DOCUMENT RESUME

ED 041 623 PS 003 167

AUTHOR Steffe, Leslie P.; Johnson, David C.
TITLE Problem Solving Performances of First Grade Children.
INSTITUTION Georgia Univ., Athens. Research and Development Center in Educational Stimulation.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Cooperative Research Program.
PUB DATE Mar 70
CONTRACT OEC-6-10-061
NOTE 25p.; Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, Minnesota, March, 1970

EDRS PRICE EDRS Price MF-$0.25 HC-$1.35
DESCRIPTORS Ability Grouping, *Arithmetic, *Conservation (Concept), Factor Analysis, Grade 1, Intelligence Quotient, Manipulative Materials, Measurement Instruments, *Performance Factors, *Problem Solving, Task Performance

ABSTRACT

This study examined differential performances among groups (categories) of first grade children when solving eight different types of arithmetical word problems under two distinct experimental conditions. The categories of children were actually 4 ability groups: (1) low quantitative comparison scores and low IQ (Lorge-Thorndike IQ Test), (2) low quantitative comparison scores and high IQ, (3) high quantitative comparison scores and low IQ, and (4) high quantitative comparison scores and high IQ. The 111 children who filled these categories were given a 48-item problem solving test, with six problems from each of the eight types presented in a randomized sequence. Half of the children in each ability group were randomly assigned to the condition of no manipulatable objects, while the other half were provided with manipulatable objects referred to in the problems and were allowed to use them any way they wanted to help solve the problems. Analysis of the data revealed that IQ was not a significant factor, that Problem Condition was significant, that there was a significant interaction due to Quantitative Comparisons and Problem Conditions for one problem type, and that there were significant main effects due to Problem Conditions for the remaining seven problem types. There was also a significant main effect due to Quantitative Comparisons for one of the remaining seven problem types. (MH)

PROBLEM SOLVING PERFORMANCES OF FIRST GRADE CHILDREN

Leslie P. Steffe David C. Johnson

University of Georgia

A paper read at the Annual 1970 AERA Meeting, Minneapolis, Minnesota.

This report was made as part of the activities of the Research and Development Center in Educational Stimulation, University of Georgia, pursuant to a contract with the United States Department of Health, Education, and Welfare, Office of Education, under Provisions of the Cooperative Research Program.

Center No. 5-0250 Contract No. OE 6-10-061

PROBLEM SOLVING PERFORMANCES OF FIRST GRADE CHILDREN

Leslie P. Steffe David C. Johnson

University of Georgia

The purpose of this study was to investigate differential performances among categories of first grade children when solving arithmetical word problems of eight different types under two distinct conditions. The categories were, in part, based on various types of quantitative comparisons which children are known to make; gross quantitative comparisons, intensive quantitative comparisons, and extensive quantitative comparisons. Elkind (1961) defines gross quantitative comparisons as "single perceived relations between objects (larger than, fewer than) which are not coordinated with each other [p. 37]." For example, a child may judge two equivalent sets of objects to be of the same number when the objects of each set are placed adjacent to one another in a one-to-one fashion; but upon moving one of the two sets of objects so that it subtends a region of greater (or less) measure than the original, the child may no longer believe the two sets possess the same number. He has not yet started to coordinate the density of the objects with the measure of the region the objects subtend. When this coordination begins, the child is capable of intensive quantitative comparisons (Piaget, 1952). He may not yet be able to grasp proportionality of differences, but is capable of logical coherence. Elkind (1961) has shown that gross quantities are easiest for children to compare, then intensive and extensive quantities in that order.

Conceptually, the nature of the transformation of the objects involved when children are asked to make a quantitative comparison is of significance. A physical movement of the objects of one or both collections may take place, hereafter termed "physical transformation." This class of transformations leaves the two collections in any one of an uncountable number of final states. Any final state has the possibility of being reversed to the initial state by a simple returning of the objects to their original position, physically or in thought, which is termed "reverse transformation." It is important to recognize, however, that a child may be presented with two patterns of objects which are equivalent but which appear nonequivalent by virtue of their configuration. In this situation, no recourse to reversibility is possible, but an analogous process does exist, termed "forward transformation" by Beilin (1969) "in which the pattern can be ideationally rearranged in anticipation of an actual or hypothetical rearrangement [p. 435]." Forward transformation is viewed as a more significant type of transformation than reverse transformation, since, as Beilin (1969) comments, "it is the basis of many kinds of problem solving [p. 435]." He also postulates the existence of an analytic set which initiates the forward transformation which activates solution strategies available to the child. Piaget's Grouping Structures may be considered to serve as bases for solution strategies to which Beilin refers. Detailed accounts of these structures and accompanying behavioral manifestations have been given elsewhere (Flavell, 1963) and will not be repeated. It is only prudent to point out that some psychologists do not admit that grouping structures form models for intellectual operations. Note Kohnstamm's (1967) comment, "We do not believe in the existence of these structures. . .[p. 145]." Theoretical relationships, however, do exist between one-to-one correspondence and quantitative comparisons and has been explicated elsewhere (Harper & Steffe, 1968). Just what sort of cognitive activity a child engages in when presented with a stimulus such as depicted by Figure 1 is open to debate, where the question "Are there more squares here (the experimenter points to one of the collections) or are there more squares here (the experimenter points to the remaining collection) or are there the same number of squares here as here?"

Insert Figure 1 about here

The fact that the child must engage in some sort of cognitive activity which involves a forward transformation to arrive at a correct answer is not open to debate (barring solution by non-cognitive means). The forward transformation could be inextricably involved with logical multiplication of relations and extensive quantitative comparisons.

Sullivan (1967) has written a critical analysis of Piaget's theory as it relates to School Curriculum, in which he states, "A substantial correlation between number readiness (e.g. conservation of number) and the achievement of addition and subtraction can be interpreted in both directions (p. 21)." Assuming that what Sullivan means by "addition" and "subtraction" is processing sums and differences and that what he means by "conservation of number" is the ability of children to make extensive quantitative comparisons involving forward or reverse transformations, then "conservation of number" is a logical prerequisite to "addition" and "subtraction." It has been shown (LeBlanc, 1968; Steffe, 1966) that differential mean performances do exist among different categories of first grade children when solving addition and subtraction problems, when the categories are based on levels of the ability to make extensive quantitative comparisons involving forward transformation. The point of view may be adopted that children who are able to make extensive quantitative comparisons involving forward and reverse transformations are in the concrete operational stage as explicated by Piaget (1952), and thereby it is hypothesized that such children are able to solve arithmetical word problems with structural types a + b = n, a - b = n, a + n = b, and n + b = a in the presence or absence of manipulatable objects during solution with equivalent success for each type. In the studies referred to above, (LeBlanc, 1968; Steffe, 1966) this hypothesis was not substantiated in the case of the two problem structures a + b = n and a - b = n. The hypothesis has never been tested, however, for problem structures a + n = b or n + b = a. The two problem structures a + b = n and a - b = n did contribute differentially to problem difficulty.

For the category of children who do not display an ability to make quantitative comparisons involving forward or reverse transformations, it is hypothesized that Problem Conditions (presence or absence of manipulatable objects during solution) is significant in favor of manipulatable objects, when children categorized as above solve arithmetical word problems of the structural types a + b = n, a - b = n, a + n = b, and n + b = a. This hypothesis has been substantiated (LeBlanc, 1968; Steffe, 1966) in the case of the two structural types a + b = n and a - b = n.

A problem statement may or may not involve a described action. For those children who do display an ability to make extensive quantitative comparisons, whether the problem statement involves a described action or not may be of little consequence, since these children are able to initiate a forward transformation. For those children who do not display such an ability, the problems which are presented under the condition of a described action may be easier for the children to solve than those problems presented under the condition of no described action. It is not known whether Action (described action or no described action) operates equivalently across the four problem structures of interest within or across different abilities to make extensive quantitative comparisons. Moreover, it is not known if Action operates equivalently within the two levels of Problem Conditions across

Problem Structures. It has been found (LeBlanc, 1968; Steffe, 1966) that problems of the structual types a + b = n and a - b = n presented under the condition of a described action were easier for children to solve than when they were presented under the condition of no described action. No interaction occurred with Problem Conditions or levels of an ability to make quantitative comparisons.

The classification variable, IQ, has been found to be significant (LeBlanc, 1968; Steffe, 1966) relative to problem solving performances of First Grade Children and was thereby included. The Lorge-Thorndike Intelligence Test; Level I, Form A was utilized, since, in the words of Lorge and Thorndike (1957), "The items . . . were selected so that, . . ., they deal with relationships.. In answering most of the items a pupil is required to find a principle and then apply it. The tests, . . ., have been designed to measure reasoning ability [p. 10]."

Method

Test of Quantitative Comparisons

A small group test which contains eight items involving a forward transformation and seven items involving a physical or reverse transformation has been utilized in other studies (Harper et al., 1968) and was deemed appropriate for utilization in this study. A performance criterion for this test of at least ten correct comparisons out of 15 comparisons was established, based on random responses. If a child responded on a random basis, the probability of him achieving ten correct comparisons out of 15 is less than .01. Moreover, a child had to respond correctly to no less than five out of the eight items which involved a forward transformation and to no less than four out of the seven items which involved a reverse transformation to meet criterion. Children who met this criterion were able to (1) make extensive quantitative comparisons involving forward transformation, (2) make extensive quantitative comparisons involving reverse transformations, and (3) conserve a one-to-one correspondence. Children who scored no more than seven correct were classified as not meeting criterion. If a child guessed, the probability of his score being in the set {3, 4, 5, 6, 7} is approximately .8 and the probability of his score being in the set {0, 1, 2} is approximately .13. Those children who scored an eight or nine were excluded from the data analysis.

Procedure

Four schools from among the elementary schools in Walton County, Georgia participated. With very few exceptions (in the case of absence), all the first grade children in these four schools were given the test of quantitative comparison during the period from November 6 to November 22, 1967. A total of eight classrooms were involved in the 199 children tested. From January 15 to January 24, 1968, the Lorge-Thorndike IQ Test; Level 1, Form A, was administered to 192 of the previously tested 199 children. The families of seven children had moved since the administration of the first test. Of the 192 children upon which two measures had been taken, only those 127 children whose IQ scores were in the range of 80-97 or 103-120 were used in the study. Those children were, by virtue of their scores, separated into four ability groups; (a) quantitative comparison scores (0-7) and IQ scores (80-97), (b) quantitative comparison scores (0-7) and IQ scores (103-120), (c) quantitative comparison scores (10-15) and IQ scores (80-97), and (d) quantitative comparison scores (10-15) and IQ scores (103-120). Sixteen of the 127 children had total scores of eight or nine. Of those not eligible for the study, excluding the 16 with total scores of eight or nine, 41 had IQ's greater than 120. From May 1 to May 28, 1968, a problem solving test was administered to 108 of these 111 children. Of the three children not tested, the family of one had moved and two were absent during the duration of the testing.

The test of quantitative comparisons was normally administered to the children in groups of five. All of the testing was done by one tester during the school day in various sized rooms but always in moderately quiet surroundings. Children were seated in a row facing the tester and were separated by 2 x 3 foot pieces of cardboard. Average testing time was 15-18 minutes per group. The Lorge-Thorndike IQ Test was also administered by one tester. Approximately fifteen children were tested at one time while the classroom teacher entertained the remainder of the class elsewhere. The IQ Test is divided into three sub-tests so that time was provided between sub-tests for the children to relax. Total time spent with each group was about 35 minutes. The final test given to each child was a 48 item problem solving test. A practice problem was used to help the student feel relaxed and also to acquaint him with certain mechanics. For the practice problem, the child was helped if he seemed slow in responding. During the administration of the other 48 problems no help was provided by the tester. Twelve problems of each of the four following problem structural types were presented to each child: $a + b = n$, $a - b = n$, $a + n = b$, and $n + a = b$. Six of the twelve problems in each problem structural type involved a described action and six did not involve a described action. The items were randomized independently for each child. The children in each ability group were randomly assigned to Problem Conditions such that within each group, the number of children assigned to each condition differed by no more than one. Each child was tested individually in two sittings by two different testers, who were randomly assigned to the first or second position for each child. Each sitting lasted from 15-30 minutes depending on the child. For each child assigned to the condition of no manipulatable objects, each problem statement was repeated once and occasionally twice if the child failed to respond. When manipulatable objects were present, time was provided for the children to manipulate the objects after each reference to a set of materials in the given problem was made. For each problem, one more object than was needed for solution of the problem was present in each material set. For example, if the problem involved "fruit", a piece of cardboard containing "oranges" and a piece containing "apples" were moved near the child, who then could make the selection of as many of these as he wanted after he heard them referred to in the problem. When aids were available, the child was encouraged but not told how to use them. Any manipulations had to be initiated by the child.

Materials

There were eight different problem types in the 48 item problem solving test, with six problems within each of the eight types. The six problems within each problem type differed on the names of the categories referred to in the problem and on the number triples assigned. Each of six different material sets and each of six different number-triples were used once for each problem type. The number-triples were; (3,5,8), (3,4,7), (3,7,10), (2,7,9), (4,2,6), and (6,3,9). The material sets used were models of the objects in the stated categories.

Data Analysis

Test statistics, including internal consistency reliability coefficients (KR-20), and a principal component analysis were computed for each six-item problem solving test and for the test of quantitative comparison, utilizing an appropriate computer program (Wolf, Klopfer, 1963). Program MUDAID (Multivariate, Univariate, and Discriminant Analysis of Irregular Data) was used for the MANOVA, where the eight problem types were the response variables. MUDAID provides a multivariate analysis on all response variables for all combinations of independent variables taken two at a time. Thus, three MANOVA's were printed out; one for each of IQ by Quantitative Comparisons (Q), IQ by Problem Conditions (C), and Q by C. For each univariate ANOVA, cell means, standard deviations, and number of respondents in each cell are reported. For each MANOVA, a discriminant function for each effect, each interaction, and for all effects are included, as well as correlations based on the error matrix.

Results of the Study

Reliability Studies

As noted previously, 199 children were administered the test of quantitative comparisons. Table 1 contains the test statistics and Table 2 the frequency distribution of total scores. The internal consistency reliability of the test was substantial, indicating good homogeneity of the test items. The mean score of 8.68 and rather large standard deviation reflect

---

Insert Table 1 about here

Insert Table 2 about here

---

the fact that a substantial number of children scored at almost every point of the criterion scale, which makes it rather difficult to categorize children as "conservers", "nonconservers", and "transitionals" on the basis of a small number of items. The item difficulties and principle component analysis are given in Table 3. Item 3, a difficult item, did not load on either factor with sums of squares greater than one. Items 7, 9, and 12 were the three items, other than Item 3, which contained unequal numbers of objects in the collections. Since the basis for the relative difficulty of Item 3 was not

---

Insert Table 3 about here

---

clear, the item was retained. The relative difficulty of Item 12 may be explained by the nature of the geometrical configuration of the objects. The children could have judged the two collections of objects as containing the same number on the basis of the fact that the length of the two line segments subtended by the objects were of the same measure. That the items which contained an unequal number of objects in the two sets to be compared and the items which contained an equal number of objects in the sets to be compared demand differential abilities, gains support from other studies (Carey, Steffe 1968; Rothenberg, 1969) and is discussed later. It is significant to note that of the children included in the MANOVA, each child who responded correctly to at least 10 items also met the additional criterion of responding correctly to no less than five out of the eight items which involved a forward transformation and no less than four items out of the seven items which involved a reverse transformation.

Item statistics and analyses are given in Table 4 for each of the eight subtests of problem solving. The number of children for the analyses was 125, since most of those children who had total scores of 8 or 9 on the test of quantitative comparisons were also given the problem solving test.

The mean scores for the two problem types a + b: A and a + b: N* were appreciably greater (no significance test performed) than the mean scores of the remaining problem types, which were all quite comparable. The mean score

---

Insert Table 4 about here

Insert Table 5 about here

---

* a + b: A   A denotes described action; N denotes no described action.
a + b: N

within the two levels of the Problem Conditions is explicated when the results of the MANOVA are presented as well as the means within the two categorization variables. It is important to note that "skewness" was significant for the problem structures a + b: A and a + b: N, while "kurtosis" was not. Just the opposite was the case for the other six problem structures. The reversal is a reflection of the frequency distribution of the test scores given in Table 5. The mean scores and standard deviations given in Table 4 and the frequency distribution given in Table 5 suggested that described action, as used in this study, did not differentially affect problem solving performances on the four basic structures. To obtain further information on this point, four principle component analyses were carried out on the four basic problem structures. The results indicated that there was no reason to believe that the items involving a described action and the items involving no described action involved differential abilities, since there was only one factor with loadings consistently greater than .5 and whose sums of squares of the loadings exceeded 1.

MANOVA

Since IQ was not significant in any of the MANOVA's and Univariate ANOVA's and did not interact with any other variable, analyses involving IQ are not reported. The following is a discussion of Analyses relative to Q and C.

The subclass means of the eight tests are presented in Table 6. The likelihood ratio test statistic, $\chi^2_{24} = 50.62$, for all effects of Q and C was significant ($p<.001$). The main effect due to Q and the interaction of Q and C were not significant. The main effect due to C was significant, $F(8,96) = 4.19$ ($p<.01$). Table 7 contains the results of the eight univariate analyses run. A notable result was that for the problem type a + b: A, the main effect of Q was significant as well as the main of C.

On this problem type, those children classified in the high category of Q performed better when solving problems of the indicated type than those children in the low category of Q, where the low mean score in the Q x C matrix was 53 percent, obtained by those children in the low category of Q and no manipulatable objects present. The high mean score in the same matrix was 86 percent obtained by those children in the high category of Q and manipulatable objects present.

Insert Table 6 about here

Insert Table 7 about here

A second notable result is the significant interaction of Q and C for the problem type a + b: N. The Q x C matrix for this problem type in terms of mean percentages is given in Table 8, which shows that the effect of C was negligible within the high category of Q, i.e. the presence of manipulatable objects was not a facilitator of mean problem solving performances for children within the high category of Q and within the problem type a + b: N. It must be remarked that these results were not a great deal different than the results obtained for the problem type a + b: A. The mean scores for that problem type may be obtained from Table 6. The differences that were present are not explainable in terms of past research.

In Table 9, a correlation matrix among the eight dependent variables is presented. These correlations are interpreted as "intrinsic" correlations (after elimination of all significant main effects) among the problem types. The two problem types a + b: A and a + b: N intercorrelated .74. The correlation of either of these problem types with any other problem type fell in the interval [.45, .59]. The remaining six problem types had intercorrelations falling in the interval [.64, .79] with only four of the fifteen less than .70.

Insert Table 8 about here

Insert Table 9 about here

Discussion

The test of quantitative comparisons had good psychometric properties and an acceptable criterion performance level which was supported by the data in that each child responded correctly to at least 10 items on the test also met the conditions of responding correctly to no less than five out of eight items which involved a forward transformation and no less than four items out of seven items which involved a reverse transformation. The principal component analysis supported a contention that the items which contained equal numbers of objects in the sets to be compared and the items which contained unequal numbers of objects in the sets to be compared would demand differential abilities. It is important to note that those items which contained an unequal number of objects varied across transformational type. Other fluctuation of item difficulty was not a function of the transformational type as Beilin (1969) found, but more a function of the final geometrical configuration of the objects. The fact that the items involving unequal numbers of objects loaded on a factor different than the factor on which the items with equal numbers of objects loaded, demands further explication. In a previous study involving length relations (Carey et al., 1968), a test of conservation of length relations was constructed which demanded only short "yes" or "no" answers. The "yes" answers were in response to a question relative to the relation involved. For example, if a child established that curve A was as long as curve B, the question asked after a transformation was, "Is curve A still as long as curve B?" The "no" answers were in response to a question in which a change of the relation or terms of the relation was involved. The second question in the example above would be, "Is curve A longer than curve B?" or "Is curve A shorter than curve B?" A principal component analysis revealed a bipolar factor with the items of the first type loading negatively and the items of the second type loading positively, with quite comparable item difficulties. It must be pointed out that it was not the relational type, i.e., equivalence vs. order relations, which determined the factors, but instead the type of inference demanded by the task. The way the questions were worded on the test of quantitative comparisons precluded the possibility that the instrument was sensitive enough to discriminate among inference types, which is now construed to be a shortcoming of the test. Moreover, the instrument was not sensitive enough to discriminate those children who were responding relative to quantitative relations from those children who were responding on the basis of set relations, a discrimination which is now viewed as important, on a theoretical basis.

In a methodological study of "conservation of number", Rothenberg (1969) commented that it "seems unlikely that the Ss can reliably answer the three-section question with a single response (p. 385)." Experimental evidence (Van Engen & Steffe, 1967) however, indicates that the order in which the three-section question is phrased is not important in terms of a biased selection. More important than the order of the questions is the consideration of just what logical property is necessary in order to respond correctly to each part. The results just given relative to the length relations leads one to the position that it is necessary to elicit responses relative to each part of a multiple section question. It is possible that the differential abilities revealed in the principal component analysis of this study were more a result of children's having to use the asymetric property on the questions involving an order relation than they were a result of having to respond in terms of an order relation per se. It is quite important, both from a point of view of understanding phenomena usually termed "conservation" and from the point of view of assessment, that the properties of equivalence and order relations and logical consequences of those relations be considered. It is apparent that work needs to be done in constructing instruments with good construct validity for the purpose of measuring different facets of the ability to make quantitative comparisons. For children who met criterion on the test used in this study, it is not at all clear that those children were able to use properties or consequences of the relations involved or whether they perceived of the relations as quantitative relations or not.

In the case of objects present, no difference existed between the mean performances of the two categories of Q for the type a + b: N. In the case of objects absent, the mean performance in the high category of Q is 75 percent and the mean performance of the low category of Q was 48 percent, a marked difference. For the problem structure a + b: A, the mean performances in the high and low categories of Q in the case of the condition, "no objects", was 73 percent vs. 53 percent, respectively; and for the high and low categories of Q in the case of the condition "objects present", 86 vs. 82 percent which suggests the presence of an interaction, but which was not significant. The mean performances in the high and low categories of variable Q for the case of "no objects present" is considered as substantively different for each problem structure a + b: A and a + b: N, while the analogous mean performances for the case of "objects present" is not considered as substantively different for the same problem structures. The analogous mean performances across levels of Problem Conditions were not statistically different for each of the other six problem types. The total mean scores for these problem structures were between 46 and 54 percent, inclusive.

Hence, the hypothesis that children in the high category of Q are able to solve arithmetical word problems across structural types in the presence or absence of manipulatable objects during solution, with equivalent degrees of success was substantiated only for those structural types on which the children had the greatest mean scores. The hypothesis that Problem Conditions is significant for those children in the low category of Q has been substantiated both by results of the MANOVA and the univariate analyses. It appears that forward transformation may be basic to solution of arithmetical word problems for which relevant solution strategies are available.

The variable, Described Action, operated differently in this study than in the studies previously mentioned (LeBlanc, 1968; Steffe, 1966). However, in those studies it was pointed out that the significance of the variable may have been the result of instruction. If the variable is to be intelligently utilized in instructional settings, further experimentation must be conducted in which it is manipulated to ascertain it's relative contribution to the problem solving abilities of young children. It is desirable that children are able to solve arithmetical word problems involving both a described action and no described action. It is also highly desirable that first grade children be able to solve arithmetical word problems at least of the various structural types as efficiently in the absence of manipulatable objects as in their presence. Experimentation needs to be conducted in which Problem Structure and Problem Conditions are systematically varied and outcomes assessed both as direct achievement tests and transfer tests. Such experimentation seems to be what Kieren (1969) was calling for when he stated, "Within the context of manipulative methodology, studies need to be done to determine the value of actual as opposed to vicarious manipulation. . . [p. 518]."

REFERENCES

Applebaum, M., & Bargmann, R.E. A fortran II program for MUAID: multivariate, univariate, and discriminant analysis of irregular data. Technical Report NONR 1834 (39), Urbana: University of Illinois Press, 1967.

Beilin, H. Stimulus and cognitive transformation in conservation. In D. Elkind, & J.H. Flavell (Eds.), Studies in cognitive development: Essays in honor of Jean Piaget. New York: Oxford University Press, 1969.

Carey, R.L., & Steffe, L.P. An investigation in the learning of equivalence and order relations by four-and five-year old children. Research Paper No. 17, Research and Development Center in Educational Stimulation, Center No. 5-0250, OE 6-10-061. Athens: University of Georgia, 1968.

Elkind, D. The development of quantitative thinking: A systematic replication of Piaget's studies. The Journal of Genetic Psychology, 1961, 98, 36-46.

Flavell, J.H. The developmental psychology of Jean Piaget. Princeton: D. Van Nostrand Company, Inc., 1963.

Harper, E.H., & Steffe, L.P. The effects of selected experiences on the ability of kindergarten and first-grade children to conserve numerousness. Technical Report No. 38, Research and Development Center for Cognitive Learning, C-03, OE 5-10-154. Madison: University of Wisconsin, 1968.

Kieren, T.E. Activity learning. Review of Educational Research, 1969, 39 (4), 509-522.

Kohnstamm, G.A. Piaget's analysis of class inclusion: right or wrong? Paris: Mouton & Co., 1967.

LeBlanc, J. The performance of first grade children in four levels of conservation of numerousness and three IQ groups when solving arithmetic subtraction problems. (Doctoral dissertation, University of Wisconsin) Ann Arbor, Mich.: University Microfilm, 1968. No. 67-68.

Lorge, I., & Thorndike, R.L. Examiner's manual: the Lorge-Thorndike intelligence tests. Boston: Houton Mifflin Company, 1957.

Piaget, J. The child's conceptions of number. London: Routledge and Kegan Paul, 1952.

Rothenberg, B.B. Conservation of number among four- and five-year-old children: Some methodological considerations. Child Development, 1969, 40 (2), 383-406.

Steffe, L.P. The performance of first grade children in four levels of conservation of numerousness and three IQ groups when solving arithmetic addition problems. Technical Report No. 14, Research and Development Center for Learning and Re-Education, C-03, OE 5-10-154. Madison: University of Wisconsin, 1966.

Sullivan, E.V. Piaget and the school curriculum: A critical appraisal. The Ontario Institute for Studies in Education, 1967, No. 2.

Van Engen, H., & Steffe, L. First grade children's concept of addition of natural numbers. Technical Report No. 5, Research and Development Center for Learning and Re-Education, C-03, OE 5-10-154. Madison: University of Wisconsin, 1966.

Wolf, R., & Klopfer, L. Program TSSA' test scorer and statistical analysis 2. Computer Program Library: Social Sciences Division, Chicago: University of Chicago, 1963.

TABLE 1

Test Statistics: Quantitative Comparisons

Mean	SE	St. Dev.	Skewness	Kurtosis	Reliability (KR-20)
8.68	.29	4.09	-.63*	-.73*	.86

*(p<.05)

TABLE 2

Frequency Distribution of Total Scores: Quantitative Comparisons

Total Score	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Frequency	5	11	9	9	8	4	9	8	16	14	22	25	25	18	13	3

TABLE 3

Principal Component Analysis: Quantitative Comparisons

ITEM		FACTORS	
NUMBER	DIFFICULTY	1	2
1	0.492	0.6651	0.1457
2	0.593	0.7668	0.0843
3	0.191	-0.1064	-0.2924
4	0.698	0.6828	0.2656
5	0.663	0.8764	0.0089
6	0.714	0.8400	0.2009
7	0.513	0.4585	-0.6205
8	0.573	0.7817	0.1563
9	0.563	0.3685	-0.6280
10	0.643	0.7068	0.2143
11	0.618	0.8220	-0.1829
12	0.417	0.0345	-0.5558
13	0.729	0.8324	-0.0020
14	0.568	0.7359	-0.2338
15	0.709	0.7763	0.0450
SUM SQ		6.9523	1.4736

TABLE 4

Test Statistics: Problem Solving

Type **	Mean	SE	St. Dev.	Skewness	Kurtosis	Reliability (KR-20)
a+b: A	4.53	.14	1.61	-.97*	-.08	.59
a-b: A	3.14	.17	1.91	-.03	-1.18*	.61
a+x: A	3.25	.19	2.13	-.23	-1.30*	.69
x+a: A	2.90	.18	1.97	.14	-1.21*	.63
a+b: N	4.45	.15	1.69	- 1.16*	.59	.61
a-b: N	2.77	.18	2.02	.17	-1.23*	.65
a+x: N	3.00	.18	2.07	- .01	-1.33*	.66
x+a: N	2.87	.19	2.10	.10	-1.36*	.69

** Six Item Subtests

* $(p<.05)$

TABLE 5

Frequency Distributions of Total Scores: Problem Solving Tests

Type \ Total Score	0	1	2	3	4	5	6
a+b: A	1	9	7	14	16	31	47
a-b: A	11	21	17	18	23	17	18
a+x: A	21	14	10	15	23	17	25
x+a: A	16	20	24	15	17	16	17
a+b: N	6	5	5	14	19	33	43
a-b: N	21	21	18	18	16	15	16
a+x: N	21	16	18	15	17	19	19
x+a: N	21	23	14	13	21	13	20

Table 6

Subclass Means: Q x C

	Type	C_1*	C_2*	Means
Q_1**	a+b: A	3.18	4.94	4.06
	a-b: A	2.19	3.94	3.03
	a+x: A	1.82	4.24	3.03
	x+a: A	2.47	3.53	3.00
	a+b: N	2.88	4.65	3.76
	a-b: N	1.82	3.41	2.62
	a+x: N	1.94	3.71	2.82
	x+a: N	1.65	4.12	2.88
Q_2**	a+b: A	4.39	5.17	4.77
	a-b: A	2.76	3.51	3.12
	a+x: A	2.76	3.89	3.30
	x+a: A	2.47	3.47	2.96
	a+b: N	4.50	4.83	4.66
	a-b: N	2.24	3.26	2.73
	a+x: N	2.40	3.74	3.04
	x+a: N	2.16	3.63	2.86
Means	a+b: A	4.02	5.10	*C_1: No Objects Present
	a-b: A	2.56	3.65	
	a+x: A	2.47	4.00	
	x+a: A	2.47	3.50	*C_2: Objects Present
	a+b: N	4.00	4.77	
	a-b: N	2.11	3.31	**Q_1: Low Category
	a+x: N	2.26	3.73	
	x+a: N	2.00	3.79	**Q_2: High Category

TABLE 7

Univariate ANOVA'S: Q x C

Type Variation	a+b: A	a-b: A	a+x: A	x+a: A	a+b: N	a-b: N	a+x: N	x+a: N
Q	5.66**	<1	<1	<1	7.10**	<1	<1	<1
C	14.53**	9.65**	15.62**	8.05**	6.10*	10.17**	16.12**	23.74**
Q x C	2.59	2.02	2.40	<1	4.43*	<1	<1	1.61

* (p<.05)

** (p<.01)

TABLE 8

Interaction Table: Q x A

Problem type a + b: N

Q \ A	No Aids	Aids	Means
Low	48	77	63
High	75	80	78
Means	67	79	73

FIGURE 1.

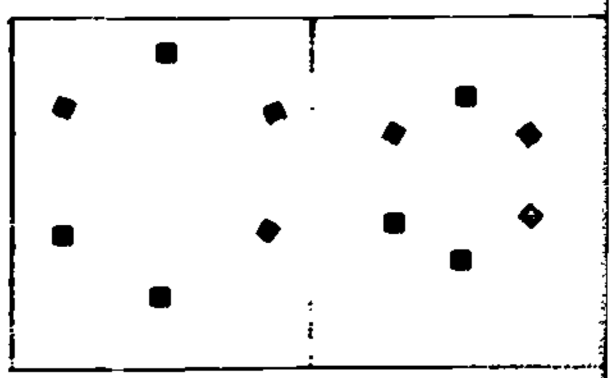

TABLE 9

Correlation Matrix

	a+b: A	a-b: A	a+x: A	x+a: A	a+b: N	a-b: N	a+x: N	x+a: N
a+b: A	1	.49	.55	.55	.74	.50	.45	.57
a-b: A			.70	.69	.52	.74	.70	.64
a+x: A				.77	.59	.64	.79	.70
x+a: A					.52	.70	.80	.76
a+b: N						.51	.58	.52
a-b: N							.75	.67
a+x: N								.73