may serve to automatize the use of such relationships.

According to the alternative model, then, the amount of computing may be less important than how the child uses the computed results. Computing can lead the child to discover relationships and strengthen semantic knowledge that may underlie number combination facility. Practice may serve to automatize such knowledge so that it may be applied to arithmetic problems efficiently. However, computing per se is not necessary for learning relationships and, in some cases, may actually interfere with mastering the basic combinations. Given the readiness to learn, a teacher's comment might help the child to notice and internalize a relationship. Especially where children are just mastering a computing procedure, the process of computing may absorb so much attention that important arithmetic regularities or relationship may go unnoticed.

The alternative model is based on the (little) existing data on children's early estimation performance and has some important advantages over the distributions-of-associations model. First, the alternative model avoids the issue of how to define the initial state of the distributions of associations. The alternative model suggests that the nature of initial estimates will depend on the developmental sophistication of the child's semantic knowledge and knowledge of the counting string. For an unknowledgeable child, the novel stimulus 2 + 4 = ? may mean little. Such a child might well fall back on some kind of response bias to manufacture an answer (e.g., repeat one of the addends). More sophisticated children may assimilate previously unseen and uncalculated problems to their general conceptual knowledge of addition and manufacture a more reasonable estimate. For instance, given the novel problem 2 + 4, a child might realize that addition involves incrementing and that the sum has to be greater than either addend. Because the child is very familiar with next-number relationships, the child advances "5" as a response. Such a model avoids hypothesizing that all known numbers are initially—with the exception of the counting-string associates—equally likely responses to ascending problems and ties.

Second, the alternative model posits a more parsimonious use of cognitive resources. When presented a problem, the child draws upon already existing resources. The estimation strategy is based on extant knowledge of arithmetic and can manufacture an answer by <u>operating on</u> the stored number sequence. In brief, the child can quickly manufacture an answer without building up and storing in LTM numerous (incorrect) problem-sum associations.

Third, the alternative model explicitly invokes use of semantic knowledge regardless of problem type. According to the alternative model, a child will tend to use the same strategy for ascending problems and ties as well as well as descending problems. That is, depending on the level of development, the child will seize on one of the addends, add one to the last addend, add one to the larger addend, or add several to the larger addend for all three types of problems. Thus, the alternative model predicts that some children will generally give impossible estimates regardless of problem type, others may tend to give impossible sums for descending problems only, and some will give very few or no unreasonable estimates.

Fourth, the alternative model can account for qualitatively different types of estimates and hence, for variability among subjects. The relative frequencies of estimation errors are not shaped by how frequently a child miscalculates a problem and arrives at erroneous answers. In large part, estimation errors are the result of a systematic strategies based on existing knowledge. Estimation errors arise because the underlying knowledge is not complete and accurate or because the strategy does not reflect completely and accurately the underlying knowledge. Differences in underlying knowledge or how it is applied account for the various kinds of systematic errors different children make. This explains why some children may choose as an estimate a number less than or equal to the larger addend but other children avoid such a response.

Fifth, because estimates are the result of a strategy based on children's conceptual knowledge rather than something retrieved from a repertoire of specific numerical associations stored in LTM, the alternative model can better account for qualitative changes in development. More specifically, as their understanding of addition evolves, children should become capable of more sophisticated mental estimates. That is, a child's most probable response (estimate) might evolve in ways that are independent of the child's previous response history ("distribution of associations"). Thus, as they become more sophisticated, a child's estimation strategies might change from naming one of the addends to incrementing by one the first or second addend. Later the child might increment the larger addend by one.

Finally, more advanced knowledge would specify that N + 1/1 + N and N + M/M + N (where N and M are greater than 1 and N is greater than or equal to M) would have to be treated differently. As a result, a very sophisticated level of estimation ability would entail responding to N + M/M + N problems with incrementing by several while still responding to N + 1/1 + N problems with incrementing by one. The evolution of estimation errors, then, may have very little to do with the types of computational errors children make.

The remainder of this paper reports on two studies that were designed to test directly key assumptions of the distributions-of-associations model and the alternative model. A training study involved normal IQ children of kindergarten age before they received any formal arithmetic training. A second study involved mentally retarded children. Because their arithmetic learning is often characterized as rote rather than meaningful learning, mentally handicapped children's estimates might more readily be explained by the distributions-of-associations model. To address the issue of qualitative differences among children, the distributions of associations were gauged for individual subjects. To address the issue of qualitative development, subjects in both studies were then given computational training and were tested on estimation tasks. This paper reports on the portions of the studies currently completed: the pretest and posttest results of the kindergarten children and the pretest data of the mentally handicapped children.

Study 1

Method

Participants

Children for Study 1 were drawn from a kindergarten class in a school that serves a working- to upper-class suburb of Rochester, NY. From a class of 21, four children were excluded from the study because they persistently tried to compute rather than estimate sums. The sample included in the study consisted of 7 boys and 10 girls (ages 4 years and 9 months to 5 years and 11 months).

Design

After a familiarization session, the children were screened on prearithmetic skills to account for any deficiencies that might impede the addition training. The children were also screened on basic addition ability to determine if a child had a mental computing algorithm and if a child could compute sums using concrete objects. The screening results provided an indication of the children's readiness for arithmetic or level of arithmetic ability. The children were pretested on a set of problems to estimate their distributions of associations. The set of problems was administered 20 times over the course of 8 weeks. One addend of the test problems ranged from 6 to 9. Thus the test problems were of a type that a child was not likely to encounter outside of the experimental setting. The subjects were required to make mental estimates—computing was not allowed. The children were then given intensive arithmetic training for a period of 8 weeks with a different set of problems—using problems with addends 0 to 5 only. The training was done in small groups of 3 or 4 about twice a week. Training focused on helping children to understand addition and to compute accurately. The trainer avoided pointing out specific short-cuts such as a zero rule. Afterward, the children were retested on the (nonpracticed) test items. In effect, the posttest gauged whether the children would or would not transfer the learning promoted by the training sessions. Each problem of the test set was administered once on the posttest. Transfer of learning was also gauged on the posttest by administering unfamiliar three-digit problems of the type 0 + 0 + N, N + 0 + 0, 1 + 1 + N, and N + 1 + 1, where N ranged from 2 to 5. All testing was done on a one-to-one basis and audiotaped.

Tasks

Familiarization Task. In groups of three or four, the children met with a tester and played a car race game. On their turn, the children would roll dice with 0 to 5 dots. If necessary, the tester prompted, "How many dots are there altogether?" The child could then move his or her car a number of spaces equal to the sum. Children who did not know what to do were shown a concrete counting-all procedure (count the dots on the first die and then continue the count as the dots on the second die were enumerated).

Screening for prearithmetic skills. Prearithmetic skills tested were oral counting to 15, enumerating counting sets of up to 15 objects, producing (counting out sets of) 1 to 5 objects, reading single-digit numerals, and applying the $N + 1 > N$ rule.

For the oral-counting task, the child was asked to count a card with 15 stars. If the child did not successfully generate the number sequence to 15, the task was readministered in a make-up session. The child's best counting performance was scored according to the following criteria: 3 points for a correct count to 15; 2 points for a correct count to 10 and one error from 11 to 15; 1 point for a correct count to 10 but more than one error in the teens; and 0 points if the child could not count to 10 correctly.

For the object-counting task, the child was asked to enumerate 5, 7, and 15 stars attached to cards in regular arrays. A trial was readministered in a make-up session if the child was incorrect. The scoring criteria were: 3 points, if the child correctly enumerated all the sets on the first try; 2 points, if the child was correct on the first effort for two trials and, for the third, was correct on the make-up or used 1-1 pointing but made a tagging error; 1 point, if the child was unable to enumerate one set; and 0 points, if the child was unable to enumerate two or three sets.

The counting out sets task involved having the subject produce one to five pegs from a disk of pegs. The criteria were: 3 points, 80% accuracy or better; 2 points 50-79% accuracy, 1 point, 20-49% accuracy, and 0 points for 0-19% accuracy.

The $N + 1 > N$ rule was evaluated by asking the subjects, in random order, which was more 1 or 0, 2 or 3, 5 or 4, 6 or 7, and 9 or 8. If the child was not correct, on all five trials, the task was readministered in a make-up. The scoring criteria were: 3

points for 5 correct responses on the intial test; 2 points for 5 correct responses on the make up; 1 point if one error persisted on the make up; and 0 points if 2 or more errors persisted.

In the reading numerals tasks, the children were asked, in random order, to read the numerals 0 to 9. If a child was incorrect, a make up trial was administered later. A correct response in the make up scored as one-half correct. The criteria were 3 points for no errors; 2 points for one-half to one error; 1 point for one and one half to two errors; and 0 points for three or more errors.

Screening for addition ability. The tester administered five change-type word problems involving the problems 1 + 2 + ?, 2 + 4 = ?, 3 + 5 = ?, 4 + 1 = ? and 5 + 3 = ?, in random order. The children were encouraged to solve the problem mentally first and, if need be, by using concrete objects. Children who did not know how to use objects to compute sums were taught a concrete-counting-all procedure: count out the number of blocks that represents each addend and then count all the blocks put out. One point was scored for correct answers regardless of solution procedure. No points were awarded for a trial on which the child had to be shown or helped with a concrete counting-all procedure. Scores could range from 0 to 5.

The children were also evaluated on proficiency with basic number facts with addends ranging from one to five. Using the estimation task described below, the children were given the following problems in random order: 1 + 3, 1 + 4, 2 + 1, 5 + 1, 2 + 3, 2 + 4, 3 + 5, 3 + 2, 4 + 3, and 5 + 2. The children were asked to respond to the problems quickly, without calculating. One point was awarded for each correct and automatic response. Scores could range from 0 to 10. (The number-fact screening also served as a familiarization round for the estimation test administration in a latter session.)

Estimation Task. The estimation task took the form of a "Quick Think" Game. The tester explained, "Let's play the quick think game and give you a chance to win some prizes. I'll give you some adding problems, and you tell me quickly what you think the answers are. In this game you don't have to use blocks or fingers to figure out the answers—just think what the answer is. If you answer all the problems quickly —before the bell on the timer rings, you win a prize. [The child was shown the timer.] I'll keep score and if you give good answers, you'll win even more prizes. There's one special rule in this game: You have to keep your hands folded. Remember, think hard and answer quickly."

A trial was readministered at a later time if the child tried to compute the sum or if the child answered correctly but took more than 3.0 second to respond. On 21 pretest trials, children responded with sums greater than 20 (e.g, "100") or with unusual answers (e.g., "8 1/2"). These trials were readministered at a later time because it was judged that the child was not making a genuine estimation effort. Three posttest trials were readministered for the similar reasons. This procedure basically worked in favor of the distributions-of-associations model and against the predictions of the alternative model.

The test problems included both ascending (small-addend-first) problems and descending (large-addend-first) problems of three different types: zero, one, and large problems. The ascending zero (0 + N) problems were 0 + 6 and 0 + 9; descending zero (N + 0) problems were 7 + 0 and 8 + 0; ascending one (1 + N) problems were 1 + 7 and 1 + 8; descending one (N + 1) problems were 6 + 1 and 9 + 1; ascending large (M + N) problems were 5 + 6, 3 + 7, and 4 + 8, and descending large (N + M) problems were 7 + 4, 8 + 5, and 9 + 3. The problems were chosen so that the sums are evenly distributed from 6 to 13. The three-digit transfer problems (0 + 0 + 2, 0 + 0 + 5, 3 + 0 + 0, 4 + 0 + 0, 1 + 1 + 3, 1 + 1 + 4, 2 + 1 + 1, and 5 + 1 + 1) were administered after the test problems on the posttest. The test and transfer problems were presented in random order. The speed of posttest test and transfer problems were rated as follows: a reaction time (RT) of less than 1.00 second was scored as 0, a RT between 1.00 and

1.99 was scored as 1, a RT between 2.00 and 2.99 was scored as 2, and a RT of 3 seconds or more was scored as 3.

Training

The training consisted of four phases, each phase lasting two weeks. In Cycle 1, each addend of the addition problems was represented by dots within a 7.62 x 7.62 cm box on a 5" x 8" card. The dots were arranged in a regular pattern as on a die. An empty box represented zero. Below each box the cardinal value of the addend was indicated by a numeral. A plus sign was positioned between the two numerals. The deck of 36 cards was shuffled and used to play a variety of math games. On their turn, the children would draw a card and asked the sum of the problem represented. If the child did not respond or used her own strategy to generate an incorrect answer, the trainer had the child count the two sets of dots.

In Cycles 2 to 4, the problems were represented on 3" x 5" cards by numerals only. In Cycle 2, blocks were provided and, if needed, the child was instructed or helped to use a concrete counting-all procedure. In Cycle 3, an abacus-like device with five red markers on one side and five green markers on other was provided. If needed, the child could compute the sum of a problem by sliding up the appropriate number of markers to represent each addend and then counting the number of markers of both colors in the up position. In Cycle 4, nonresponders or incorrect responders were encouraged or helped to use their fingers to compute the sums of problems.

Results

Pretest

There was little evidence of learning over the 20 repetitions of the test problems that constituted the pretest. To check for learning, the number of correct responses to a problem on the first 10 trials was compared to the success rate on the trials 11 to 20. For the zero problems, 14 subjects exhibited great consistency: The difference in success rates for each half of the pretest was 0 $\pm$ .10. There was one case (S 04) in which a modicum of improvement was registered: The child, who had given no correct zero-problem sums for trials 1 to 10, had success rates in trials 11-20 for 0 + 6, 0 + 9, 7 + 0, and 8 + 0 of .20, .10, 0, and .10, respectively. Two subjects had moderate improvement that was due to their systematic but erroneous estimation strategy. S 05's success rates went from .50 to .70, .60 to .50, .40 to .80, and .40 to .70, and S 21's success rates went from .70 to 1.00, .70 to 1.00, .90 to 1.00 and .70 to 1.00. In both cases, the change could be attributed to the fact that the child used a state-the-larger-addend strategy more consistenty with <u>all</u> types of problems.

For one problems, there was only one clear case of genuine general improvement. S 14's success rates for 1 + 7, 1 + 8, and 6 + 1 went from .70 to 1.00, .80 to .90, and .70 to .90. The success rate for 9 + 1 remained constant at .80. S 03 improved slightly on 1 + 8 (.80 to .90) and 6 + 1 (.80 to 1.00), while remaining consistent on 1 + 7 and 9 + 1 (.90 and 1.00, respectively, on both halves). Two subjects improved dramatically on a single one problem but not the other three. S 04's success rate for 1 + 8 went from .30 to .80 but remained the same (at .80) for 1 + 7 and actually dropped for 6 + 1 and 9 + 1 (.80 to .30 and .80 to .60, respectively). S 09's improvement on 1 + 7 (.20 to .70) appeared to be due to the child's increased reliance on the response bias of saying 8 for all one and large problems.

For large problems, five subjects showed some improvement on isolated problems. S 03's success rate for 5 + 6 went from .30 to .90. S 04's success with 8 + 5 went from .20 to .40. S 07's success with 3 + 7 and 9 + 3 jumped from 0 to .30 and 0 to .40, respectively. S 07 showed some gain on 8 + 5 (0 to .30), and S 20 had a modest gain on 5 + 6 (0 to .20). In all, there appeared to be a minimum of learning on the pretest. It appears that the pretest can be taken as a reasonable measure of the children's distributions of associations at the beginning of school.

For combinations which they had not mastered, the kindergarten children did respond with a range of answers but not in a manner that was consistent with the predictions of the distributions-of-associations model. Typically, a large portion of their responses could be accounted for by positing a general estimation strategy or even a specific strategy (a strategy that involved choosing just one or a few numbers). In Table 1, the most accurate estimator (S 03), who was generally correct on zero and one problems, typically seemed to add several to the larger addend for larger problems. For 3 + 7, for example, the child responded "9," "10," or "11" 75% of the time (.20, .50, and .05, respectively). Indeed, this hypothesized strategy would account for nearly three quarters (.72) of all responses given for the six N+M/M+N problems. This, in itself, is not inconsistent with the distributions-of-associations model.

Insert Table 1 about here

What is difficult to explain in terms of the distributions-of-associations model is the fact that one answer accounts for nearly half (.46) of the child's responses. As can be seen in Figure 2, 10 is the most frequent response for N+M/M+N problems, except for 5 + 6. These results cannot be explained by hypothesizing the mechanical production of counting-string associates or answers whose associative strength has been built up because of counting errors. In the case of this subject, it may be that larger problems are associated with a known fact such as 5 + 5. This factual knowledge can be the basis for quickly reasoning that 11 would be a good estimate for 5 + 6 and that 10 would be a good estimate for other large problems such as 4 + 8.

Insert Figure 2 about here

Consider another case (S 15) that cannot be easily explained by the distributions-of-associations model. It appears that, for all types of problems, the child's general strategy was to add one to an addend. Moreover, she favored a specific version of this strategy: add one to the larger addend (advance the counting-string associate). As Table 1 shows, the general or specific strategy could account for over half of the child's responses. Further analysis indicated a secondary tendency. Despite knowledge of number-after relationships and magnitude ($N + 1 > N$) comparisons, this girl responded with the number one less than the larger addend on some problems. For example, as Figure 3 shows, she typically responded with seven to problems with an addend of eight. Thus, it seems that three strategies (smaller plus one, larger plus one, and larger minus one) accounted for an astounding 90% of her responses to zero problems and 97% of her responses to one and large problems. It is not clear how the distributions-of-associations model would account for the discontinuous, highly peaked unimodal and bimodal plots in Figure 3.

Insert Figure 3 about here

As Table 1 shows, six children (S 14, S 12, S 06, S 09, and S 07) apparently had a specific response bias—a particular number they favored. In three cases, a single number accounted for at least one quarter of the subject's estimates. Some of these children would favor a number with one type of problem but another for other types. Three numbers accounted for three fifths of all S 06's responses, but the preferred response shifted with problem type. Though he used 7 and 8 with some frequency (.10 and .11, respectively), he clearly favored 6 for zero problems (.40). For one problems, he relied equally on his three favorite numbers: 6, 7, and 8 (.22, .21, and .20). For larger problems, he favored seven (.25 of the time) and six and eight to a less extent (.20 and .14, respectively). A shifting preference among three number accounted for three eights of S 07's responses. For zero problems, she relied on 19 (.20) but used 12 and eleventeen with some frequency (.07 and .09, respectively). For one problems, she shifted preference from 19 (.09) to 12 and eleventeen (.15 and .14, respectively). For large problems, she used the three favorite numbers with about equal frequency (.14, .12, and .14, respectively).

Another type of response bias—simply stating an addend—was used with great consistency by two children. One child (S 21) typically chose the larger addend. The second child (S 05) tended to choose the larger addend for zero problems, chose either addend half the time for one problems, and chose the smaller addend somewhat more often on the larger problems.

As Table 1 shows, six subjects—including four of the best estimators (S 01, S 02, S 10, and S 04)—typically responded to the estimation task with an answer in the teens and typically not just any teen. Stating one of the addends as a teen appeared to be a favorite way of manufacturing an answer. For instance, note that in Figure 4, S 10 nearly always responded with an answer greater than 10—most usually with N or M + teen. For example, for 7 + 4, the boy responded "17" 85% of the time and 14 10% of the time. It is unlikely that such frequent teen responses are the result of computing errors. Most teen responses—and certainly those constructed from the addends—are greater than the sum, even in the case of the larger problems.

Insert Figure 4 about here

Qualitative analyses indicates that the children sometimes responded with a series of numbers from the standard sequence that were basically unrelated to the problems presented. Consider the otherwise mysterious responses of S 05 on repetition 1 to the following four problems:

1 + 7 -- 5;

9 + 1 -- 6;

0 + 9 -- 7;

9 + 3 -- 8.

Seven children appeared to manufacture answers in this way on only one or two occasions. A handful of children (S 14, S 06, S 20, S 07, and S 17), however, apparently used the strategy with some frequency. Indeed, as Table 1 reflects, it was S 20's most frequently used strategy—accounting for about one third of her responses. If pairs of numbers from the standard sequence are included in the tallies (e.g., 0 + 9—"6", 9 + 3—"7"), a generate-a-number-sequence strategy would account for half of the girls' responses (52%, 54%, and 47% of zero-, one-, and large-problem responses, respectively).

Posttest

According to the distributions-of-associations model, the distributions of asscociations for the test problems should have remained unchanged because of the absence of computational practice. Thus the distributions of associations gauged by the pretest should be predicative of the posttest responses. However, the pretest distributions of associations did not provide a good indication of how the children responded on the posttest. Consider first the zero problems. There were ten subjects who were correct on less than 90% of the pretest zero trials. This includes one boy (S 05) who achieved 57% accuracy on zero trials and one girl (S 21) who achieved 87% accuracy by virtue of the estimation strategy they used for all problems: choose the larger addend.

As Table 2 shows, on average, the most frequent responses to the zero problems had an estimated associative strength of about .55. Because the zero problems were not practiced, the mean associative strength of the posttest responses should be about the same for subjects who did not know the zero comb' .ions at the time of the pretest. Moreover, given associative strengths ranging from .54 to .59, the model would predict that the most frequent pretest response would be given about 22 times (.55 x 40 responses). Yet the mean associative strength of the posttest responses was dramatically lower, and the subjects responded on the posttest with an answer with the greatest associative strength only eight times. In fact, the children responded with novel answers twice as often and with their third-most-frequent pretest answer more often than their favorite pretest response.

Insert Table 2 about here

Furthermore, the distributions-of-associations model predicts that the higher the associative strength of an answer, the more likely it is that a child will respond with that answer. Thus, favorite responses on the pretest that had relatively high conditional probabilities should be more likely on the posttest than a favorite response that had a relatively low associative strength. Thus the Glass rank-biserial pretest-posttest correlation should be moderate (.4 to .6), if not high. However, except for 7 + 0, the correlations were low for the ten subjects who initially did not know the zero combinations. The correlations for the seven children who genuinely made progress are overwhelmingly negative: -0.26, -1.00, -1.00, and -1.00, respectively.

The discrepancies between the pretest and posttest zero-problem data are due to the fact that the subjects were considerably more accurate on the posttest than on the pretest. The ten subjects responded correctly to 19% of the pretest zero trials and 95% of the posttest trials. Their mean absolute error for the zero trials fell from 2.9 on the pretest to 0.2 on the posttest. However, three subjects improved their accuracy in zero problems by indiscriminantly using a state-the-larger addend with consistency. S 05 and S 21 had used this strategy regularly on the pretest; S 18 apparently adopted this strategy for the posttest. If these three cases are excluded from the analysis, the accuracy rates for the remaining seven children improved from 6% correct on the pretest to 96% on the posttest.

Apparently, the seven subjects learned an $\underline{N} + 0/0 + \underline{N} = \underline{N}$ rule. As Table 3 shows, these subjects (S 04, S 06, S 07, S 12, S 15, S 17, and S 20) consistently responded to the zero problems incorrectly on the pretest. Without practicing these problems, they were consistently accurate on the posttest. Moreover, unlike S 05, S 18, and S 21, these seven subjects used a "pick-the-larger-addend" strategy with only zero problems. The results of the three addition transfer problems also depicted in Table 3 support this analysis. The subjects in this study probably had never been exposed to written three-digit addition problems. Yet, the seven childen under discussion responded quickly and accurately to these novel zero problems. Moreover, they used a $\underline{N} + 0/0 + \underline{N} = \underline{N}$ rule selectively.

Insert Table 3 about here

The results delineated in Table 2 for the 11 subjects who initially did not know the one combinations and the 16 children who initially did not know the larger combinations were also difficult to reconcile with a distributions-of-associations model. In both cases, the mean associative strengths of the posttest responses were considerably lower than were those of the most frequent pretest responses. This is due to the fact that, with the one problems, responses with very low or zero conditional probability were nearly as frequent as those with the greatest conditional probability. For larger problems, responses with the third greatest, very low, or no conditional probability each outnumbered those with the second greatest associative strength. Indeed, responses on the pretest estimated to have no associative strength were nearly as frequent as those deemed to have the greatest. Finally, for both one and large problems, the generally low Glass rank-biserial correlations suggest that the degree of associative strength was not especially predictive of a child's posttest response.

The discrepency between pretest and posttest results for one and large problems is not as easily explained as in the case of zero problems. For both one and large problems, estimates were somewhat more accurate. For the 11 subjects who initially did not know the one combinations, the mean number of correct responses on the pretest and posttest were .22 and .30, respectively. The pretest answers of these subjects differed from the correct answers, on average, by 2.9. On the posttest the mean absolute error dropped to 1.8. For the large problems, the mean number of correct responses on the pretest and posttest were .05 and .12, respectively. The margin of error fell from 4.0 on pretest to 3.2 on the posttest.

For one- and large-type problems, there was little evidence that supported the alternative model's contention that qualitative changes in strategy account for improvement in estimation ability. Though specific estimation strategies did change, the children's hypothesized general estimation strategy typically remained unchanged. For example, the best estimator on the pretest (S 03) not surprisingly continued to use the general approach of adding several to the larger addend (5 of 6 or 83% of the large-problem posttest trials). However, she stopped favoring the specific number 10 (a new and more reasonable favorite, 11, was given 50% of the time). State a teen remained a favorite approach for large problems for children on the posttest. For S 01, S 02, S 10, and S 04, the strategy accounted for 75%, 100%, 100% and 40% of their incorrect responses, respectively. All four of these children, however, no longer relied on the specific and somewhat mechanical strategy of changing one of the addends into a teen. Stating a particular number apparently remained a favorite strategy for three children. S 14 responded with 9 on 4 of the 6 large problem trials, S 12 responded with 10 to large problems 50% of the time. S 09 used 8 or 10 83% of the time with large problems. S 14 and S 12 used the same particular numbers that they had used on the pretest, and S 09 dropped 6 as favorite, continued using 8, and also adopted the more reasonable estimate 10. State an addend remained the predominant strategy for two subjects. S 21 used this approach for all types of problems with 100% consistency. S 05 used this strategy on all 14 posttest trials but one. In terms of a specific strategy, S 05 actually appeared to regress. Though on the pretest she relied on stating the larger addend, on the posttest she always seized upon the last addends.

Three subjects appeared to have switched the type of general strategy they used, but only in one case might this be considered an improvement. S 09, who on the pretest seemed to rely on stating particular numbers (6 and 8) for zero and one problems, on the posttest rather consistently gave the larger addend as an answer for all eight zero and one trials. S 18, who had earlier relied almost exclusively on adding one to an addend—especially the larger addend—relied exclusively on stating an addend when she encountered one and large problems on the posttest. Indeed she nearly always simply chose the first addend (75% of the one trials and 83% of the large trials). S 17, who had relied on a state-a-teen strategy on the pretest for all types of problems, switched to an add-one-to-the-larger approach on the posttest for (3 of the 4) one problems and (3 of the 6) large problems.

Comparisons With Screening Results

Table 4 compares the quality of the subjects' pretest and posttest estimates with the screening results. On the pretest, five subjects demonstrated considerable flexibility in their estimation response—that is, they differentiated among zero, one, and large problems. Two children demonstrated some flexibility (S 04 responded in qualitatively different way to larger problems, and S 09 differentiated between zero and other types of problems). Most subjects tended to use a (hypothesized) strategy that was reasonable—that generated answers greater than the larger addend (or, in the case of zero problems, greater than or equal to the larger addend). As Table 4 shows, most subjects responded with reasonable estimates on a majority of the trials. A handful of subjects appeared to rely on a strategy that violated the concepts that addition involves incrementing and that the sum must be greater than the larger addend. Though she tended to favor 12, 19, and the made-up number 11-teen, one child (S 07) otherwise did not appear to respond systematically.

Insert Table 4 about here

As hypothesized by the alternative model, most children either did or did not avoid impossible sums for both ascending and descending problems. On the pretest, five kindergartners (S 02, S 01, S 10, S 03, and S 13) almost never gave impossible sums. Eight children gave a roughly equal number of impossible sums for ascending and descending problems. Two subjects (S 12 and S 06) gave somewhat more impossible sums for descending problems than for ascending problems (.44 vs. .34 and .64 vs. .51 respectively). Two participants (S 14 and S 18) responded with considerably more impossible answers to descending problems than to ascending problems (.31 vs. .14 and .94 vs. .14, respectively). In the case of S 18, at least, the discrepency could be attributed to the estimation strategy the child seemed to be using: adding one to the last addend.

As Table 4 shows, children who tended to be better estimators generally had sound prearithmetic skills, while the poorest six estimators tended to have weaknesses. Note especially that differences in the use of the $N + 1 > N$ rule. Furthermore, six of the seven flexible estimators were quite successful (4 or the maximum 5 correct) on the word problems. On the other hand, the poorest estimators tended to have difficulty with the addition word problems. Four of these children could not solve a majority of the trials on their own—even though they were shown or helped with a concrete-counting-all procedure after each unsuccessful attempt. It should be pointed out that the two children (S 21 and S 05) who mechanically stated an adend for most of their estimates on the pretest and the posttest were among the children with the least amount of pre-arithmetic and general arithmetic ability. This evidence is consistent with the argument that a state-an-addend response bias is a very early estimation strategy.

Study 2

Method

Participants

Children for Study 2 were drawn from 11 classes in an upstate New York county-wide special education service agency. From this subject pool, a total of 30 children were identified for the study. Qualifying children passed a screening test that indicated a readiness for arithmetic training but had not yet mastered the basic number combinations. There were an equal number of males and females. The sample consists largely of moderately retarded children: 24 children have IQs ranging between 31 and 49. Six children are classified as mildly retarded (IQs ranging from 52 to 66). The IQs were taken from school records and, for the most part, were scores on the WISCR-R or Stanford-Binet test. Children ranged in chronological age from 6 years and 10 months to 20 years and 10 months.

Design

The children were screened on prearithmetic and basic addition to identify children whose skill level was too low or too high for the study. The addition screening also served to gauge the subjects' level of arithmetic ability. The subjects were then tested on a set of problems to estimate their distributions of associations. A set of 16 basic additions combinations were administered 20 times in 7 or 8 sessions over a period of a month.

Tasks

Screening for prearithmetic skills. To be included in the study, the subjects had to demonstrate competence in using the $N + 1 > N$ rule with numbers 1 to 5, reading numerals to 10, producing sets of 1 to 5 objects, and enumerating 1 to 12 objects.

The $N + 1 > N$ rule was assessed by randomly presenting a child with eight number pairs, such as 3 or 2, and 4 or 5, and requiring the child to pick the bigger number. Criterion was seven of eight correct ($p < .05$, Sign test). If a child did not meet criterion on the first try, the task was readministered at a latter time.

The children were presented numerals in random order and asked to read them. The criterion for the numeral-reading task was 10 of 10 correct.

In the context of a game, children were asked to count out 3 and 6 miniature cowboys from a pile of cowboys. If a child produced the incorrect number of objects, the trial was readministered. Two points were scored for a correct initial response; one point for correct response on a second try. Criterion was defined as 3 or 4 points.

There were two enumeration trials in which the child counted stars attached to a 5" x 8" card. One card had 6 stars; the other had 12. If a child incorrectly enumerated a set, the trial was readministered. Two points were granted for a correct initial response; one for a correct response on the second try. Competence was defined as 3 or 4 points.

Addition sceening. The addition screening consisted of four tasks: 8 combine-type word problem, 8 change-type word problems, 12 concrete addition problems, and 16 commutativity questions. The addition screening was done over three 20 minute sessions. The trials for a task were divided between two sessions. Session 1 consisted of combine-type word problems (Trials 1 to 4), change-type word problems (Trials 1 to 4), and commutativity questions (Trials 10 to 8). Session 2 consisted of combine-type word problems (Trials 5 to 8), change-type word problems (Trials 5 to 8), and concrete addition problems (Trials 1 to 6). Session 3 entailed concrete addition (Trials 7 to 12) and commutativity questions (Trials 9 to 16).

The combine-type word problems were embedded in a game in which the child noncontingently won money. The tester placed a number of play dollars equal to the first addend in his or her right hand (the child's left side) and said, "I have X dollars in this hand. See?" Next the tester closed his or her right hand and repeated the process for the second addend with the left hand. Then the tester clasped his or her hands together and said: "Now I put the dollars together. How much are X and Y altogether?" A trial was readministered if the child did not respond or could not remember the addends. A practice trial of 1 + 1 was administered before the test trials. If the child responded incorrectly on the pratice trial, the tester commented, "No, I have two altogether. I started with one here and one here, so there are two in my hands altogether." The subjects were allowed to mentally compute the sum or use their fingers. The trials for Session 1 consisted of 1 + 2, 4 + 1, 2 + 4, and 5 + 3; for Session 2, 2 + 1, 1 + 4, 4 + 2, and 3 + 5. Within a session, trials were administered in random order. Scores could range from 0 to 8 correct.

The change-type problems were presented as stories. The tester would read, for example, "After school, Cookie Monster runs home for his snack. His mom gives him four cookies and he sneaks one more. How much is four and one more altogether?" A problem was reread if the child did not respond or could not remember the addends. The subjects were allowed to mentally compute the sum or use their fingers. The trials for Session 1 consisted of 2 + 1, 1 + 4, 3 + 5, and 4 + 2; for Session 2, 1 + 2, 2 + 4, 4 + 1, and 5 + 3. Within a session, the trials were given in random order. Scores could range from 0 to 8 correct.

Concrete addition was checked in the context of a game by asking the child to compute the sums of problems presented verbally and in written form. The tester showed the child a card with a arithmetic sentence and said: "This card says, 'X and Y.' How much is X and Y altogether? You can figure this out any way you want: with block, fingers, or in your head." If a child was not successful in using a mental procedure, the child was asked to figure out the problem using fingers or blocks. If the child was still unsuccessful, the tester demonstrated a concrete-counting-all procedure: "What does this number [the first addend] say. X, O.K., let's put X blocks under the number. What does this number [the second addend] say. Y, O.K., let's put Y blocks under the number. Now let's find out how much X and Y are altogether by counting all the blocks." The trials for Session 2 were 1 + 2, 2 + 3, 3 + 5, 3 + 1, 4 + 2, and 5 + 4; for Session 2, 2 + 1, 3 + 2, 5 + 3, 1 + 3, 2 + 4, and 4 + 5. Within a session, problems were presented in random order. Scores could range from 0 to 12 computed correctly—either mentally or by using concrete objects. Trials on which a concrete-counting-all procedure had to be demonstrated by the tester were scored as incorrect.

Commutativity was assessed by asking the child to help a muppet Cookie Monster with his arithmetic homework. The tester wrote an equation on a magic slate board while saying, "Cookie Monster says that X cookies and Y cookies make Z cookies altogether." Directly beneath the first arithmetic sentence, the tester wrote either a commuted problem (or a different problem). The tester asked, "Does Y (A) cookies or X (B) cookies make Z cookies or a different number of cookies?" The trials for Session 1 were 3 + 4 & 4 + 3, 5 + 2 & 2 + 5, 9 + 3 & 3 + 9, 6 + 12 & 12 + 6, 0 + 4 & 3 + 4, 5 + 3 & 5 + 0, 2 + 9 & 2 + 12, and 10 + 5 & 9 + 5. The trials for Session 2 were 4 + 2 & 2 + 4, 4 + 5 & 5 + 4, 12 + 8 & 8 + 12, 14 + 5 & 5 + 14, 1 + 5 & 1 + 0, 4 + 8 & 1 + 8, 4 + 3 & 5 + 3, and 9 + 1 & 9 + 4. A correct response to commuted trials involved indicating that the sum of the second problem was the same as the first either by saying, "The same," or by stating that Z was the sum. Commuted trials were scored as incorrect if the child had to compute the sum or indicated that the sums were different. Different-problem trials were scored as correct if the child indicated that the sum would be different from Z or specified a sum other than Z. Scores for each type of trial could range from 0 + 8. Success was defined as 7 or 8 points ($p < .05$, Sign-test) for <u>both</u> types of problems.

<u>Estimation Task</u>. Basically, the same estimation task procedures used in Study 1 were used in Study 2. Trials on which a child responded with an answer greater than 20 were not readministered because it was not safe to assume that such answers were clearly performance failures. There were two descending zero problems (4 + 0 and 6 + 0) and two ascending zero problems (0 + 5 and 0 + 9), two one problems of each type (3 + 1, 8 + 1, 1 + 4, and 1 + 7), and four large problems of each type (4 + 2, 5 + 3, 8 + 6, 9 + 3, 2 + 5, 3 + 4, 5 + 8, 7 + 9).

Results

The results with the mentally handicapped children were similar to those of the kindergarten children. Again there was little evidence of learning during the estimation task. There were some improvements in the success on the last ten trials in a number of isolated instances. (S 01 improved from .0 to .6 on 1 + 7; S 02, from 0 to .3 on 0 + 5; S 03, from 0 to .2 on 1 + 4; S 05, .3 to .5 on 1 + 7, S 06, .6 to .8 on 1 + 7; S 10, .6 to .8 on 0 + 9; and S 21, 0 to .2 on both 1 + 7 and 2 + 5.) Small amounts of improvement were registered by a number of children on particular types of problems because of the estimation strategy they adopted. For example, S 09 appeared to improve on the zero problems (.6 to .7, .5 to .7, .7 to 1.0 and .7 to .9) because the child adopted a state-the-larger-addend strategy for all types of problems. Other cases are even less dramatic (S 07 improved from 1 to .4 on 4 + 0 because of a generally used state-an-addend strategy; S 19 improved from .3 to .6 on 7 + 9 because of the response bias of saying "16;" S 25 apparently improved from .3 to .7 and .1 to .6 on <u>N</u> + 1 problems the child fell into the habit of adding one to the larger addend; and S 27 went from .6 to .9 on 4 + 0 because of a tendency to state an addend for all types of problems). Two children appeared to show genuine improvement on the zero problems. S 12, who consistently got about 70% of the <u>N</u> + 0 (descending zero) problems correct, improved from .4 to .6 and .6 to 1.0 on the 0 + <u>N</u> problems. S 24 appeared to consolidate his knowledge of the zero combinations—improving from .7 to .9 on 0 + <u>N</u> (ascending zero) trials and from .6 to 1.0 and .7 to .8 on <u>N</u> + 0 problems.

Even more so than the kindergarteners, the mentally handicapped children tended to use one strategy or two that accounted for a large proportion of their responses (see Table 5). For five of the six subjects who typically responded correctly to zero and one problems (S 28, S 12, S 20, S 05, and S 10), adding one to an addend accounted for a substantial proportion of the large-problem estimates. Adding one to either addend also accounted for a large portion of large-problem estimates for subjects who were generally correct on the zero problems (S 24, S 13, and S 21).

Insert Table 5 about here

Four subjects (S 19, S 23, S 15, and S 04), did not know the zero combinations but systematically treated zero problems differently from other problems. The next seven subjects listed in Table 5 used a variety of strategies, which included generating number pairs or sequences of three or more numbers that were unrelated to the problems presented. The strategies identified were used with at least some consistency and, in most cases, accounted for at least half a subject's responses to a particular type of problem. Moreover, the last 10 subjects listed in Table 5 tenaciously used the response bias of stating an addend.

Table 6 summarizes the results of the arithmetic screening and lists data that indicate the quality of the subjects' estimates. Note that children with a better performance on the arithmetic screening tended to make estimates that were more reasonable and more accurate. The commutativity results are not listed in Table 6 because only one child (S 05)—one of the more accurate estimators—met criterion. Note especially that the children who relied upon a state-an-addend strategy (the last 10 children listed in Table 6) were, as a group, among the weakest in arithmetic ability.

Insert Table 6 about here

As in the case of the kindergarten children, the mentally handicapped children tended to respond to ascending and descending problems in a similar manner—except when their estimation strategy favored one type of problem. There were eight instances in which the proportion of impossible sums were very discrepant (differed by a factor of at least 6). Seven of these children (S 12, S 15, S 04, S 14, S 18, S 02, S 06) gave many more impossible sums to descending problems because of their tendency to add one to the last addend. One child (S 19) gave many more impossible sums to ascending problems because of the tendency to state the first addend when presented a zero problem. There were five instances (S 20, S 05, S 10, S 13, and S 03) in which the proportion of impossible sums was somewhat to moderately different (differed by a factor of 0.2 to 2.0). In three of these cases, the larger proportion of impossible sums given for descending problems can be attributed to the tendency to add one to the last addend.

Discussion

The distributions-of-associations model does not adequately account for either the kindergartners' or the mentally handicapped children's estimates. Quite frequently, when children's estimation performance is examined individually, the range of estimates is far narrower and the strength of certain incorrect responses is far greater than the model predicts. In particular, the model does not explain why the children with the lowest arithmetic ability in this study should seize upon one of the addends as an estimate or otherwise responded mechanically (to ascending and descending problems alike). Moreover, the model does not account for the qualitatively different responses among individuals or for the qualitative changes in some of the kindergarteners' responses to the zero problems.

It does not appear that the distributions-of-associations model is applicable to children's earliest estimation efforts and may only provide a partial account of later developments. Though a longitudinal study is needed to examine directly preschoolers' initial estimation efforts, the results of these studies suggest that children's earliest estimates are not drawn randomly from all known numbers as Siegler and Shrager's (1984) computer simulation model implies. Furthermore, the results of these studies suggest that the amount computing experience is insufficient to account for the type of estimation errors that occur before mastery of the basic number combinations (cf.

Ilg & Ames, 1951; Olander, 1931; Thiele, 1938; Thorndike, 1922; Wheeler, 1939). However, further research is needed to test directly the predictions of the distribution-of-associations model concerning the role of response frequency in establishing both incorrect and correct responses.

The limited range of responses and the nature of the responses suggest that the kindergarten and mentally handicapped children tended to rely on a strategy or two to manufacture estimates. Presumably these estimation strategies were shaped by the children's semantic and procedural knowledge of arithmetic. The ability of kindergarten children to respond efficiently to unpracticed single-digit and unfamiliar three-term zero problems is consistent with the hypothesis that these subjects learned a general zero rule. Though training resulted in some improvement on nonpracticed one and large problems, there was little evidence of qualitative changes in estimating non-zero problem sums. However, the training was of limited duration and some kindergartners already seemed to be using relatively sophisticated estimation strategies. Needed is a long-term study that follows children from the time they produce unsophisticated estimates to the time they produce good estimates. Further research is also needed to determine what role general arithmetic knowledge plays in the development of mental arithmetic. For example, whether or not knowledge of commutativity plays a role in mastering the basic number combinations needs to be examined directly.

In conclusion, the development of mental arithmetic—even that involving the basic single-digit combinations—cannot be understood entirely in terms of forming specific numerical associations. No doubt computing experience plays a role in mastering the basic number combinations, but it does not appear to be the only process at work. The evidence of these studies is consistent with the view that children do not simply recall estimates from a repertoire of specific associations. Rather they seem to manufacture estimates by employing strategies that are based on their knowledge of arithmetic and which operate on their representation of the counting string. Just how previous learning and practice interact to promote learning the basic number combinations remains an open and important theoretical and educational question.

References

Baroody, A. J. (1983). The development of children's informal addition. In J. C. Bergeron & N. Herscovics (Eds.), Proceeding of the Fifth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, (Vol. 1, pp. 222-229). Montreal: Université de Montréal, Faculté des Sciences de l'Education.

Baroody, A. J. (1985). Mastery of the basic number combinations: Internalization of relationships or facts? Journal for Research in Mathematics Education, 16, 83-98.

Ilg, F., & Ames, L. B. (1951). Developmental trends in arithmetic. The Journal of Genetic Psychology, 79, 3-28.

Olander, H. T. (1931). Transfer of learning in simple addition and subtraction. II. Elementary School Journal, 31, 427-437.

Siegler, R. S., & Robinson, M. (1982). The development of numerical understandings. In H. W. Reese & L. P. Lipsitt (Eds.), Advances in child development and behavior, Vol. I (pp. 241-312). New York: Academic Press.

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition: How do children know what to do? In C. Sophian (Ed.), Origins of cognitive skills (pp. 229-293). Hillsdale, NJ: Erlbaum.

Thiele, C. (1938). The contribution of generalization to the learning of the addition facts. New York: Bureau of Publications, Teachers College, Columbia University.

Thorndike, E. L. (1922). The psychology of arithmetic. New York: Macmillan.

Wheeler, L. R. (1939). A comparative study of the difficulty of the 100 addition combinations. The Journal of Genetic Psychology, 54, 295-312.

Table 1

Proportion of Pretest Trials in Which Kindergarten Subjects Used Their Most Frequent Global and/or Specific Estimation Strategy for Three Problem Types

		Problem Type[c]		
		N+0/0+N	N+1/1+N	N+M/M+N
S#[a]	Hypothesized Strategy[b]	(80 trials)	(80 trials)	(120 trials)
03	Add several to the larger addend			
	(particular number: 10)	NA	NA	.72(.46)
01	State a teen (N or M + teen)	NA	NA	.97(.37)
02	State a teen (N or M + teen)	NA	NA	.86(.37)
14	State a particular number: 9	NA	NA	.31
10	State a teen (N or M + teen)	NA	NA	.93(.86)
04	Add one to the larger addend	.65	.65	-
	State a teen (N or M + teen)	-	-	.68(.33)
15	Add one to an addend (larger addend)	.57(.57)	.61(.51)	.67(.55)
21	State an addend (larger addend)	.94(.87)	1.00(.87)	1.00(.84)
12	State a particular number: 10	.34	.30	.31
06	State a particular number: 6	.40		
	State particular numbers: 6, 7, and 8		.64[d]	
	State particular number: 7			.25
09	State particular numbers: 6 and 8	NA	.60[d]	.53[d]
20	Generate a number sequence	.37	.35	.29
05	State an addend (larger addend)	.84(.57)	.84(.42)	
	State an addend (Smaller addend)			.85(.44)
18	Add one to an addend (last addend)	.97(.90)	.97(.89)	.97(.87)
07	State a particular number: 19	.20		
	State particular numbers:			
	12 and 11-teen		.29[d]	
	State particular numbers:			
	19 and 11-teen			.28[d]
17	State a teen (N or M + teen)	.82(.59)	.72(.49)	.82(.55)
13	State a teen	.99	.97	.96

[a]Subjects are listed from the most accurate estimator to the least accurate. Accuracy was computed by summing the mean absolute error scores (correct sum minus estimate) for the 14 problems.

[b]In cases where a global strategy and a more specific version of the general strategy are reported, the latter is included in parenthesis.

[c]NA indicates not applicable. That is, the child was consistently correct on that type of problem and an incorrect procedure could be discounted (e.g. always picking the larger term for one and larger problems as well as zero problems).

[d]The frequency of each particular number used was equal or very nearly equal.

Table 2

Pretest and Posttest Responses in Terms of Distributions of Associations

Problem	Pretest: Mean associative strength of the most frequent response	Posttest: Mean associative strength of response	Responded with an answer with the greatest associative strength	Responded with an answer with the second greatest associative strength	Responded with an answer with the third greatest associative strength	Responded with an answer with low but some associative strength	Responded with a novel answer (with conditional probability = zero)	Total	Pretest-posttest (Glass rank-biserial) correlation
					0+N/N+0				
0+6	.54	.21	3	2	0	0	5	10	.37
0+9	.59	.17	2	1	1	2	4	10	.31
7+0	.57	.15	1	1	4	0	4	10	1.00
8+0	.58	.18	2	0	4	1	3	10	.25
Total			8	4	9	3	16	40	
					1+N/N+1				
1+7	.59	.29	4	2	2	0	3	11	.07
1+8	.48	.34	4	5	0	0	2	11	.50
6+1	.47	.27	4	2	1	2	2	11	.39
9+1	.55	.36	4	1	1	1	4	11	.36
Total			16	10	4	3	11	44	
					M+N/N+M				
3+7	.47	.21	6	1	2	3	4	16	.03
4+8	.48	.16	5	2	3	2	4	16	-.02
5+6	.49	.23	3	5	2	2	4	16	.54
7+4	.51	.19	4	0	2	5	5	16	.44
8+5	.44	.19	4	4	4	2	3	16	.05
9+3	.51	.28	3	2	2	2	3	16	.08
Total			27	14	15	16	24	96	

BEST COPY AVAILABLE

Table 3

A Comparison of Pretest, Posttest, and Transfer Task Performance for Subjects Who Did and Did Not Learn a Zero Rule

		Posttest			Three-Digit Addition Transfer		
S#	Pretest accuracy rate	Accuracy rate	Mean RT rating	Used zero rule discriminantly	Accuracy rate	Mean RT rating	Used zero rule discriminantly
		Children Who Appeared to Learn the Zero Rule					
04	.05	1.00	0.25	Yes	1.00	0	Yes
06	.17	1.00	0.25	Yes	1.00	0	Yes
07	.02	1.00	0.25	Yes	1.00	0	Yes
12	.01	.75	0.75	Yes	1.00	1.25	Yes
15	.01	1.00	0	Yes	1.00	0	Yes
17	.01	1.00	0.50	Yes	1.00	0.50	Yes
20	.12	1.00	0	Yes	1.00	0.75	Yes
		Children Who Did Not Appear to Learn Zero Rule					
05	.57	.75	0.25	No	.50	0.25	No
18	.00	1.00	1.00	No	1.00	0	No
21	.87	1.00	0	No	1.00	0.50	No

Table 4

Indicators of Estimation Quality and Screening Results

Type of Pretest Estimates	S#	Proportion of impossible sums: Pretest	Proportion of impossible sums: Posttest	Mean absolute error: Pretest	Mean absolute error: Posttest	Prearithmetic Skills: Counts to 15	Prearithmetic Skills: Counts sets 1 to 15	Prearithmetic Skills: Counts out sets of 1 to 5	Prearithmetic Skills: N+1>N rule 1 to 10	Prearithmetic Skills: Reads numerals 0 to 9	Arithmetic Skills: Word Problems correct (0 to 5)	Arithmetic Skills: Automatic basic facts (0 to 10)
Flexible and Reasonable	02	.00	.00	1.6	0.9	3	3	3	3	3	5	3
	01	.01	.00	1.5	0.6	3	3	3	3	3	4	6
	10	.01	.00	2.0	0.2	3	0	3	3	3	2	4
	03	.02	.00	0.7	0.6	3	3	3	3	3	5	5
	04	.11	.21	2.1	1.5	3	3	3	3	3	4	4
	14	.23	.21	1.9	1.5	3	3	3	3	2	5	4
	09	.42	.43	2.8	1.6	3	2	3	3	3	5	3
Reasonable	13	.02	-	6.2	-	2	2	3	1	2	3	8
	17	.16	.07	6.1	2.0	3	3	3	2	3	1	0
	12	.39	.21	2.7	1.8	1	2	3	2	0	4[a]	4
	15	.42	.14	2.5	1.0	3	3	3	3	3	4[a]	5
Mechanical	18	.54	.71	4.3	3.4	3	3	3	2	3	2	3
	06	.57	.21	2.8	1.6	2	2	3	2	3	5	3
	20	.57	.43	3.1	2.1	2	2	3	0	1	3	4
	21	.75	.71	2.7	2.0	1	2	3	1	0	1	0
	05	.84	.79	4.3	4.3	3	2	3	0	3	2	2
Other	07	.24	.50	4.7	2.9	1	1	1	2	2	1[a]	0

[a]On an additional problem, the child spontaenously used a correct procedure but miscalculated the sum.

Table 5

Proportion of Estimation Trials in Which Mentally Handicapped Subjects Used Their Most Favored General and Specific Estimation Strategy for Three Problem Types

		Problem Type[c]		
		N+0/0+N	N+1/1+N	N+M/M+N
S#[a]	Hypothesized Strategy[b]	(80 trials)	(80 trials)	(160 trials)
01	Add several to larger addend	NA	NA	.71
28	Add one to either (the larger) addend	NA	NA	.54(.37)
12	Add one to either (the last) addend	NA	NA	.61(.51)
20	Add one to either (the last) addend	NA	NA	.46(.33)
05	Add one to either (the last) addend	NA	NA	.98(.47)
10	Add one to either (the last) addend	NA	NA	.72(.49)
24	Add several to larger addend	NA	.41	-
	Add one to either (the larger) addend	NA	-	.57(.33)
13	Add one to either addend	NA	.60	.63
21	State a teen (addend + teen)	NA	.51(.45)	-
	Add one to either (the last) addend	NA	-	.32(.21)
19	State (first) addend	.99(.96)	NA	-
	State a particular number: 16	-	NA	.44

23	State smaller addend	.97	-	-
	State a particular number: 10	-	.29	.28
15	Add one to larger addend	.99	-	-
	Add one to last addend	-	1.00	1.00
04	Add one to larger addend	1.00	-	-
	Add one to last addend	-	.99	.99
14	Larger addend plus or minus one (add one to larger)	.80(.69)	-	-
	Add one to either (the larger) addend	-	.94(.72)	-
	Add one to either (the last) addend	-	-	.93(.91)
25	State a particular number: 8	.29	-	-
	Add one to either (the first) addend	-	.57	.49
18	Larger addend plus or minus one (add one to larger)	1.00(.85)	-	-
	Add one to the last addend	-	1.00	.99
02	State a particular number: 10	.61	-	-
	Add one to either (the last) addend	-	.86(.81)	-
	Larger addend plus or minus one (add one to last)	-	-	.89(.55)
08	Generate number pairs (sequences)	.39(.07)	.37(.15)	.40(.17)
06	Add one to either (the last) addend	.79(.75)	.79(.69)	.83(.74)
03	Generate number pa.rs (sequences)	.39(.14)	.40(.16)	.42(.14)
11	State (larger) addend	(1.00)	(.96)	.97(.90)
30	State (larger) addend	(.95)	(.94)	.95(.81)

16	State (larger) addend	1.00(.95)	1.00(.90)	.99(.74)
22	State (larger) addend	.95(.91)	.97(.90)	-
	State (first) addend	-	-	.97(.81)
27	State (larger) addend	.96(.81)	.97(.90)	.96(.51)
26	State (larger) addend	1.00(.84)	1.00(.87)	1.00(.56)
29	State (larger) addend	.95(.79)	-	-
	State (last) addend	-	.99(.95)	1.00(.55)
17	State (smaller) addend	.97(.77)	-	-
	State (larger) addend	-	.95(.92)	.96(.79)
09	State (larger) addend	.99(.73)	-	-
	State (last) addend	-	.99(.69)	-
	State (smaller) addend	-	-	.99(.65)
07	State (smaller) addend	1.00(.64)	-	-
	State (last) addend	-	1.00(.57)	1.00(.57)

[a]Arranged in order of estimation strategy quality, as judged by appropriateness of sums and discriminant application to problem types.

[b]In cases where a general strategy and a more specific version of the general strategy are reported, the latter is included in parenthesis.

[c]NA indicates not applicable. That is, the most frequent strategy was to produce the correct sum.

BEST COPY AVAILABLE

Table 6

Mentally Handicapped Subjects' Arithmetic Screening and Estimation Task Results

	Arithmetic Screening			Estimation Task		
	Word problems correct	Concrete Addition Task				
S#	(0 to 16)	Spontaneously used a correct procedure	Number of correct answers	Proportion of impossible sums	Mean absolute error	Overall rank: Impossible sums & absolute error
01	4	12	12	.04	1.1	1
28	15	12	12	.19	1.9	2
12	3	12	11	.25	2.2	3
20	11	12	12	.27	2.3	4
05	16	12	12	.31	2.4	5
10	6	8	7	.35	2.4	6
24	5	12	11	.32	2.6	7
13	7	12	12	.44	2.8	15
21	8	11	11	.15	8.5	19.5
19	12	12	12	.14	3.4	8
23	1	6	6	.37	3.9	12.5
15	6	11	11	.37	3.0	12.5
04	6	11	9	.38	3.0	17
14	8	10	10	.33	2.9	9
25	3	10	7	.35	2.8	10
18	6	8	8	.42	3.0	18
02	6	7	7	.44	3.7	23
08	5	10	9	.05	8.0	16
06	6	11	9	.38	3.4	22
03	4	11	11	.72	4.8	26
11	1	10	9	.73	2.3	11
30	8	7	7	.74	2.4	14
16	2	11	11	.76	2.7	19.5
22	6	2	2	.75	2.8	21
27	5	11	10	.77	3.1	24
26	1	11	11	.79	3.1	25
29	0	4	1	.80	3.8	27
17	0	0	0	.92	3.5	28
09	0	10	10	.82	4.2	29
07	4	10	6	.91	4.4	30

Figure 1: Siegler and Shrager's (1984) Distributions-of-Associations Data

PROBLEM	ANSWER												
	0	1	2	3	4	5	6	7	8	9	10	11	OTHER
1 + 1		.05	.86		.02		.02					.02	.04
1 + 2			.09	.70	.02		.04			.07	.02	.02	.05
1 + 3		.02		.11	.71	.05	.02	.02					.07
1 + 4					.11	.61	.09	.07				.02	.11
1 + 5					.13	.16	.50	.11		.02	.02		.05
2 + 1		.07	.05	.79	.05								.04
2 + 2	.02		.04	.05	.80	.04		.05					
2 + 3			.04	.07	.38	.34	.09	.02	.02	.02			.04
2 + 4		.02		.07	.02	.43	.29	.07	.07				.04
2 + 5		.02		.05	.02	.16	.43	.13			.02		.18
3 + 1		.02		.09	.79	.04		.04					.04
3 + 2			.09	.11	.11	.55	.07						.07
3 + 3	.04			.05	.21	.09	.48		.02	.02	.02		.07
3 + 4				.05	.11	.23	.14	.29	.02				.16
3 + 5				.07		.13	.23	.14	.18		.05		.20
4 + 1			.04	.02	.09	.68	.02	.02	.07				.07
4 + 2			.07	.09		.20	.36	.13	.07		.02		.07
4 + 3				.05	.18	.09	.09	.38	.09		.02		.11
4 + 4	.04			.02	.02	.29	.07	.07	.34		.04		.13
4 + 5					.04	.09	.16	.09	.11	.18	.11	.04	.20
5 + 1			.04		.04	.07	.71	.04	.04		.04		.04
5 + 2			.05	.20	.02	.18	.27	.25	.02		.02		
5 + 3			.02	.11	.09	.18	.05	.16	.23		.05		.11
5 + 4					.11	.21	.16	.05	.11	.16	.04		.16
5 + 5	.04					.07	.25	.11	.02	.04	.34	.04	.11

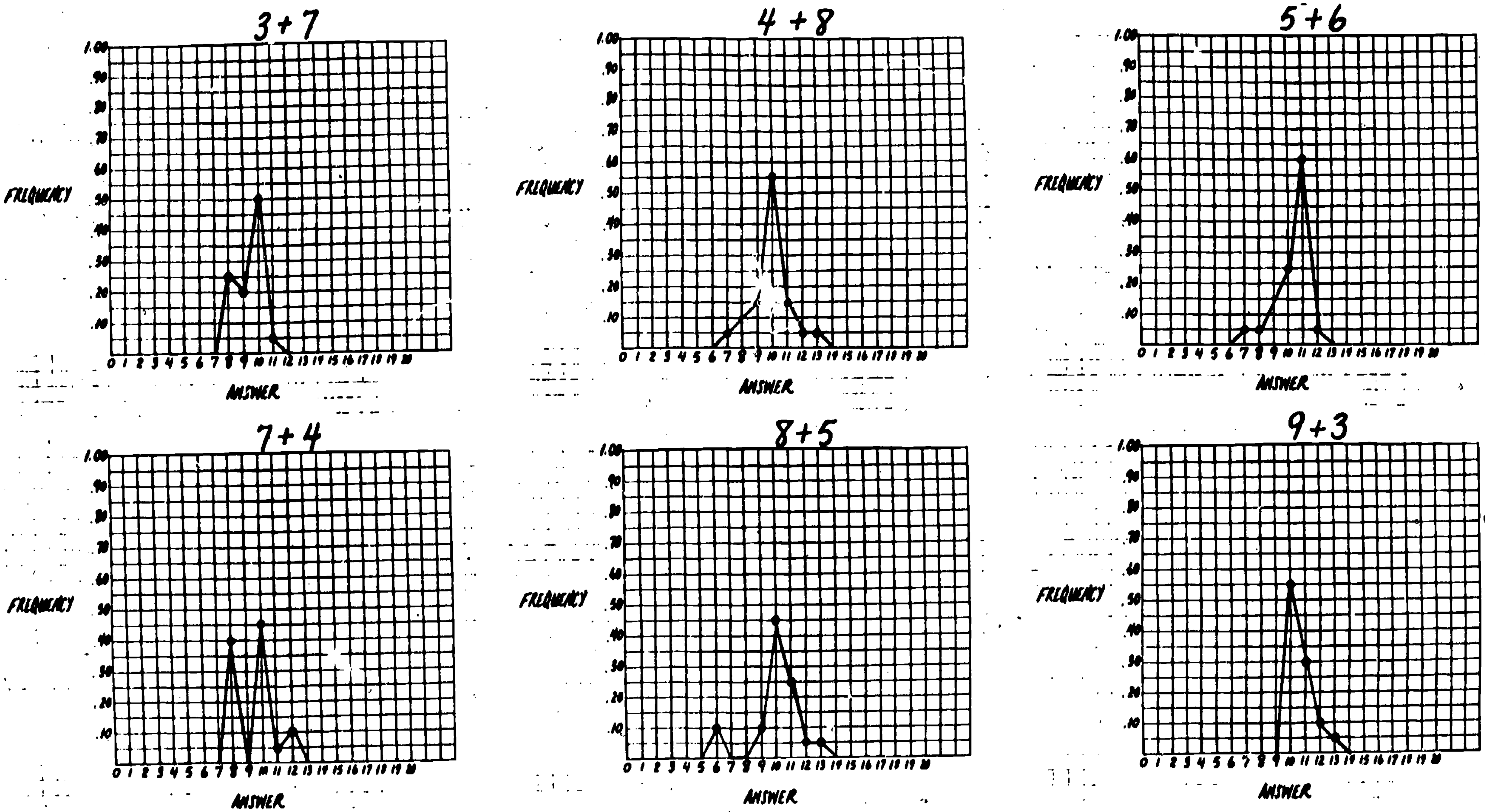

Figure 2: S.03's Pretest Responses to Large Problems

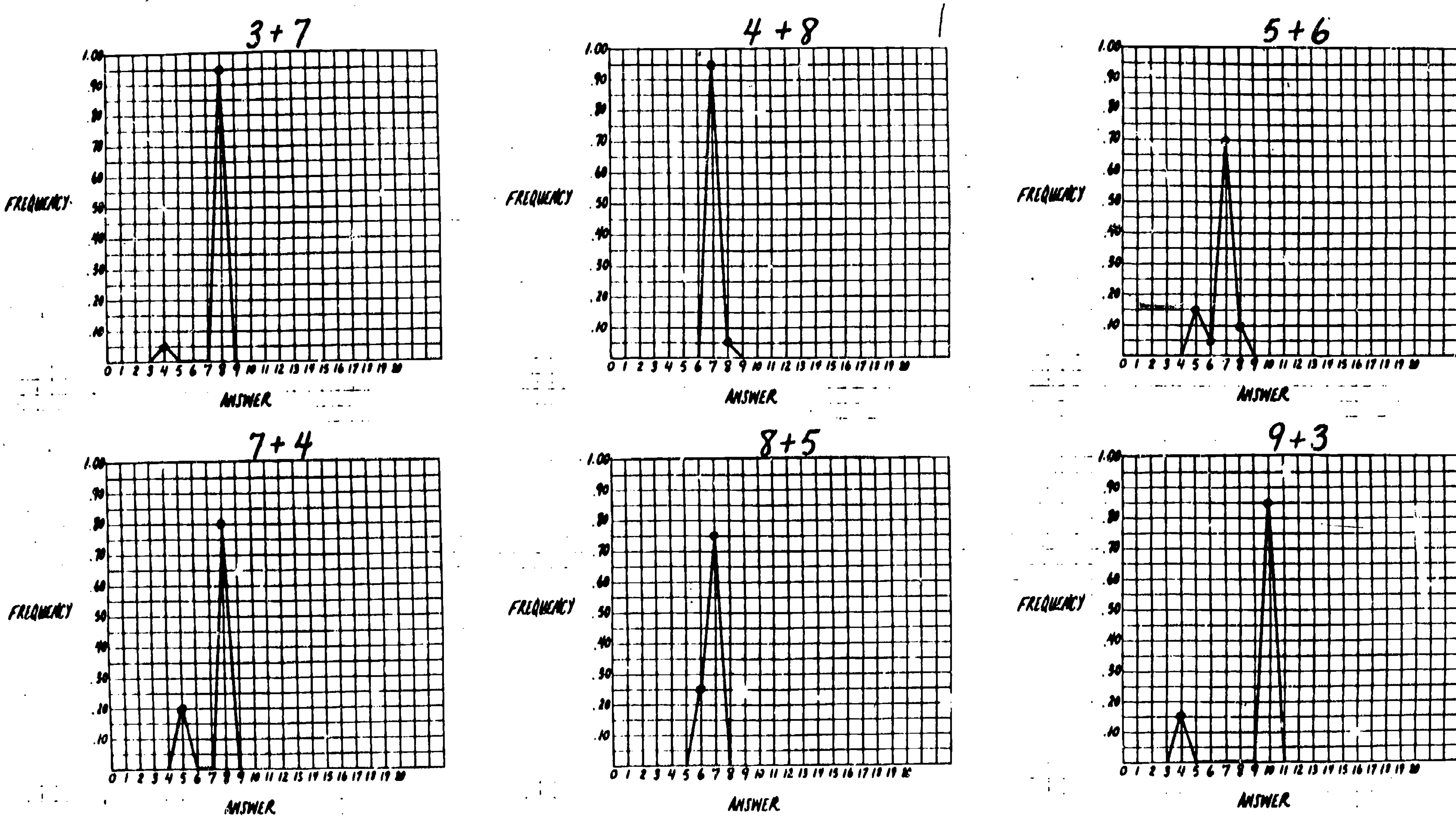

Figure 3: S-15's Pretest Responses to Large Problems

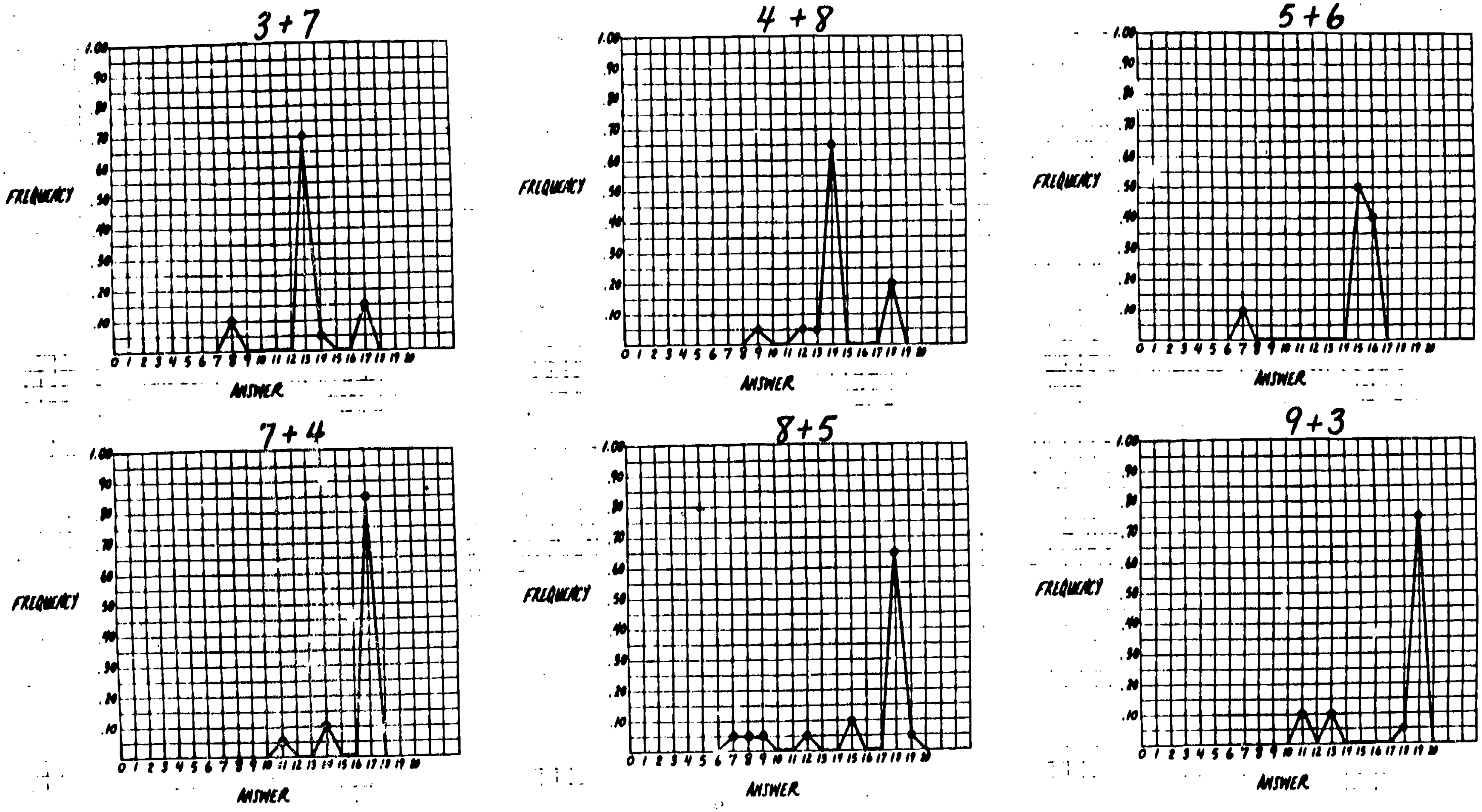

Figure 4: S 10's Pretest Responses to Large Problems