DOCUMENT RESUME

ED 137 139 SE 022 428

AUTHOR Osborne, Alan R., Ed.
TITLE Investigations in Mathematics Education, Vol. 10, No. 1.
INSTITUTION Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE 77
NOTE 71p.
AVAILABLE FROM Information Reference Center (ERIC/IRC), The Ohio State University, 1200 Chambers Rd., 3rd Floor, Columbus, Ohio 43212 (Subscription $6.00, $1.75 single copy)

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Abstracts; College Mathematics; Elementary School Mathematics; *Elementary Secondary Education; *Higher Education; *Instruction; *Learning; *Mathematics Education; Number Concepts; Research Reviews (Publications); Secondary School Mathematics

ABSTRACT

Eighteen research reports related to mathematics education are abstracted and analyzed. Studies include elementary, secondary, and college mathematics education areas. A majority of the studies relate to instruction and learning. Research related to mathematics education which was reported in RESOURCES IN EDUCATION and CURRENT INDEX TO JOURNALS IN EDUCATION between October and December, 1976, is listed. (RH)

INVESTIGATIONS IN MATHEMATICS EDUCATION

Published quarterly by

The Center for Science and Mathematics Education
The Ohio State University
1945 North High Street
Columbus, Ohio 43210

With the cooperation of the ERIC Science, Mathematics and Environmental Education Information Analysis Center

Volume 10, Number 1 - Winter 1977

Subscription Price: $6.00 per year. Single Copy Price: $1.75
Add 25¢ for Canadian mailings and 50¢ for foreign mailings.

A note from the editor . . .

I have recently received a preliminary announcement of the formation of a new organization concerning research in the psychology of learning and teaching mathematics. The formation of this interdisciplinary group is an outgrowth of the activities of one of the working groups at the Third International Congress on Mathematical Education (1976) at Karlsruhe this past summer. The major goals of the organization will be:

"1. To promote international contacts and exchange of scientific information in the area of mathematical education.

"2. To promote and to simulate interdisciplinary research in the aforesaid area with the cooperation of psychologists, mathematicians and mathematics teachers.

"3. To further a deeper and more correct understanding of the psychological nature of teaching and learning mathematics and the implications thereof.

"4. To design new productive methods for the psychological training of mathematics teachers."

These goals are perceived to be addressing the problem of "how to improve the strategy of psychological research in order to make it useful for mathematical education."

Members will be sought among "all those who by virtue of their training or professional activity are interested in the psychology of learning mathematics (psychologists and mathematicians, college professors, teachers and researchers)."

Professor Efraim Fischbein of Tel Aviv University was elected president of the provisional planning committee in Karlsruhe. He would welcome inquiries concerning the nature of this new group and its expected future activities. His address is School of Education, Tel Aviv University, Tel-Aviv, Israel. Among other activities being considered is a first meeting to take place sometime during 1977.

It is heartening to recognize the interest that is developing in the psychology of learning and teaching mathematics. Heinrich Bauersfeld's summary paper at Karlsruhe entitled "Research Related to the Mathematical Learning Process" appears to have struck a responsive chord in stating "The reality of mathematics instruction and mathematical learning particularly has been influenced by research outcomes on a very small scale." I believe that mathematics education can come of age as a discipline only by concerted, collaborative interchanges similar to those envisioned by this group. Thus, it is easy to support the development of this new international organization concerned with the learning and teaching of mathematics.

Alan Osborne
Editor

4

Winter 1977

THE ANALYSIS OF A PROPOSITIONAL LOGIC GAME BY ORDERING THEORY. Airasian, P. W.; Bart, W. M.; Greany, B. S. Child Study Journal, v5 n1, pp13-24, January 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Theodore Eisenberg, University of Northern Michigan.

1. Purpose

The purpose of this study was to organize hierarchically a set of propositional statements. Specifically, a set of sixteen conjunctive and disjunctive statements were hierarchically ordered by the relationship: "Understanding statement i is prerequisite for understanding statement j."

2. Rationale

The authors claim that research has verified that Piaget's stages of human cognition (sensory-motor, preoperational thought, concrete operations, and formal operations) are distinct, occur in their posited order, and that each stage must be attained prior to a subsequent stage. However, little attention has been focused upon the sequence of cognitive developments within a single stage. Therefore, this study investigated the extent to which skills representative of the formal operational stage are hierarchically ordered.

3. Research Design and Procedure

A game entitled "Butch and Slim" was individually played with 60 children (34 males, 26 females) who were approximately 14 years of age. It was assumed, because of their age and Piaget's findings, that these children were operating within the formal operational stage. It was further assumed that the disproportionate sex representation could be ignored because no sex differences in formal operational thinking have been identified in previous research.

Butch and Slim were two criminals being interrogated as to whether or not they robbed a bank. The subjects were asked 16 questions dealing with compound propositions. The response options were four cards, each with a drawing of Butch and Slim and on each a statement representing the four possible combinations: (1) Butch, yes; Slim, hes, $(p \wedge q)$; (2) Butch, yes; Slim, no, $(p \wedge \tilde{q})$; (3) Butch, no; Slim, yes, $(\tilde{p} \wedge q)$; (4) Butch, no; Slim, no, $(\tilde{p} \wedge \tilde{q})$. If Butch said, "Slim and I robbed the bank together," then card (1) was the correct response. If Butch said "I am certain that Slim did not rob the bank," then cards 2 and 4 constituted the correct response. (Note that $(p \wedge \tilde{q}) \vee (\tilde{p} \wedge \tilde{q})$ is equivalent to $\tilde{q}$.) For most statements a correct response required the selection of more than one card. A subject had to select all the cards comprising a correct response to be given a score of 1 for the item; otherwise he or she received a zero on the item.

No *a priori* ordering of the items was hypothesized. Subordinancy was determined by the frequencies in the (0,0), (1,0), (0,1), and (1,1) cells for each item pair (i,j). The subordinancy algorithm was not well defined.

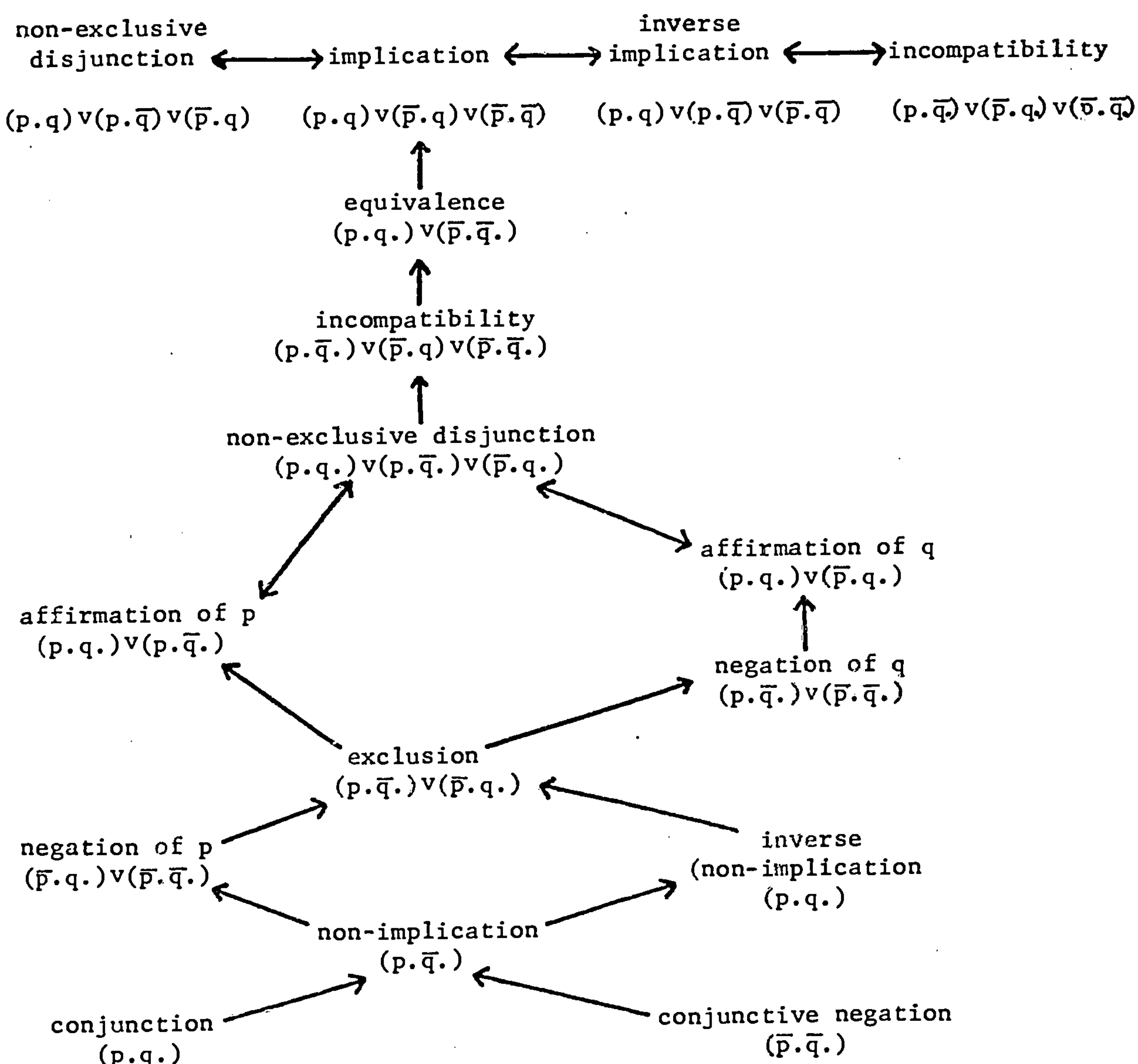
All possible subordinancies were considered and then (somewhat mysteriously) arranged into a hierarchy of dependencies. The hierarchy was tested for reproducibility using Torgerson's algorithm.

4. Findings

The hierarchy of dependencies is presented in Figure 1.

Figure 1

Hierarchy Among Logical Statements



"The reproducibility coefficient indicates the extent of instances where a subject failed an item at a lower hierarchical level but passed an item at a higher level for which the failed item was prerequisite" (p. 22). The reproducibility for this hierarchy was 0.94.

5. Interpretations

> "The general finding that atomic propositions are prerequisite to bi-atomic propositions which in turn are prerequisite to tri-atomic propositions does lend credence to the hypothesis that the psychological order of comprehension for propositions has an additive atomic basis." (p. 22)

Critical Commentary

Mathematics educators have long been concerned with identifying dependency relationships and using them in developing curriculum materials. This study employed an interesting game to ascertain dependency relationships in a specific set of logical statements. Nevertheless, there are several flaws in the reporting of this research.

(1) The algorithm used for determining task dependency was not well defined. This was a serious oversight on the part of the authors and the referees of the journal. Indeed, there seemed to be two different algorithms: (a) If $f(1,0) \geq 0.9(60-N)$ and $f(0,1) < 6$, then i is subordinate to j (p. 16). (b) Item i is subordinate to j if, and only if, $f(0,1) < 6$ (p. 17).

(2) The concept of not having an *a priori* specification of the hierarchy merits special study. For example, consider the tasks:

i) The person can identify the integers one through ten.
and
j) The person can do ten pushups.

One would expect $f(1,0)$ to be high and $f(0,1)$ low. But is i prerequisite for j? What are the implications for teaching under such an assumption? By not specifying at least a skeletal hierarchy, one can subordinate unrelated tasks.

(3) Other researchers have shown that task subordinate hierarchies are population dependent. For example, consider task i the ability to add algebraic fractions and task j the ability to integrate a function by partial fractions. If the tasks were given to a group of high school algebra students, one would expect $f(1,0)$ to be high. For a group of grade school students or of calculus students, $f(0,0)$ and $f(1,1)$ should be high. In the former case we conclude that i is subordinate to j. In the latter case we are instructed by this study to consider them equivalent.

Since hierarchies are population dependent, the results of this study might be different when tested on other populations. The authors' claim of 0.94 validity seems spurious.

(4) The hierarchy seems to appear out of the air. The algorithm for its construction is not stated. Moreover, it is implied that all dependencies are of equal strength--and that certainly cannot be true.

(5) It is implied that inverse non-implication $(\tilde{p} \wedge q)$ and negation of p ($(\tilde{p} \wedge q) \vee (\tilde{p} \wedge \tilde{q})$) are equivalent since they are at the same level in the hierarchy, and are both subordinate to affirmation of q, $(p \wedge q) \vee (\tilde{p} \wedge q)$. This seems to be reversed from previous studies which have shown that the length and the number of negations in a structure are two variables which influence the difficulty of the structure. It seems that with few exceptions, one can eyeball the levels in the hierarchy presented: the more disjunctions the higher the level.

The strength of this study lies not in the results, but in the procedure. It exposes one to the concept of not specifying a hierarchy *a priori,* to Torgerson's reproducibility coefficient, and to a very interesting game. These deserve further study.

Theodore A. Eisenberg
Northern Michigan University

AN INVESTIGATION OF ONE ALTERNATIVE BASIS FOR TEACHING MATHEMATICS. Arnold, William R. Colorado Journal of Educational Research, v15, pp38-46, Spring 1976.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Terrence G. Coburn, Oakland Schools, Pontiac, Michigan.

1. Purpose

The study sought to determine whether a teacher could efficiently and effectively develop, execute, and evaluate a mathematics program that would adequately enable pupils to learn. Subsidiary, related questions addressed were:

a. To what degree can a teacher use certain models of learning activities of a mathematics program, foregoing any reference to a school's existing program?

b. To what degree are children taught by a teacher's program, as opposed to being taught by an existing school program, able to: (1) master basic skills, (2) comprehend mathematical concepts, (3) solve mathematical problems, and (4) use induction to produce mathematical generalizations?

2. Rationale

The author maintains that mathematics education in the recent past has failed. He cites declining standardized test scores and some literature which implies that the basis for mathematics education circa 1955-70 was inadequate. The author's major criticisms of mathematics education appear to center on the use of commercial textbooks, the traditional grade placements of topics, the rigor and formalism of the new mathematics, and some of the pedagogy related to ability grouping and self-paced instruction. The purpose of his study is to seek alternative bases for curriculum and pedagogy.

3. Research Design and Procedure

No formal design was used in this action research. The subjects in the experiment were thirty fourth-grade pupils in a self-contained classroom in a rural elementary school. The teacher was a university professor who taught for just the mathematics period (60 minutes, 5 days per week).

The researcher-teacher determined the content, sequence, and pacing of the mathematics program. He also created the instructional materials and activities during the experimental period, which lasted eleven weeks (October through December, 52 class periods). Lesson plans were constructed for all of the developmental activities; worksheets were constructed for drill, and problem or activity cards were used for enrichment. A "laboratory table" was established for readiness activities. No textbooks or other commercially prepared materials were used.

The pupils investigated a wide-ranging assortment of topics including quite a bit of geometry, e.g., transformations, symmetry, and coordinate systems.

Pupils were involved in readiness, developmental, drill, and enrichment activities. The sequence related to order of involvement and not to mastery of skills. Each child had an opportunity to learn each concept and acquire related skills at least two times in each of four different types of activities.

The selection, order and presentation of concepts and skills was conducted according to models developed by the researcher:

A Model of Mathematics Education

I. The function of the teacher is to develop, execute, and evaluate learning activities. A program is a sequence of learning activities, perhaps partitioned into units.

II. Learning activities are created by having the teacher coordinate and manipulate five teaching variables. The five teaching variables are (a) the concept to be learned; (b) the conditioned learning environment; (c) the teacher's behavior (teacher moves); (d) the pupil's behavior (pupil actions); and (e) the vehicle. As these variables are manipulated, what and how children are taught are changed.

III. By manipulating the variables listed in (II) above, four distinct types of learning activities can be created and, subsequently, executed and evaluated. The four types are: (a) readiness activities, (b) developmental activities, (c) drill activities, (d) enrichment activities. Collectively, the learning activities provide for consideration of major domains and dimensions of learning.

IV. The effect of involving children in a sequence of learning activities is evaluated. This required utilization of check lists, tests, and observation. Diagnosis, if deemed necessary, is part of evaluation.

V. Learning activities are executed in accordance with four principles. Specifically, the teacher assumed a leadership role; children learn actively; children have choices within constraints; and both concepts and skills are emphasized.

In conjunction with the preceding model, the researcher also used a 3-dimensional arrangement consisting of "categories of learning", "categories of activities", and "concepts to be taught."

The study was made realistic in that no more than 30 minutes was allotted to planning each day. Planning was done in the morning before class.

Three examinations, one for each of 3 units of study, were administered. Each contained items to assess skills, comprehension, and problem solving; "a third type assessed ability to use inductive processes to produce a generalization". Aside from these examinations, informal evaluation was used to assess pupil work "in progress".

4. Findings

(a) Teacher's development, execution, and evaluation of a mathematics program.

Total planning time was 760 minutes for a mean development time per period of about 14 minutes. Seventy-two learning activities were developed (ten minutes per activity, average).

Only 3 of the 30 pupils had consistent difficulty attending to or being at a task.

The researcher found the ongoing, informal evaluation of pupil progress to be "efficient and effective".

(b) Performance on the testing. The researcher-selected criterion for satisfactory achievement was 80 percent.

The results (in percentages) are given below:

UNIT TESTS	I	II	III
Pretest results	21	12	7
Posttest results	20.8	85.7	83.4

5. Interpretations

The researcher feels that the pupils were "successful" and it is possible for a teacher to develop, execute, and evaluate a program particularly suited to that teacher's circumstances. This finding coupled with the supposed failure of established programs suggests to the researcher that many alternative bases for mathematics education exist. He urges that mathematics educators seek to identify and research potential alternatives.

Critical Commentary

This study has little significance for mathematics education. The researcher's suggestion that a teacher-developed program can be successful is neither startling nor revealing. Teachers commonly develop units (sans textbook) for their pupils and pupils frequently learn from this instruction. The description of this study, however, is not sufficiently clear as to what was learned nor did it indicate the stability of the learning. The limited nature of the sample and the lack of controls obfuscates interpretation of the results.

The researcher has shown commendable skill in implementing his education models for developing instructional materials and activities. I disagree, however, with the implication that a teacher should forego commercially prepared materials in favor of developing worksheets and activity cards. The study does have some implication for inservice and preservice education. Teachers should be encouraged to supplement their commercially prepared program with materials and activities designed for their pupils.

Terrence G. Coburn
Oakland Schools

STORY PROBLEMS: MERELY CONFUSING OR DOWNRIGHT BEFUDDLING? Blankenship, Colleen S.; Lovitt, Thomas C. Journal for Research in Mathematics Education, v7 n5, pp290-299, November 1976.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by W. George Cathcart, University of Alberta.

1. Purpose

This study was designed to answer two questions:

(a) What is the effect of manipulating problem vocabulary on students' accuracy and speed?

(b) Will the techniques of Reread, Write, and Re-examine Solution improve students' accuracy in solving story problems?

2. Rationale

Earlier studies have shown that problems with extraneous information are very difficult for educable mentally retarded students. The performance of other types of exceptional children on problems of this nature has not been well researched. In addition there have been only a few attempts to modify directly the performance of students through different teaching techniques. The clinical approach was used because this approach to studying problem solving was recommended in 1969.

3. Research Design and Procedure

Seven boys, ages 9-12, of normal intelligence but functioning at least one year behind their peers in reading and arithmetic, were asked to work through 12 classes of story problems. Classes of problems were developed by varying the number and type of nouns (3 classes), presence and placement of extraneous information (3 classes), type or tense of verb (2 classes), presence of an introductory sentence (1 class), and mode of numeral (1 class). Each class contained addition only, subtraction only, and addition/subtraction basic facts with the basic facts randomly assigned to problem sheets within each class. Subjects had 100% accuracy on the basic facts.

Each day students were given 15 problems to solve. All students worked through the 12 classes in the same order. The criterion for moving from one class to the next was three successive daily assignments at 100%. The time to complete each daily assignment was taken by the students themselves.

Each class of problems was introduced by a baseline phase that provided a measure of the students' ability to solve that class of problems.

If a student did not reach criterion within 6 days, a "reread" technique was started on the 7th day. In this technique the student was asked to read each incorrectly solved problem orally and give the correct answer. If criterion was reached within 6 days, the next class of problems was assigned; if not, a "write" technique was scheduled on the 13th day. In this technique the student had to write out each incorrectly solved problem and reanswer it. If criterion was not met within 6 days, a "re-examine solution" technique was started on the 19th day. Here the student had to read orally each incorrectly solved problem, explain how he or she solved the problem, and offer an alternate solution. If criterion was not met the "reread" technique was reported.

Satisfactory reliabilities were obtained on teacher procedures, accuracy of students' timing, and paper correction.

4. Findings

A comparison of total mean baseline performance to performance in each class of problems yielded the following results:

(a) Varying the number and type of nouns was usually associated with a high degree of accuracy and a slight reduction in computational speed.

(b) Problems containing the passive verb tense "was given" did not impair accuracy nor computational speed. However, the presence of verbs such as "purchased" or "bought" was associated with a decrease in computational speed and accuracy and an increase in number of days to reach criterion.

(c) The use of an introductory sentence was associated with a slight reduction in computational speed; however, accuracy was not impaired.

(d) Changing the form of verbal cue, the question, or the mode of numeral did not impair computational speed or accuracy.

(e) Extraneous information was first introduced in a constant position. Its presence was associated with decreased accuracy and computational speed. When the placement of the extraneous information varied, the pupils' performance was further impaired. Once the students had mastered extraneous information in varied positions, its effect on their performance on similar problems that differed only in their use of pronouns was less severe.

The performance of one typical student was described in some detail in the article.

The application of the three teaching techniques enabled all students to master classes of problems that were originally difficult for them. The reread, write, and re-examine solution techniques were used 30, 11, and 3 times respectively.

5. Interpretations

Blankenship and Lovitt list two limitations to the study, a possible ordering effect for the problem classes and for the teaching techniques. Six EMR students worked through the first seven classes of problems in reverse order and their performance was similar to the performance of students in the study.

The inability of the students to solve problems containing extraneous information might be explained by the "rote computational habit" theory. This implies that children do not read the entire problem, but read only until they encounter numbers and then use the numbers to arrive at a solution. If this is true, the authors suggest that teachers ought to use a teaching technique which stresses the careful reading of problems.

Students' accuracy averaged 95.6% on a mixture of twelve problem classes. The authors say that this indicates that students could discriminate between classes of problems.

The following method of instruction is recommended to teachers: identify the problems to teach; "arrange" problems of the selected types from different arithmetic textbook series; begin with one class of problems and continue until mastery is achieved before beginning the next class; and use procedures such as rereading, writing, and re-examining solutions.

Critical Commentary

The students used in this study are to be pitied. It is difficult to imagine anything more boring than working 15 artificial problems a day, day after day, and if a mistake were made, being told to reread the problem, write it out, or tell how the solution was obtained and suggest a different solution. This may be forgiveable in a research study but it is going much too far for the authors to recommend this routine to teachers in general. One explanation for the difficulty children have with story problems, as traditionally taught, might be that they are just plain dull. If this is true, then this study adds nothing to our search for better teaching techniques.

Most of the quantitative data are presented in one table which is difficult to interpret. For example, for one class of problems students were able to solve correctly an average of 27.2 problems per minute. This is one correct solution every 2 to 3 seconds. Impossible! Therefore, all the data in this column must reflect the score for all 7 subjects combined. Applying the same logic to the column headed "days to criterion" presents another impossibility. The most difficult problem class took an average of only 14.7 days to criterion. This is an average of just over 2 days per student. This implies that for the majority of students, none of the teaching techniques would have been applied, which is not the case. It seems, therefore, that in one case the data are the total for all subjects and on the other hand they are the mean. A little confusing!

The authors recognize the limitations due to ordering effects. In addition to ordering effects, the amount of practice children had with earlier classes of problems, regardless of order, could inflate the baseline data for later classes.

The three teaching techniques were considered effective. An alternate explanation might be that children got the answers correct simply by chance or by trial and error. They worked 15 problems from one class per day for 6 days before the reread technique was applied. Each of the three techniques were applied for a maximum of 6 days each. From doing such a large number of similar problems sooner or later children should hit on the right combination of numbers and operations regardless of the kind of teaching technique used.

A high score on a test containing mixture of problems from each of the 12 classes is interpreted to mean that children could discriminate between classes of problems. Is the ability to discriminate among the classes based on such things as type of noun, tense of verb, extraneous information, numeral mode, et cetera, a necessary condition for solving story problems?

Good research on problem solving is greatly needed. The authors are to be commended for looking at a problem of great concern to mathematics educators. However, we need more imaginative and interesting research than this.

W. George Cathcart
University of Alberta

SYSTEMATIC ERRORS IN THE FOUR VERTICAL ALGORITHMS IN NORMAL AND HANDICAPPED POPULATIONS. Cox, L. S. Journal for Research in Mathematics Education, v6 n4, pp202-220, November 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Doyal Nelson, University of Alberta.

1. Purpose

To identify the most frequently occurring systematic errors in whole number addition, subtraction, multiplication, and division algorithms.

2. Rationale

Although the analysis of errors in the use of computational algorithms has a long history, no studies have specifically analyzed the systematic errors by skill levels, by grade levels, and by normal and handicapped samples. Cox has set out to do such an analysis.

3. Research Design and Procedure

Levels of computational skill were identified for the four operations. There were eight levels for addition, six for subtraction, ten for multiplication, and ten for division. These levels were determined by number of digits involved and whether there was carrying or borrowing. To determine systematic errors, each child in the sample was given five sample exercises for each of the levels of computational skill. Following is one example of such a set of five sample exercises.

$$\begin{array}{r} 24 \\ \underline{+\ 67} \end{array} \qquad \begin{array}{r} 56 \\ \underline{+\ 28} \end{array} \qquad \begin{array}{r} 17 \\ \underline{+\ 13} \end{array} \qquad \begin{array}{r} 25 \\ \underline{+\ 16} \end{array} \qquad \begin{array}{r} 32 \\ \underline{+\ 29} \end{array}$$

If a child made the same error in three or more of these five examples, it was classified as a systematic error for the particular algorithm and level. Any other error pattern was considered to be random or careless.

There were 564 children in the normal sample ranging in grades from 2 to 6. One hundred eighty handicapped children made up the other sample. The handicapped children were classed as Primary Special Education, Intermediate Special Education, or Junior High Special Education.

Each week from September on, data sheets from one of the computational levels were distributed to the classrooms and the test therein administered to the children. For any child's paper to be analyzed, he or she must first know the basic facts associated with the particular algorithm and must have had instruction in the algorithm involved. Data were gathered first for addition and subtraction and later in the year for multiplication and division.

4. Findings

The percentages of systematic errors for the normal sample were 4% for addition, 11% for subtraction, 5% for multiplication, and 5% for division. The percentages for the handicapped group for the four operations were 5%, 17%, 19%, and 17%. It should be noted that none of the second-grade children in the normal sample qualified for multiplication or division tests at any level. It should also be noted that none of the primary special education handicapped children were trained to do multiplication or division at any level.

Two hundred twenty-three systematic errors occurred for the four algorithms: 51 in addition, 67 in subtraction, 52 in multiplication, and 53 in division. The errors were classified according to type of dysfunction. The table in the report showing the classes of systematic errors is reproduced below.

Classification of Nature of Dysfunction Resulting in Systematic Errors for the Four Algorithms

	No. of Different Systematic Errors		No. of Different Systematic Errors
Addition		Subtraction	
Renaming	23	Renaming	28
Concept of Addition	17	Concept of Subtraction	16
Wrong Operation	6	Wrong Operation	6
Place Value	5	Place Value	2
Total	51	Total	52
Multiplication		Division	
Concept of Multiplication	19	Concept of Division	21
Partial Products	13	Estimation	9
Mult. Process After Renaming	10	Partial Quotients	5
Add. Process After Renaming	7	Remainders	5
Renaming	6	Zeros in Quotients	5
Mult. with Zeros	6	Errors in Mult. or Sub.	4
Wrong Operation	4	Zeros in Dividend	2
Reversal of Digits	2	Partial Dividends	2
Total	67	Total	53

To answer the question "Do children who make systematic errors continue to make them one year later?", there was a follow-up study. One hundred fifteen children who had previously made systematic errors and who were still available were retested. Twenty-three percent of this group were making the identical error or some other error one year later.

20

The most common error was in subtraction and occurred because the smaller number was always subtracted from the larger as in this example:

$$\begin{array}{r} 43 \\ -\ 17 \\ \hline 34 \end{array}$$

Tables (5-8) showing the details of systematic errors which occurred were presented in the original report (they have not been reproduced in this summary). Two or three examples of each error were given in the tables along with descriptions in words of how the error was made.

5. Interpretations

Cox indicates that there is a possibility that some systematic error types were not identified in the analysis. However, most of them seem to have been identified. In the analysis, it is observed that the children making errors seemed to know that some procedure or pattern had to be followed in the algorithm but simply adopted the incorrect one.

Various errors fell under the general categories of misconceptions of the nature of number, the nature of the operation, the function of place value, or the function of renaming. Cox also notes that in almost every case the number facts were correct but the process involved in using the algorithms was wrong. She urges teachers to recognize the problem when systematic errors occur in algorithms and not to mistake it for lack of knowledge of basic facts. More drill on basic facts in these cases would do more harm than good.

Systematic errors were still occurring one year later. Although no measures were taken in this study to find out if specific instruction was given to eliminate them, it appears that when systematic errors occur they have the potential to persist. There is a definite pattern in systematic errors as compared with random errors,so proper diagnosis and remedial instruction should be effective means of eliminating them.

Critical Commentary

The results of the study reported here would provide a reliable guide for teachers interested in the diagnosis of systematic computational errors. The method of finding systematic errors appears to be a good one and the illustrations given make it easy to identify the error type. My recommendation would be that every teacher involved in the teaching of computational algorithms should obtain copies of Tables 5, 6, 7, and 8 in the original report and that these tables be kept where they are immediately accessible.

Although Cox obtained information on the errors made by normal and handicapped samples, there is nothing in the discussion to help the reader understand the rather striking differences in the computational behavior of these two samples. I would have liked for Cox to suggest reasons for these differences and to suggest how remedial procedures might differ for normal and handicapped children.

In spite of such criticism, this represents the type of research that is sorely needed in mathematics education and Cox is to be commended for her design and analysis.

Doyal Nelson
University of Alberta

THE EFFECT OF SCHOOL ATTENDANCE ON INTELLECTUAL DEVELOPMENT IN NORTHERN NIGERIA. Fahrmeier, Edward D. Child Development, v46, pp281-285, March 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Jerry P. Becker, Northern Illinois University.

1. Purpose

The purpose of the study was to collect data, the analysis and interpretation of which would throw light on the question, "Does schooling have an effect on cognitive development?" A related working hypothesis was that performance should improve at a faster rate for the children who attend school than for those who do not, and the effect of schooling should be cumulative with time.

2. Rationale

The researcher states that the question, "Does schooling have an effect on cognitive development?", cannot be answered in societies with universal education. Thus, an opportunity existed to investigate the question in a city in Nigeria with a significant number of both schooled and non-schooled children available as subjects.

Other such studies have been carried out, but results have not provided a clear answer to the question. Earlier investigators using a small number of tasks have found major differences or no difference, whereas investigators using a greater variety of tasks have found mixed results. The present investigation involves study of a wide range of skills with a large sample using individual testing in the primary language of the subjects. Since the school system in which the research was being done was expanding and educational opportunity was not open only to the socially elite, the investigator throught the selection bias was minimal. Thus, the study possessed a number of advantages over other earlier ones. The study also related to a number of factors presented by Scribner and Cole (1973) which they believed should be kept in mind in order to explain the effect of schooling on the cognitive development of children in various cultures.

3. Research Design and Procedure

Subjects were children between the ages of 6 and 13 who lived in the central ward of the city in which the data were collected. Nearly all 230 school children were located and tested; only 175 of 270 unschooled children were given any part of the test battery, since it was difficult to find them and some were unwilling to cooperate. Of the 175 children, 30 received the entire battery of tests, but 70 or more protocols were obtained for each test given to the entire age group.

Ages of 100 of the children were obtained from the town birth register. For the others, however, the height of the children was used to provide an estimate of age (correlation = .92); this estimate was averaged with the father's estimate and the result in each case was considered as a child's age.

Two types of tests were used in the study. The first type touched on those areas necessary for successful activity in the culture: vocabulary, arithmetic, general information, and memory. A second type was a set of tests developed in the West and previously used in this type of study (e.g., Kohs Blocks, Raven's Matrices). Skills related to literacy but not including reading and writing, such as attention to detail when examining drawings, strategies for remembering, experience with tests, and willingness to perform arbitrary tasks, were also used. Twelve tests were used in all: six were developed by the investigator and six were tests developed in the West.

All tests were individually administered. The tests were given to the entire age range, except for three which were thought appropriate only for children of at least age 9. Four male Nigerian secondary-school graduates, fluent in the primary language, administered all tests. Testing conditions were similar for both schooled and unschooled children. All of the tests were administered in a time span of two months.

An analysis of variance was carried out on the data with main factors of age, schooling, and sex. Subjects were classified into four age levels to reduce the effects of error in the age estimates. Orthogonal polynomials were used with the age-level variable in order to determine age trends in the data.

4. Findings

The data analysis showed that the expected age and schooling effects were generally present, but that the expected age-by-schooling interaction did not generally occur. The interactions that did occur were regarded as artifacts. More particularly, however, there were no age and schooling effects on the Spatial Relations test; there were age but not schooling effects on the Figure Grouping and Memory of Objects tests. Sex differences were found favoring the girls on the Figure Grouping test and favoring the boys on the Cube and Memory of Objects tests. This latter effect, favoring the boys, was not duplicated on other similar tests. Also the superior performance of the girls on the Memory of Objects test was explained by their probable greater familiarity with the material. The expected advantage of schooled children on the Western tests did not emerge. Four of the remaining tests showed evidence of either ceiling or floor effects. Finally, unexplained poor performance among the oldest, unschool children as compared to younger unschooled children was exhibited on one other test.

The investigator summarizes his findings by commenting that age and schooling effects appear on most tests, but the difference between schooled and unschooled children tends to be constant at all ages, rather than increasing as schooling increases.

5. Interpretations

There was an initial difference between children starting in school and children not going to school favoring those starting school. The investigator explains this by commenting that school starters may be more intelligent, alert, or attentive than the others. He comments that even though the unschooled children seem to be at a disadvantage at age 6, the rate of improvement is, in general, the same as that of the schooled children. Since one might expect more rapid improvement among schooled children if the selection effect was based on differences in native intelligence, a more likely explanation is that the initial difference between schooled and unschooled children is a motivational one based on a desire to do well on things which seem more important because of school attendance. The motivational edge would be maintained over time.

The investigator feels there is clear evidence that the skills represented by the tests are not peculiar to the school setting. While schools might be responsible for literacy, they do not seem to be engendering cognitive growth that would set schooled children apart from unschooled children. Finally, the investigator comments that the experience of attending classes in which topics are handled in a more abstract manner than might be typical of everyday life, which is a likely characteristic of all school systems, seems not to have a major influence on cognitive growth.

Critical Commentary

A. I do not see any direct relevance of the study to mathematical education. Perhaps an examination of the testing instruments would change this observation, but I doubt it.

B. The investigator comments that there is clear evidence that the skills represented by the tests are not peculiar to the school setting. This would seem to have to be true, to the reviewer, if unschooled children are performing with any degree of success on the tests. It should be noted that there may be many aspects of cognitive behavior treated by the curriculum and/or teachers on which the tests did not touch. Thus, the investigator should have been more careful in commenting that schools do not seem to be engendering cognitive growth that would set schooled children apart from unschooled children.

C. The tests themselves seemed to exhibit results which were questionable in one way or another and thus, in the mind of the reviewer, the results should be viewed with some skepticism.

D. I would tend to be even more careful than the investigator regarding generalizability of the study. If the results in the study were known to be valid, which they are not, even then the results could possibly be attributed to a poor school system in which the study took place.

E. The investigator makes reference to the study being a cross-cultural study. In what sense is it cross-cultural? This is unclear in the mind of the reviewer.

F. No reference was made to the process by which the investigator developed his tests. This information would be valuable. Also, it would be valuable to have seen some sample items from the tests in the research article.

Jerry P. Becker
Northern Illinois University

COGNITIVE EMPHASIS OF GEOMETRY TEACHERS' QUESTIONS. Friedman, Morton. Journal for Research in Mathematics Education, v7 n5, pp259-264, November 1976.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by James M. Moser, University of Wisconsin - Madison.

1. Purpose

The study was undertaken to determine what relative emphasis teachers of high school geometry give to each of four categories of questions: Memory, Comprehension, Application, and Higher Level.

2. Rationale

In studying teaching strategies, some researchers have analyzed classroom behavior using an observational system based on various categories. Those who were concerned with the intellectual component have used systems in which the question is the basic unit of analysis. These systems have evolved from Bloom's Taxonomy. Sanders has modified the Bloom model to develop a taxonomy of questions for classroom use. The author used the Bloom-Sanders model as the basis for his own system. Because the teachers involved taught geometry, a subject that traditionally stresses deductive reasoning, it was hypothesized that there would be fewer straight memory or recall-of-facts type of questions than had been reported by Kleinman, Gusaak, Godbold, and Gall dealing with other subject matter areas.

3. Research Design and Procedure

Thirteen teachers in eight different schools in New York City and its suburbs were asked to tape-record three different lessons in which a required theorem was introduced and proved. The theorems were the Isosceles Triangle Theorem, the 180° Angle Measure Sum in a Triangle Theorem, and the Hypotenuse-Leg Right Triangle Congruence Theorem. The teachers' questions occurring in each of the three lessons were coded by the investigator directly from the tapes using the four categories listed in Section 1. Objectivity for coding was checked using inter-observer agreement between the observer and a high school geometry teacher. An agreement of 94 percent was reported. Six of the teachers were only able to secure two usable recordings leading to a total of 33 class sessions available for analysis. By computing mean scores for each teacher in each of the four question categories, differences in question categories were analyzed using a one-way repeated measures ANOVA design.

4. Findings

The total mean number of questions asked was 52.6. Of these, 30.7 were Comprehension, 12.7 were Memory, 9.1 were Application, and 0.1 were Higher Level. ANOVA showed significant differences between categories at the .01 level. Duncan's Multiple Range Test showed that significantly more Comprehension questions were asked than questions at any other level.

In addition, Memory and Applications questions were asked significantly more times than Higher Level questions, but there was no significant difference between the frequency of Memory and Application questions.

5. Interpretations

The relatively small number of Memory questions asked suggests that geometry teachers may be making greater intellectual demands upon their students than teachers of other subjects. If it is true that geometry teachers ask their students to think on a higher intellectual level, researchers who are interested in the cognitive emphases of teachers' questions should look to the geometry classroom to determine if they have found a productive focus for their investigations.

Critical Commentary

This study was extremely limited in its scope. As the researcher duly noted, the teachers participated on a voluntary basis in lessons in which a required theorem was considered. These lessons may be very specialized and not typical of most lessons. Should the lessons have been randomly selected for analysis? The article did not describe how the thirteen teachers were selected. Even though they all volunteered, were they randomly selected from a larger population of teachers who were willing to participate? Were these teachers known in advance to the researcher as superior teachers? I am afraid that the results are not too generalizable, and thus are of very limited value to the mathematics education community.

James M. Moser
University of Wisconsin - Madison

ARITHMETIC COMPUTATION SCORES: OR CAN CHILDREN IN MODERN MATHEMATICS PROGRAMS REALLY COMPUTE? Garigliano, Leonard J. School Science and Mathematics, v75, pp399-412, May-June 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Karen Fuson, Northwestern University.

1. Purpose

This paper reports the effects of two variables in the administration of the Computation subtest of the Stanford Achievement Test. The variables examined were the effect of summer forgetting (October versus May administration of the test) and the timing of test (timed and untimed). Subjects were suburban fourth graders who had always used a modern mathematics text.

2. Rationale

In the 1970's the popular press has frequently published articles asserting that children from a modern mathematics program cannot compute. The 1968 through 1972 test scores of a New Jersey suburb seemed to support this contention. Scores on the arithmetic Concepts and Applications subtests were above average, while those on the Computation subtest were below grade-equivalent performance. A modern mathematics elementary series (publisher unspecified) had been used during this time. However, this school system felt that all possible sources of data relevant to the question of computational ability should be reviewed before accepting the tests results as conclusive. Specifically, the author investigated (1) the possibility that the score from October administrations of standardized computation tests were deflated by "summer forgetting" and (2) the possibility that modern pupils would perform as well as traditional pupils if computation tests were untimed.

3. Research Design and Procedure

All (n = 215) fourth graders in the system took the arithmetic Computation, Concepts, and Applications subtests of the Stanford Achievement Test early in October and again in late May. In May one class chosen by the principal in each school (four classes, n = 91) was given an unlimited amount of time to finish the Computation subtest. The other classes (six classes, n = 124) had only 40 minutes to finish the 39 computation problems. Computation worksheets for the untimed sample were collected and analyzed. Summary descriptive statistics are reported for these tests.

4. Findings

From October to May the mean scores on the Concepts and Applications subtests increased by one grade level, while the scores on the Computation subtest increased by 1.25 grade levels. However, the mean grade equivalent gain for the timed group was 1.02 while that for the untimed group was 1.49.

The October class means for the May untimed and the timed groups were not different, so the difference between groups was not due to initial ability. Thus an untimed administration of the test seems to result in a greater mean grade level increase.

An analysis of the percentage of children below grade equivalent level on the computation subtest reveals that in the timed group, 55% were under grade equivalent in October and 35% in May, while those figures for the untimed group were 42% and 25%. Thus a May administration results in a considerably smaller portion of students with scores below grade level equivalent.

Item analysis data indicated that the timed and untimed groups performed about the same on items in the first half of the test, with the untimed pupils scoring more than ten percentage points higher on four items in the middle of the third quarter of the test and on eight of the ten items in the final quarter of the test. This is accounted for by the fact that the untimed group attempted virtually all items (a total of five items were not reached) while the timed group failed to reach a total of 348 items.

Fifty-one of the 90 untimed working papers were complete enough for evaluation. Thirty-eight answers were incorrectly transferred to the IBM answer paper. In addition, nineteen problems were incorrectly transferred from the test booklet to the scrap paper where they were correctly computed. Thus pupils missed, on the average, one problem which in fact they had computed correctly. A single raw score point is equivalent to from .3 to .9 grade equivalents, so these technical errors do lower the scores.

5. Interpretations

Teachers in the school system were informed of the retesting program in January, so they may have increased the drill on computation somewhat. This fact combined with the group means and grade equivalent data lead the author to assert the possibility of an interaction effect between May testing, untimed testing, and increased drill by teachers.

The total of 66 children below the grade equivalent in Computation in May compares favorably with the number of children below the grade equivalent scores in the reading and other arithmetic subtests in October. Thus, "summer forgetting" seems to affect scores on computation much more than on other subtests. The author concludes that children using a modern mathematics series can compute, if they are given sufficient time to do so and if they are tested at the end rather than at the beginning of the year. He argues that students have higher scores in the areas of concepts and applications than they did ten years ago, and that this gain more than offsets the loss in speed in computation.

Critical Commentary

The questions raised by the author are insightful ones, and the carefulness of his examination of possible sources (not all mentioned in this summary) of the lower scores in computation are praiseworthy. Unfortunately one error and one seeming contradiction in his data cloud the results of his study. The error is that the classes taking the untimed test were chosen (by unspecified criteria) by the principals of their building rather than being randomly selected. The author does not acknowledge possible problems due to the non-random selection and does not even discuss the comparability of the timed and untimed groups. Fortunately, sufficient data are presented to permit the reader to make some of these comparisons for himself or herself. The timed and untimed groups had the same (.02 difference) mean grade equivalent scores on the October test, but the untimed group had only 42% of their number under the expected grade equivalent of 4.1 while the timed group had 55% under this grade equivalent. Thus there were more students just above grade equivalence in computation in the untimed than in the timed group. The effect of this difference in group composition is unclear. If one would expect students just above grade equivalence to benefit the most from extra working time, the difference may have inflated the untimed scores. Alternatively, it might be argued that students just below grade equivalence would gain the most from extra working time, so the under-representation of this type of student in the untimed group actually may have lowered the untimed scores.

Although there is no clear evidence that the pupils in the timed and untimed groups were different (and indeed the mean October grade equivalents were the same), there is some evidence that the teachers of the untimed group might have been superior. The arithmetic Concepts and Applications subtests, as well as the Computation subtest, were given in May. Pupils were allowed unlimited time only on the Computation subtest. However, the mean grade equivalent scores for the untimed group were greater than those for the timed group on all three subtests: by .63 in Computation, by .43 in Concepts, and by .58 in Applications. Because October scores of the two groups did not differ, this uniform superiority of the untimed group would seem to be due to better teaching of these pupils. Because the timed pupils failed to answer 348 items and the untimed pupils, only 5 items, the untimed conditions did contribute to the higher computation score. But the amount of this increase due to the untimed condition and the amount due to better mathematics teaching cannot be determined from the data given.

The seeming contradiction in the data pertains to whether the summer-forgetting or the timed nature of the test contributes more to the low October computation scores. If one uses mean grade equivalent scores, the timed nature of the test seems all important, for the May mean timed score is 1.02 higher than the October score, while that of the untimed group is 1.57 higher. If, however, one examined the percentage of pupils who have scores under the grade equivalent score at the time of testing, the timed and untimed groups show little difference: they both have about 20% fewer scores under the grade equivalent score at the May testing. These data seem to indicate that the time of testing, and the associated summer-forgetting problem, are responsible for the low October computation scores. The author does not acknowledge this confusion explicitly, though his final conclusion naming both variables as contributors does indicate implicit awareness of this problem.

In summary, this was a noteworthy effort to clarify the often emotional issue of the effect of modern mathematics instruction on children's computational ability. It raised several interesting alternative explanations for the standardized test results in computation and presented data which supported the argument that an end-of-the-year (versus a beginning-of-the-year) test administration and an untimed (versus a timed) test will raise computation scores considerably. Because of the ambiguities in the design and analysis discussed earlier, the relative importance of these two variables and that of teacher skill in teaching mathematics cannot be assessed. An interesting result of the analysis of the student working papers is that students miss an average of one problem (out of forty) which they have worked correctly due to miscopying of the initial problem prior to working it or to marking of the incorrect slot on the IBM answer sheet; these errors depress scores from .3 to .9 of a grade equivalent. And finally, the author does an important service by pointing out that while modern mathematics students do compute more slowly (though as accurately) as their counterparts did ten years ago, they score higher in tests of concepts and applications. Because, however much mathematics educators may wish otherwise, there will always be a fairly small and certainly finite amount of school time available for instruction in mathematics, one should expect that increases in learning in one area will be accompanied by decreases in other areas. With the widespread availability of calculators, speed in computation is certainly less important than both an understanding of the processes involved and a firm sense of when to use these processes.

Karen Fuson
Northwestern University

DIVISION INVOLVING ZERO: AN EXPERIMENTAL STUDY AND ITS IMPLICATIONS. Grouws, Douglas A.; Reys, Robert E. Arithmetic Teacher, v22 n1, pp74-80, January 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by J. Dan Knifong, University of Maryland.

1. Purpose

(a) To determine the success rate for students in grades 4, 6 and 8 on single-digit division problems, some involving zero and some not.

(b) To evaluate the effectiveness of two teaching strategies for division involving zero.

2. Rationale

Division involving zero either as divisor or dividend causes more than its share of problems among elementary and junior high school students. However, research on division has ignored this important problem and dwelled almost exclusively on the whole number division algorithm.

3. Research Design and Procedure

Six hundred seventy-two children from thirty classrooms, representing grades 4, 6 and 8 were pretested on the first day. On the second day they were given one of two instructional treatments on division involving zero. On the third day they completed a posttest. Immediately following the posttest, four students from each of the classes were interviewed concerning their understanding of division with zero. Approximately six weeks later a retention test, similar to the pre- and posttests, was given to the entire sample.

Both instructional sequences were developed from the definition of division as the inverse of multiplication. In the Division-Multiplication lesson (D-M lesson), children first reviewed single-digit division problems. Then they were presented with $15 \div 3 = \square$ and led to discover the related multiplication sentence, $15 = 3 \times \square$, and its inverse relationship. At this point children were presented with a division problem involving zero as either the divisor or dividend and led to conclude (through the help of the multiplication sentence) that in the first case there was no solution; and in the second case, that the solution was zero. The case of $0 \div 0 = \square$ was not considered in this study. The rest of the time was spent on practice.

The Multiplication-Division lesson (M-D lesson) was exactly the same as the D-M lesson except that the multiplication sentence was always presented before the related division sentence. During the practice session emphasis was on working the multiplication sentence _first_ and then turning to its division pair.

All forms of the 18-item tests contained six problems with zero as the divisor, six with zero as the dividend, and six without zero. All three of the common division symbols (e.g., $12 \div 4 = \square$, $12/4 = \square$, and $4\overline{)12}$) were evenly distributed among the test items.

4. Findings

(a) Improvement in test scores after treatment was statistically significant, but small. Although type of treatment was shown to be significant in the posttest (favoring D-M), this difference "washed out" in the retention test.

SUMMARY OF TEST MEANS

	Pre Test	Post Test	Retention Test
Total Sample	10.4	11.9	11.2
D-M Treatment Group	*	12.4	11.2**
M-D Treatment Group	*	11.4	11.2**

*This summary of means is not possible from the text. Nevertheless, it is clear that the posttest differences are not due to sample variations.
**These scores were reported as "very close to 11.2."

(b) "There was considerable variation among (the thirty) classes." Although the individual means were not reported, the upper-grade students did better.

(c) There were no sex differences.

(d) On the same test there were small differences among the notations for division.

(e) Students were _roughly_ 90% successful with problems not involving zero, but only _about_ 50% successful with problems involving zero, _even after treatment_. After treatment students improved their scores on problems involving zero as the divisor but missed more problems involving zero as the dividend.

MEAN SOLVING PERFORMANCE
BY TEST AND THE OCCURRENCE OF ZERO

	Zero is the dividend	Zero is the divisor	Zero is not involved
Pretest	4.79	.32	5.33
Posttest	3.20	3.00	5.71
Retention test	3.16	2.46	5.60

5. Interpretations

The data are interpreted to indicate:

(a) Focusing on the division sentence first and then on the multiplication sentence (the D-M sequence) is a better instructional strategy than vice versa.

(b) Instruction on division involving zero might best be delayed for some fourth graders.

(c) Systematic practice and review are needed to retain concepts.

(d) Children will perform best using notation with which they are most familiar. Hence, each of the notations should be used in division lessons.

(e) Children in grades 4, 5 and 6 can learn division involving zero.

(f) "Children tend to over-generalize." The example is given of the posttest case where the children reasoned that $0 \div 5 = \square$ has no solution because $5 \div 0 = \square$ has no solution.

(g) "Using the inverse relationship between multiplication and division seems to hold much promise in helping these children (learn cases involving zero.)"

Critical Commentary

The overgeneralized interpretations of the data are in keeping with the hortatory tone of the paper and are no doubt due to an attempt to provide practical advice for teachers. Whether this is acceptable is another matter.

Two of the interpretations which relate directly to the data deserve comment. First, it is not at all clear that either of the instructional strategies is superior to the other since the retention test means are almost identical. Second, it also is not clear that "using the inverse relationship...holds much promise..." as a teaching approach. Students missed nearly 50% of the posttest problems involving zero, the topic of the lesson. This score is very discouraging since some students must have been using the popular rule: "When zero is involved the answer is either 'zero' or 'no solution'. I can never remember which so I guess." This rule of chance will yield nearly the same 50% correct response mean found by the investigators after treatment.

Finally, one other misleading feature of the study stems from its early and cavalier dismissal of non-inverse explanations.

> It is only through this inverse relationship that students can meaningfully comprehend that sentences such as $12 \div 0$ have no solution. .... An attempt to interpret division by zero in terms of concrete objects, repeated subtraction, and other such techniques inevitably results in confusing situations and student failure.

There are at least five such different explanations of division involving zero. For example, see Robert Wirtz's use of toothpicks in the "Lines and Crossing Points" development presented in his second-grade text, Individualized Computation b_2, especially pp. 41, 45, and 51, Curriculum Development Associates, 1974. It may be that these other explanations are successful or it may be that they are not successful. There are no data. But there are data from this study suggesting that the inverse of multiplication approach is only moderately successful. It would seem that these other methods deserve more consideration.

J. Dan Knifong
University of Maryland

ON LEARNING "LESS": LANGUAGE AND COGNITIVE DEVELOPMENT. Holland, V. Melissa; Palermo, David S. Child Development, v46, pp437-443, June 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by William Nibbelink, The University of Iowa.

1. Purpose

The primary purpose was to determine whether a training procedure would lead kindergarten children to discriminate "less" from "more". The secondary purpose was to determine whether learning to discriminate "less" from "more" is correlated with conservation performance.

2. Rationale

Several studies support the proposition that many three- to seven-year-olds treat "less" and "more" as synonyms. "Relatively little attention has been given to the relationship between this confusion and such cognitive developmental phenomena as conservation." Clearly, success with teaching children to discriminate "less" from "more" coupled with a simultaneous or immediately-following gain on conservation tasks would speak to the broader question of whether stage development can be hastened via instruction.

3. Research Design and Procedure

Eighty-four kindergarten children were pretested for knowledge of "less" to secure a sample of 32 males and 30 females who exhibited the "more-less" confusion. The males and females were randomly assigned to two groups. While one group received no instruction in using "more" and "less", the other group received a sequence of tasks and/or instructions which was basically as follows:

(a) "More" was associated with "a lot" and "less" was associated with "a little bit" relative to examining two bowls with unequal numbers of marbles.

(b) S was required to say "more" or "less" when one of two such bowls of marbles was identified by the experimenter (E).

(c) S was required to respond to an imaginary inequality upon being asked questions such as "If this glass had more water, what would the other have?"

(d) Steps 1 through 3 were repeated using pictorial stimuli, followed by questions involving the functional implications of the "more" - "less" relationship relative to bad-tasting medicine, ice cream, etc.

(e) 3-D stimuli in non-static comparisons were presented, using addition, subtraction, and transfer with both discrete and continuous quantities.

(f) S was required to defend verbally the assignment of "more" or "less" in the face of counter-suggestions and illogical proposals by E.

(g) Using three quantities, S was required to use "more" and "less" in labeling the middle quantity relative to the extremes.

All children were pretested and posttested for both knowledge of "less" and ability to conserve, with all tests involving both discrete and continuous quantities.

4. Findings

The table below shows means and standard deviations for pretests and posttests. The "less" test had a possible 8 points. The conservation test was Form A or Form B of the Goldschmid-Bentler "Concept Assessment Kit: Conservation."

Means and SDs for Correct Responses

	Pretest		Posttest	
	M	SD	M	SD
"Less"				
Trained	0.90	1.15	7.42	1.48
Control	0.84	1.28	1.29	1.46
Conservation				
Trained	1.77	3.21	1.87	3.13
Control	1.87	2.75	1.35	3.20

No systematic variations were observed in either group relative to the conservation tests. Children who were trained learned quickly, showed confidence, were able to deal with more difficult tasks without further training, and showed no surprise upon acquiring the "more" - "less" distinction nor an awareness of previous incorrect usage of "less".

5. Interpretations

Since instruction can correct "more" - "less" confusion without improving conservation performance, and since 19 subjects scored above 0 on conservation while scoring 0 on the "less" test, the capacities active in distinguishing "less" from "more" are not sufficient for operational progress and apparently not necessary. Also suggested is that "language is not the source of logic but is structured by logic."

Since the first task seemed to be the crucial heuristic for the children's surprisingly quick learning, the "more" - "less" confusion may be a relatively superficial problem, perhaps nothing more than response bias which chooses the larger quantity as the more desirable regardless of the verbal instructions. If such is the case, then what the child knows about "less" remains unclear.

Critical Commentary

The study is well-designed, offering conclusions solidly supported by the findings. The suggestion that the "more" - "less" confusion may be something superficial like the response bias favoring the larger quantity is interesting. Relative to that question, using "short" - "tall", "skinny" - "fat", et cetera, would probably hold more promise for the following reasons: (1) "More" is often used in what the authors call "the additive, nonrelational sense" as simply a demand for continuous favor. The word "less" is not often used by children or by adults dealing with children. (2) Idiomatic usage may be, more or less, another confounding factor regarding the "more" - "less" distinction.

William Nibbelink
The University of Iowa

AN INVESTIGATION OF MATRIX TASK CLASSIFICATORY AND SERIATION ABILITIES. Hooper, Frank H.; Sipple, Thomas S. Wisconsin Research and Development Center for Cognitive Learning Technical Report No. 348, September 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by William E. Geeslin, University of New Hampshire.

1. Purpose

The purpose of this study was to determine if earlier results concerning cross classification and double seriation matrix tasks presented to children were replicable and to examine the relationship of these measures to a cross classification task which utilized the double seriation task stimulus materials. It was hypothesized that: (1) matrix tasks and age-grade level would be positively related; and (2) the following orders of task difficulty would occur: (a) transposition > reproduction, (b) cross classification II > cross classification I > double seriation for the reproduction case, and (c) cross classification II > cross classification I = double seriation for the transposition case.

2. Rationale

The ability to deal with multiplicative classes and relations is a common behavioral indicator of concrete operational thought as proposed by Piaget. Piaget's theories claim developmental synchrony for performances on task settings derived from the classificatory and relational "groupements." Both Inhelder and Piaget (1964) and Hooper et al. (1974) indicate that children reach an operational level in the multiplication of series about the same period as cross classification with respect to the transposition case. Although format and instructional variations appear to influence children's performances on matrix tasks, relatively few studies have examined the development of classification abilities with respect to relationality abilities. Previous studies, with some contradictory results, have found that ability to classify a simple arrangement precedes success with multiple classification matrices and development of simple serial skills. McKay et al. (1970) concluded that discrete category matrix tasks can be performed successfully prior to tasks involving relational categories; matrices with discrete categories in both directions are equivalent in difficulty to matrices with a discrete category in one direction and a relational variable in the other direction; with discrete categories, reproduction and transposition are equally difficult; and continuous categories in a matrix cause transposition to be more difficult than reproduction. Finally, the Hooper et al. (1974) study which immediately preceded the present investigation found age level was related to success on all tasks (older performing better). Matrix reproduction cases were: cross classification > double seriation > class/series. Likewise, for the transposition cases: cross classification = double seriation > class/series.

3. Research Design and Procedure

Forty subjects from each of four levels (preschool, kindergarten, first grade, and second grade) were selected for this study. With the exception of a 19/21 split in grade one, there were 20 males and 20 females at each level. Mean ages for the four levels, respectively, were 4-5, 6-1, 7-5, and 8-3.

Each matrix task made use of a square board sectioned into a 3x3 array of individual squares. The cross classification I matrix had nine blocks arranged on color and shape dimensions. The double seriation matrix had nine blocks arranged serially on width and height dimensions. The cross classification II matrix used the same blocks as the double seriation matrix. Although blocks were arrayed so that each row had a constant height and each column had a constant width, the blocks were not seriated on either dimension.

Matrix Tasks were presented in one of six possible orders. Within each matrix type, subtasks involving replacement, reproduction, and transposition were presented in that order. For the replacement subtask, the experimenter removed one, two, and three blocks and asked the subject to put them back where they were before. Reproduction required subjects to put blocks back in the original array after the experimenter had removed all blocks. The transposition subtask involved the experimenter removing all the blocks and then replacing one block in a position equivalent to a 90° rotation. The subjects were asked to place the blocks so they made a pattern like they did before.

Successful completion of the cross classification matrix reproduction subtasks was considered to be replacing the blocks with one dimension classified in each direction. Seriating a dimension in each direction was the criterion for passing the double seriation matrix reproduction subtask. Success on the transposition subtasks was indicated by fulfillment of the same criterion as noted above without moving the block replaced by the experimenter.

Chi-square tests and McNemar Tests for the significance of changes were used in making grade level and task difficulty comparisons, respectively. The general design of the study was grade level (preschool, kindergarten, first, second) x matrix type (cross classification I, double seriation, cross classification II) x matrix subtask (replacement, reproduction, transposition).

4. Findings

Order of presentation of matrix types did not affect subject success significantly with one exception. Higher numbers of subjects passed the cross classification II involution subtask in orders where the cross classification II matrix was presented last. No significant sex differences were found. There was a positive relationship between grade level and the number of subjects who were successful on the various matrix tasks. With the exception of the cross classification I transposition case, chi-square comparisons of passing frequencies across the four grade levels were significant. Reproduction tasks were consistently less difficult than transposition tasks. For the reproduction tasks, the order of matrix difficulty was: cross classification II > cross classification I = double seriation. Similarly the difficulties of the three matrix transposition tasks were: cross classification II > cross classification I = double seriation. Color dimension errors were the predominant misplacements for the cross classification I task. For the cross classification II and double seriation matrices, the tendency was to err on the height dimension.

5. Interpretations

The patterns of numbers of successful subjects on the types of matrices were in accord with earlier results of MacKay (1970) and Hooper et al. (1974). The present subjects' placement errors were also in agreement with Hooper et al. (1974) but contradictory to MacKay's (1970) results. The noted growth of multiplicative logical skills during this age range is consistent with previous normative data. However, these subjects did less well than the previous Hooper et al. (1974) subjects.

The fact that transposition tasks were more difficult than reproduction tasks replicates previous findings. Most children who succeeded on the reproduction tasks produced an identical array. With the exception that no differences in difficulty were found between cross classification I and double seriation for the reproduction task, the difficulty patterns were consistent with Hooper et al. (1974) and in direct contradiction to Lagatutta (1970, 1974). The present results appear to agree with the work of Inhelder and Piaget (1964) who contended that "children reach an operational level in the multiplication of series about the same period (7-8) as cross-classification." Inhelder and Piaget remarked that in spite of perceptual differences all four principal groupings in the logic of classes and relations became operational during the same period. However, further examination of the present data does not support this conclusion. The cross classification II tasks were notably more difficult. Yet the stimulus materials for cross classification II and double seriation were identical and only the initial arrangement and instructional set varied between the two matrix tasks. The large majority of children who passed one of these tasks and failed the other did so by succeeding on the double seriation task and failing the cross classification II task. Other investigations are consistent with these results indicating relational understanding is a developmentally earlier acquisition than conservation or class inclusion (cf. Wang, 1975).

Critical Commentary

It should be noted that the paper reviewed here is a technical report rather than a journal article. A technical report ordinarily includes more information, but is written in a less "polished" form than a corresponding journal article. Hooper and Sipple conducted an interesting study and wrote a rather good technical report. Nonetheless, certain features of the report cause difficulties for an "outside" reader.

Hooper and Sipple apparently assumed that the reader was familiar both with technical terms such as double seriation and the connection of these terms to classificatory and relational abilities in children. The terms were not always consistent. For example, "involution" task appears on page 15 of the result section apparently in place of transposition. The literature review was extensive but the discussion of this literature was too brief. Both the introduction and discussion sections failed to emphasize the significance of this work with respect to the school and curricula. It might be interesting to look at the results by age in addition to the grade-level means, particularly in the section on ability development (across grade-level comparison). Minor difficulties were: some tables were difficult to read/compare; the reported significance levels varied; the replacement subtasks were never mentioned in the results and discussion sections; the sampling procedures were noted partially in a separate technical report; and more information concerning the length of time a student could use in completing a task and interaction between experimenter and subject could have been given.

It should be emphasized that the above criticisms are "unfair" to a degree since this was a technical report rather than a formal article. The research reported may lead to work that could not only cause revision in a major psychological theory, namely Piaget's, but also suggest significant changes in mathematics curricula for young children. Hooper and Sipple's work in conjunction with the literature noted by them deserve careful consideration by the mathematics education community.

William E. Geeslin
University of New Hampshire

PEERS, PERSUASION, AND PIAGET: DYADIC INTERACTION BETWEEN CONSERVERS AND NONCONSERVERS. Miller, Scott A.; Brownell, Celia A. Child Development, v46, pp992-997, December 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Thomas P. Carpenter, The University of Wisconsin - Madison.

1. Purpose

Two issues provide the focus for this study. The first concerns the certainty with which children maintain conservation judgments. The second is the role of peer interaction in the development of conservation.

2. Rationale

Although Piaget's theory and common sense suggest that conservation judgments should be held with more certainty than nonconservation judgments, previous research indicates that this may not be the case. Contrary to the earlier results of Smedslund (1961), Miller (1971, 1973) has found that conservers are not particularly resistant to the extinction of their beliefs. However, the authors of the current study conclude that there are certain logical and methodological limitations to this research that warrant further investigation of the problem.

Although Piaget attributes a central role to peer interaction in cognitive development, this variable has been largely ignored in the research literature. Only three studies were identified that paired conservers with nonconservers, and these studies failed to control for general differences in social influence. Furthermore, none of these reports contain much information about the interaction between the children.

3. Research Design and Procedure

The procedure used to investigate these issues was to examine carefully the interaction between pairs of subjects consisting of one conserver and one nonconserver who had been required to discuss a conservation problem and arrive at a common solution.

The study was divided into two parts. In the first session the pretest, which included two typical conservation-of-length tasks and two typical conservation-of-weight tasks, was administered to an initial sample of 223 second graders from a predominately middle-class school. A child was classified as a conserver for a given concept if he or she answered both questions correctly and also gave correct logical explanations for these answers. He or she was classified as a nonconserver if both questions were answered incorrectly. Any other pattern was considered transitional. Children were also asked to identify the most dangerous animal in the world and the best show on television. These questions were used to control for differences in social influences.

On the basis of the pretest a final sample of 100 subjects (58 boys and 42 girls) was identified. The sample was divided into 50 pairs, with each pair consisting of one conserver and one nonconserver.

During the second session, which occurred from one to eight days after the pretest session, 14 of the pairs were tested on length, 17 on weight, and 19 on both concepts. In all three cases only the first question used in the pretest for a given concept was administered. All pairs were asked the questions about the most dangerous animal and the best television show. For each question the children in each pair were required to discuss the question with each other and agree on a common answer. At the conclusion of the session, the child who had changed his or her conservation answers was retested individually on the conservation concepts for which he or she had changed the response. Both pretest questions were readministered for this posttest.

Two judges independently rated the adequacy of the conservation explanations and classified the interaction into categories: who uttered the first relevant statement, who offered an explanation of his answer, et cetera. The interrater agreement on 12 such categories ranged from 77 percent to 99 percent, and there was 96 percent agreement on the adequacy of conservation explanations.

4. Findings

Conservation of weight was found to be significantly easier than conservation of length ($p < .001$) on the pretest. Based on the control questions administered during the interaction session, there was no evidence that conservers and nonconservers differed in social influence. However, on both length and weight questions, conservers' judgments prevailed significantly more often than nonconservers ($p < .001$). Significant differences in interaction patterns emerged on several measures. For length, conservers were more likely than nonconservers to explain answers, counter the other person, or suggest moving the stimuli.

The same results were found when subjects were partitioned on the basis of winners and losers, which is not surprising since conservers were usually winners and nonconservers losers. However, on the weight trials, conservers and nonconservers only differed on the basis of who asserted their answer from the pretest while winners were more likely to utter the first relevant statement, give the first explanation, assert the answer, explain their answer, and counter the other.

On the posttest, 73 percent of the nonconservers conserved on the first problem, which had been used in the interaction session; 31 percent conserved on the second problem, which had not. Ninety-three percent of the judgments were accompanied by adequate explanations. Since nonconservers prevailed on only eight trials, posttest data for conservers are limited. Six children gave nonconservation responses on the first problem and one on the second.

5. Interpretations

In contrast to data from extinction and surprise studies, the interaction paradigm used in this study indicates that a belief in conservation is more firmly held than a belief in nonconservation. The qualities that seemed to distinguish the conservers were the variety and adaptability of their approach.

The degree of genuine training or extinction was modest at best, and it is doubtful that it resembles the way in which children acquire conservation outside the laboratory. One implication is that research should focus less on posttest gain and more on identifying the potentially generalizable aspects of the process by which change comes about.

Critical Commentary

The whole issue of relative certainty of the judgments held by conservers and nonconservers is built on somewhat shaky ground. The authors compare their findings to the results from extinction studies, with the implication that they are focusing on the same general question. Extinction studies, however, are not generally concerned with the relative certainty of conservation and nonconservation beliefs. They are concerned with the relative certainty with which natural conservers and trained conservers hold their beliefs. The point of these studies is to determine whether trained conservers have attained the genuine operational level of natural conservers, which is very different from the issue identified in the above study.

Furthermore, part of the justification of the study is based on the fact that Piaget's theory indicates that conservation beliefs should be held with more certainty than nonconservation beliefs. However, Piaget describes nonconservers in Stage I as being extremely resistant to any suggestion that there may be some inconsistency in their judgments. Given the age of the subjects in the study, it is likely that a more extensive and probing pretest would have identified many of the nonconservers as being in transitional stages. Somewhat different interaction patterns might have been found if younger subjects were included and greater care were taken to identify nonconservers.

With respect to the effects of peer interaction, it is doubtful that any significant cognitive growth could result from such short interaction (training) sessions. Since the posttest immediately followed the instructional treatment, there is no evidence of any lasting effect. Furthermore, since the posttest consisted of the problem used in the interaction session and one similar problem that differed only in the type of transformation employed, there is little evidence that subjects were doing any more than parroting the response that they had been induced to agree to in the interaction session. This explanation is supported by the fact that 73 percent of the subjects gave the correct response to the problem used in the interaction sessions, while only 31 percent correctly answered the second problem even though the problems were substantially the same. In addition, since no attempt was made to control for the effect of the pretest, it is not clear whether the gains resulted strictly from the interaction session. This problem is especially acute given the fact that the pretest included both problems on the posttest, while the interaction session consisted of a single problem. Thus, although the effect of peer interaction is an issue of some significance for mathematics education, this study provides little support one way or the other.

On the whole it is doubtful that the results of this study have any serious implications for teaching or research in the learning of mathematics. The authors themselves note that it is questionable whether the peer interaction described in the study "has much direct resemblance to the way in which children acquire conservation outside the laboratory." However, the authors' attention to the specific process of interaction within the treatment provides a paradigm worth considering, and their recommendation that research "focus less exclusively on posttest gain and more on identifying the potentially generalizable aspects of the process by which change comes about" is a good one.

Thomas P. Carpenter
The University of Wisconsin - Madison

References

Miller, S. A. The extinction of conservation: A methodological and theoretical analysis. Me[illegible] [illegible]almer Quarterly, v17, pp319-334, 1971.

Miller, S. A. Contradiction, surprise, and cognitive change: The effects of disconfirmation of belief on conservers and nonconservers. Journal of Experimental Child Psychology, v15, pp47-62, 1973.

Smedslund, J. The acquisition of conservation of substance and weight in children: III. Extinction of conservation of weight acquired "normally and by means empirical controls on a balance." Scandinavian Journal of Psychology, v2, pp85-87, 1961.

THE DIFFERENCE IN LEVEL OF ANXIETY IN UNDERGRADUATE MATHEMATICS AND NON-MATHEMATICS MAJORS. Ohlson, E. LaMonte; Godwin, Wanda. Journal for Research in Mathematics Education, v8 n1, pp48-56, 1977.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Suzanne K. Damarin, The Ohio State University.

1. Purpose

To determine whether undergraduate students majoring in mathematics differ from nonmathematics majors in level of anxiety, and to examine relationships of anxiety to certain status variables (sex, academic level), aptitude, and achievement.

2. Rationale

Although the effects of anxiety on learning and academic performance have been subject to numerous studies over a period of almost 50 years, very little research relating anxiety to specific academic disciplines has been done. Especially in the area of mathematics little is known about the effects of anxiety. Existing literature does suggest that relationships among aptitude, achievement, and anxiety do exist.

3. Research Design and Procedure

Sixty seven mathematics majors and 124 nonmathematics majors were randomly selected from the undergraduate student body of a university. Each student was asked to complete both subscales (A-State and A-Trait) of the State Trait Anxiety Inventory (STAI). Approximately half of the students from each group were given the questionnaire in a mathematics classroom; the remainder were tested in another (unspecified) classroom. Data on each of the following variables were obtained for each student: sex, academic major, academic level, cumulative grade point average, previous quarter grade point average, ACT: mathematics score, ACT: composite score, and a composite academic level by academic major score. These variables together with the testing condition (mathematics vs. nonmathematics classroom) and trait anxiety scores were treated as independent variables. Regression analyses were used to predict state anxiety from various combinations of these variables.

Factor analysis was used to determine five clusters of variables (mathematics major or classroom, achievement, aptitude, academic level, and general (trait) anxiety). Regression analysis was performed using the full set of variables; additional analyses were performed on data sets from which scores on each of the five clusters were deleted. The squared multiple r (RSQ) from the full analysis was compared with the RSQs obtained from each of the restricted analyses to determine the unique contribution of each cluster of variables to state anxiety. In each comparison a difference of .05 between the RSQs for the full and restricted data sets was considered to be significant.

4. Findings

The squared multiple r for the full set of predictor variables was .4692. When the general (trait) anxiety score was removed from the set of independent variables the RSQ was only .0567. When each of the other variable clusters was removed the RSQ remained between .4575 and .4688.

5. Interpretations

None of the independent variables except trait anxiety contributed significantly to the state anxiety of the subjects tested. In particular, mathematics majors were not found to be more anxious than nonmathematics majors; nor were students tested in mathematics classrooms more anxious than students tested in nonmathematics classrooms. Failure to find differences in state anxiety attributable to the major of the student or to class in which testing takes place may reflect the students' success in mathematics courses, the attitudes of their teachers, or the absence of stressful conditions (e.g., imminent evaluation).

Abstractor's Notes

The description of the selection of subjects and administration of the STAI is extremely meager. It is difficult for this reader to imagine how the study was conducted. If subjects were indeed randomly selected from the university population, how was in-class testing arranged? What mathematics classes were used? What nonmathematics classes? Were tests administered during regular class sessions? What instructions were given to subjects? Numerous other questions might be asked about the population and the testing conditions. When the dependent variable in a study is State Anxiety, the reader needs and deserves more information about the states under which tests were administered than is provided in this paper.

Trait anxiety accounts for 45% of the variance in state anxiety in this study. Taken together the full set of independent variables (including trait anxiety) account for 47% of the variance. This figure is far lower than one would hope for in a study involving 10 predictor variables and 191 subjects. What is the source of the remaining 53% of the variance? Is there an underlying variable (or set of variables) which was not examined in this study? Or was there an inconsistency in testing conditions which contributed heavily to the variance?

Suzanne K. Damarin
The Ohio State University

VERBAL-NONVERBAL CONSERVATION AND PRIMARY MATHEMATICS. Riggs, F. T.; Nelson, L. D. Journal for Research in Mathematics Education, v7 n5, pp315-320, November 1976.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Joe Dan Austin, Emory University.

1. Purpose

This study investigates the relation of verbal and nonverbal measures of length conservation to achievement, intelligence, age, sex, and socioeconomic status (SES).

2. Rationale

The researchers cite the conflicting evidence on when children conserve length. Studies by Braine, Elkind, Piaget, and Smedslund yield a range of 4 to 8 years for the threshold age of length conservation. Although Olson attributes this wide range to cultural differences, the authors argue it may be due to differences caused by using both verbal and nonverbal conservation tests. If this is true then there is a need to investigate these two types of conservation tests.

3. Research Design and Procedure

The researchers randomly divided 100 first-grade children of similar SES into two groups of fifty students each. Students in one group were individually given a 24-item length conservation test in a nonverbal manner. This test was developed by Sawada as a non-verbal test and revised for this study. Students in the other group were individually given the same test but in a verbal manner approximating Piaget's type of interview. Children with 17 or more correct responses were termed conservers and children with fewer than 15 correct responses were termed nonconservers.

A correlation matrix was computed for the variables: achievement, nonverbal conservation, verbal conservation, IQ, age, sex, and SES. The classification conserver-nonconserver was also correlated with these variables.

Kuder-Richardson Formula 20 estimates of reliability for the three measures constructed or modified by the researchers were: nonverbal conservation 0.63, verbal conservation 0.79, and mathematics achievement (25 items) 0.79.

4. Findings

All statistical tests were made at the 5% level.

The two groups did not differ on mean achievement scores but did differ on mean conservation scores. In the nonverbal test group 45 of the 50 children were labeled conservers and in the verbal test group only 23 of the 50 children were labeled conservers. These classifications were significantly different using a Kolmogorov-Smirnov test.

The correlation matrix yielded significant correlations for (a) achievement and the variables nonverbal conservation (0.316), verbal conservation (0.623), and intelligence (0.560); and (b) verbal conservation and the variables intelligence (0.555) and sex (0.312). Correlating with the classification conserver-nonconserver yielded significant correlations for the verbal group with the variables achievement (0.564), verbal conservation (0.814), and sex (0.325), and the nonverbal group had a significant correlation with the classification and nonverbal conservation (0.661). Since no children took both forms of the test, no correlations could be computed between verbal and nonverbal conservation measures.

Thus the verbal conservation measure was a better one-variable (linear) predictor of achievement than either the nonverbal measure or IQ. The partial correlation of verbal conservation score and achievement controlling for IQ was significant. Replacing the verbal conservation score by the nonverbal score did not yield a significant partial correlation. In the correlation between sex and verbal conservation the girls were superior to the boys. This was also true for the verbal group conserver-nonconserver classification correlation with sex.

5. Interpretations

The following conclusions were drawn from the study:

(a) Many children who were classified nonconservers on the verbal test may have been classified as conservers on a nonverbal form of the same test.

(b) A verbal factor permeated all the findings as subjects equated "higher" with "taller" and "larger" with "longer".

(c) A nonverbal setting enabled more children to exhibit their competence on conservation tasks than did the verbal setting.

(d) It is unwise for teachers to judge young children's competencies solely on the basis of their performance through verbalization.

(e) Both verbal and nonverbal activities should be used in the early stages of primary grades to aid in developing mathematical concepts and language improvement.

Critical Commentary

This study considers an important question in primary education--the relation of conservation measures to achievement. If measures of length conservation do not have a significant relation to achievement, it is doubtful they would relate to a child's learning of topics requiring length and measurement-related topics.

Several issues need to be considered before the validity of the results can be considered. This reviewer had some difficulty understanding the differences between the verbal and nonverbal methods of administering the conservation test. Although the authors do have a section relating to this problem, in both ways of administering the conservation test the student was asked "Are they (the rods) the same length?" The referenced paper by Sawada and Nelson (1967) seems to give the better explanation that nonverbal refers only to the student response. Thus the student still has to respond to a verbal question. Another question deserving some attention refers to the criteria for labeling students as conservers or nonconservers. Why were the particular numbers 15 and 17 chosen? The analysis of the number of conservers in the two groups depends of course on the definition used. Finally, if the nonverbal form of the conservation test is superior to the verbal form, and it may be superior, why does it have a lower predicted reliability than does the verbal form of the same test?

Several minor questions also arise. Some of these are: What IQ test was used? What SES measure was used? What statistical test was used to compare group means? When during the first-grade school year was the testing done?

Given that the above problems can be addressed satisfactorily, and this reviewer feels this can be done, the findings appear to be valid. The authors seemed to be careful and uniform when administering the conservation tests to both groups; therefore, the results would likely be similar with similar groups of children. The actual results are encouraging for the usefulness of conservation measures. The fact that verbal conservation had a higher correlation with achievement than did IQ is interesting. This may depend, of course, on the particular IQ and achievement tests used but the result is encouraging. Several important followup questions arise. How does the conservation measure relate to learning and how stable is it during instruction? Girls and boys seem not to differ on the achievement measure used here; does the difference on the two conservation measures (girls superior to boys on verbal test) give a way to predict or explain future achievement patterns?

Finally two of the conclusions made by the authors, while probably true, do not seem to follow entirely from the analysis and data reported in the study: conclusions (b) and (e). Conclusion (b) is a probable explanation for much of the difficulty with conservation of length. However, there are no data or tests relating directly to this conclusion. Conclusion (e), while again very likely true, does not seem to be based on the experimental evidence cited. The study considers measurement achievement and not instructional procedures. These two conclusions would seem to be very reasonable recommendations to make. Although this may be a fine line to draw between conclusions and recommendations, they do not seem to be based entirely on this study or the literature cited in the development for the study.

Joe Dan Austin
Emory University

CONCRETE INSTRUCTION IN ELEMENTARY SCHOOL MATHEMATICS: Pictorial vs. Manipulative. Scott, Lloyd F.; Neufeld, Herman. School Science and Mathematics, v76, pp68-72, January 1976.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by J. Larry Martin, Missouri Southern State College.

1. Purpose

To answer the question, "Does the use of concrete manipulative materials contribute more to children's mathematical concept formation than does the use of pictures?"

2. Rationale

While there is increasing support for "concrete" instruction in mathematics education, there is no standard definition of "concrete." Concrete approaches range from the use of manipulative materials to computer assisted instruction. Research results favor concreteness but pay little attention to the nature of concreteness. This study focuses on two common instructional modes, pictorial and manipulative. These modes have been differentiated in some instruction sequences. For these sequences, manipulative materials are used in early stages of the sequence and considered concrete. Pictorial representations follow and are considered semi-concrete. Other writers have given special attention to the value of pictures in concept formation. The current study investigates whether the two approaches, manipulative and pictorial, are equally concrete, or at least equally advantageous to children's conceptual growth in mathematics.

3. Research Design and Procedure

Three second-grade classes in each of three West Coast elementary schools participated in the study. The schools were selected on the basis of their representativeness, teacher and principal willingness to participate, availability of three second grades in the school, and rated competence of potential participating teachers. Within each school classrooms were randomly assigned to use one of three modes of instruction, (concrete manipulative (M), pictorial (P), or abstract (A)) to teach beginning multiplication. The A mode was used as a control for the M and P modes.

Curriculum materials were specially prepared to vary only on the variables indicated by their labels. That is, mode M children consistently used objects in their study. They never used pictorial materials and used abstract representations "only after concepts had been assimilated." Similarly, P children used only pictures and A children used no physical or pictorial referents. Instruction was spread over one month, with 20 sessions of 25 minutes each. Participating principals and teachers were given advance orientation to the procedures; teachers were encouraged to remain faithful to their assigned mode of instruction and principals were urged to monitor their participating teachers. During the experimental period, the investigators continued to supervise instruction and assist with problems as they arose.

An investigator-prepared test was administered before and after the instructional period. The instrument is described as being modeled after Diagnostic Tests to Accompany Modern School Mathematics Structure and Use 3 (Duncan, 1967). Test scores for the two groups M and P were compared using analysis of covariance with I.Q. and pre-test scores as covariates. In addition, the investigators interviewed 15 students from each group and used questionnaires completed by six participating teachers to determine affective consequences of the three modes of instruction. Post hoc comparisons were made between the P and A groups and between the M and A groups.

4. Findings

No significant difference was found between the pictorial (P) and manipulative (M) modes. The post hoc comparisons revealed no significant differences between P and A or between M and A. Analysis of the data from the student interviews and teacher questionnaires indicated positive affective responses favoring M and P over A for both the students and teachers, and P over M for the students. No clear distinction in affective response was found between the teachers in the M and P groups.

5. Interpretations

At least for the acquisition of beginning multiplication concepts, pictures seem to be as effective as manipulative materials. Both modes appear equally concrete. The investigators did not report any inferences or interpretations from their interview - questionnaire findings.

Critical Commentary

As the authors point out, the test used is susceptible to the criticism that it is principally abstract and hence does not measure understanding. It is possible that the test measured only ability to recognize and manipulate symbols. If the abstract nature of the test was recognized by the investigators at the outset, a reasonable question would be, "Could there have been a test given which would have made an attempt to measure understanding as well as manipulation of symbols?"

One of the claimed advantages of "concrete" instruction is that it aids accommodation, as compared to mere assimilation, and hence promotes retention. A follow-up evaluation would be valuable in comparing the merits of the three instructional approaches. What are the effects on retention of the pictorial and the manipulative lessons?

There are several questions related to the interviews and questionnaires. How were the student participants selected? What were the questions asked? How were responses evaluated? How were "clear differences" defined?

Two other questions are: (1) How large were the groups? and (2) Why not compare all three modes of instruction at once, rather than comparing two at a time?

The study is an interesting one and focuses on a major problem in mathematics education, namely ill-defined terms. Bandwagon phrases do little to advance the state of the art. Mathematics educators have recently begun questioning the role that manipulative materials play in instruction. The nature of the specific manipulatives could be important. What distractors are inherent in the manipulatives? What underlying mathematical principles do the manipulatives utilize? For example, many manipulatives presuppose spatial concepts. Does the child have these concepts? How are the aids being used to develop or to demonstrate? Is the child being asked to "read out" from the aids principles or concepts which are not inherently there but which have been "read in" by the instructor? More descriptive information concerning the nature of the manipulatives and their use in this study would be helpful.

J. Larry Martin
Missouri Southern State College

RELATIONS AMONG PIAGETIAN GROUPING STRUCTURE: A TRAINING STUDY. Silver, Edward. Journal for Research in Mathematics Education, v7, pp308-315, December 1976.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Billie Earl Sparks, University of Wisconsin-Eau Claire.

1. Purpose

The purpose of this study was to investigate the relationship of the various grouping structures as described by Piaget.

2. Rationale

The logico-mathematical structures which Piaget called groupings are structures that correspond to the organization of children's cognitive and behavioral actions. A grouping usually possesses closure, completeness, and reversibility but not all properties of a mathematical group. This study utilized training on behaviors in one grouping (primary addition of classes); on behaviors in the same grouping (class inclusion); and on behaviors in other groupings such as number and substance conservation. This was undertaken to provide additional information about Piaget's theory and because of the importance of these grouping structures in the mathematical knowledge of the child.

3. Research Design and Procedure

Fifty-five lower- and lower-middle-class first-grade children from a parochial school in New York City were the subjects used in the study. Ages of the subjects ranged from 6 years, 3 months to 7 years, 6 months.

All subjects were pretested on number conservation, substance conservation, class inclusion, and transitivity. The experimental subjects were then exposed to a two-session training process with a 2- or 3-day gap between the sessions. The training consisted of work with attribute blocks varying by size, shape, and color. The sessions consisted of the subjects identifying attributes possessed by the objects, naming classes to which they could belong and classes to which they could not belong, and distinguishing the objects by their attributes. The control group worked with the attribute blocks by building shapes but their attributes were never discussed. There were two posttests. One was given within 3 days of the training and the second was given 3 weeks later.

The subjects were classified as operative or non-operative in each of the four areas on the pretest. This categorization was done again for each posttest. The McNemar chi-square test for significance of changes was used to test the data. A sign test was also conducted on number of subjects improving or regressing. A Wilcoxon Matched Pairs Test was conducted on pairs of experimental and control subjects matched by pretest scores.

4. Findings

The changes from operative to non-operative for the trained group were found to be significant for class inclusion (McNemar $X^2 = 6.1$, $p < .01$) and for number conservation (McNemar $X^2 = 3.2$, $p < .05$).

Twelve subjects in the training group improved and none regressed on class inclusion tasks (sign test, $p < .001$). On number conservation nine improved and none regressed (sign test, $p < .01$).

The training group subjects scored significantly higher than their control paired mates on both class inclusion and number conservation (Wilcoxon, $p < .005$).

There were no significant findings on substance conservation or transitivity.

Due to virtually identical scores on the two posttests no additional analysis was conducted for the second posttest.

5. Interpretations

The author states that the data support the hypothesis that training on certain behaviors (additivity of classes) in a grouping strengthens other behaviors (class inclusion) in the same grouping. However, the findings for behaviors in other groupings were equivocal. Number conservation scores were improved but not substance conservation or transitivity scores. The author states that the lack of influence of increased number conservation on substance conservation leaves open the question of the strength of the relationships within a single grouping.

The lack of increase in transitivity scores may have been due to the high number of correct responses on the pretest; little room was left for improvement.

The author indicates the data show the groupings to interrelate in a very complex way although they are easily distinguishable from each other.

Critical Commentary

The study apparently was done in a sound, "clean" manner and the results are consistent with many other findings in relation to working on certain logical structures to induce such concepts as number conservation.

The study does add theoretical knowledge in showing some interrelationships between and within grouping structures. The study indicates these groupings are not hierarchical in nature but are complexly related. This complex interrelationship could probably best be explained through an analysis of those abilities necessary to complete each grouping to the point of providing effective adaptation. These abilities include but are not limited to reversibility.

The study also hopes to add to practical knowledge because of the importance of these structures in acquiring mathematical knowledge. For most studies on the induction of conservation skills, the practical question is: Why bother? The developmental process will take care of this. There is also the question of the "quality" of the induced conservation. Even though several additional trained children were judged number conservers, will they still break their cookies into several pieces so they will have more to eat?

To the abstractor the primary practical importance of the study is with the particular materials and approaches used. Every kindergarten teacher should be able to defend a mathematics curriculum which does not start with number skills but instead with prenumber activities on classification and seriation to help build a stable concept of number. The attribute materials and questioning procedures used in this study should certainly be of value in this area. Since several studies have shown that induction of conservation and other skills is possible, perhaps future work should center on the development and testing of number readiness tests and materials and finally move Piaget into the classroom.

Billie Earl Sparks
University of Wisconsin-Eau Claire

THE EFFECTS OF THREE INSTRUCTIONAL STRATEGIES ON PROBLEM SOLVING BEHAVIORS IN SECONDARY SCHOOL MATHEMATICS. Vos, Kenneth. Journal for Research in Mathematics Education, v7, pp264-276, November 1976.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Stephen S. Willoughby, New York University.

1. Purpose

To investigate the effects of three instructional strategies (repetition, list, behavioral instruction) used to increase the effective use of five problem-solving behaviors (drawing diagrams, approximating, constructing equations, classifying data, constructing charts).

2. Rationale

Problem-solving ability is generally recognized as an important goal of mathematics education. Most research on the subject has concentrated on translating word problems into number sentences and on answers. This research was designed to study ways of affecting the methods of solution.

3. Research Design and Procedure

The independent variable was instructional strategy. The three strategies were:

(1) Repetition: Only problem tasks were given.

(2) List: The problem task was presented. A checklist of suggested procedures to follow was provided and children were told to check off all procedures they "tried or thought about trying." Some behavioral instruction was provided and they were instructed to return to the problem task.

(3) Behavioral Instruction: Behavioral instruction was provided and then the task was presented.

Five behaviors identified in a pilot study and taught in the treatment period were: drawing diagrams, approximating and checking, writing equations, classifying data, and constructing charts.

Instructional material was presented through self-directed, written material in the form of twenty problems (about 20 minutes each) used over a fifteen-week period. Six classes at three grade levels in a private school in Iowa were involved. The classes (with numbers of students) were: Algebra II (25); Geometry (29); Algebra I, section 1 (22); Algebra I, section 2 (28); Math Survey (21); and Elementary Algebra (8). The total N was 133. Low-ability students were assigned to the last two classes. The very small number of students in Elementary Algebra resulted from unexpected schedule conflicts. In each case, students were randomly assigned to one of the three treatments.

The dependent variables were:

(a) Scores on STEP forms 2A and 3A, Mathematics Part II (pretest); and Part I (posttest)

(b) Scores on a Problem-Solving Approach Test (PSAT) constructed by Vos (posttest). The students were not required to solve the problems, but rather were to choose from a list of five approaches the best and next-best approaches to each problem.

(c) Scores on a Problem-Solving Test (PST) apparently constructed by Vos (posttest). Students were to show their work and were scored both on a right-wrong basis and partial score basis.

Analysis of variance was used to analyze the data. For the PST there also was an analysis of the five instructed problem-solving behaviors and of other problem-solving behaviors. For the PSAT, a Newman-Keuls method was used to analyze the differences between all pairs of posttest means within each class.

4. Findings

There were no significant differences between treatment groups on the STEP. The author reports differences significant at the .20 level for three of the six classes on the PST. On a purely descriptive basis, the author reports that in the low-ability classes the list treatment had the highest proportion of occurrences of the taught behaviors, while in the other algebra classes the behavior instruction treatment had the highest proportion, and in the geometry class the repetition treatment had the highest proportion of occurrences.

Some differences on the PSAT were reported at the .05, .10, and .20 levels of significance generally favoring the behavior instruction and the list procedures over the repetition.

5. Interpretations

The author calls attention to the unusual nature of the tests he constructed and the instructional materials. He mentions the short time of treatment and the possibility that some differences reported were more a result of regular classroom instruction by the teacher than a result of the treatment. He also calls attention to the fact that students in more advanced classes tend to be better problem solvers.

Critical Commentary

Presumably, if we knew that certain strategies were more useful than others, knowing how to get students to use them would be important information. Therefore, the question asked in this study appears to be a reasonable question. However, forcing students to choose strategies from a multiple-choice list seems to be a doubtful procedure for determining what strategies they use. In fact, I used a strategy of counting on one of the test items shown and of drawing a picture on the other and in neither case was my strategy listed as a possibility. I suspect some students may have had a similar difficulty. A case study procedure would seem to have been more appropriate than the experimental procedures used, especially in light of the "fishing trip" nature of the entire study.

As to the statistical procedures used, given the large number of times analysis of variance was tried, the omission of post analysis of variance tests, the apparent complex lack of any hypotheses, and the willingness of the author to report levels of significance as great as .20, there appears to be a rather high probability that the reported results could have been produced by a set of random variables.

Stephen S. Willoughby
New York University

THE PHILOSOPHY AND APPLICATION OF AN AUDIOVISUAL LABORATORY IN TEACHING BASIC MATHEMATICS. Williams, Eddie R.; Mick, Harold W. The MATYC Journal, v10 n3, pp167-172, Fall 1976.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Sigrid Wagner, University of Georgia.

1. Purpose

Although there is no explicit problem statement in this article, the authors' "main contention . . . is that the possible instructional contribution of the audiovisual laboratory in basic mathematics is unique and significant" (p. 169). The results of a study "measuring the effectiveness of the audiovisual laboratory" are reported.

2. Rationale

Students in basic mathematics courses at the college level often suffer from negative attitudes toward mathematics and deficiencies in reading and writing skills. In order to achieve successfully, these students usually need continual practice, immediate reinforcement, and ample opportunity for review.

An audiovisual laboratory helps to satisfy these students' needs by providing variable pacing of instruction and frequent evaluation of progress. Therefore, an audiovisual mathematics laboratory should be considered for inclusion in any combination of instructional modes designed to maximize the effectiveness of instruction in basic mathematics.

3. Research Design and Procedure

A pretest-posttest control group design was used. Two groups of 30 students each were randomly selected from one basic mathematics lecture course. All students attended three hours of lecture per week. In addition, students in the experimental group were required to spend one hour per week in the audiovisual laboratory; students in the control group were required to attend one hour of recitation per week. Students in both groups were given pretests and posttests over five of the last eight topic modules of the course.

Within both the control group and the experimental group, students were classified into four levels of achievement according to their "prior course performance" (p. 170). Mean pretest and mean posttest scores on each of the five modules were computed for each achievement level within each group. The mean improvement score (posttest score minus pretest score) at each level within the experimental group was then compared with the mean improvement score of the corresponding level of the control group.

4. Findings

The mean pretest, posttest, and improvement scores are listed by achievement level and group for each of the five modules and for the "course." One example will suffice to illustrate the nature of the results: The "collective mean" course score for the experimental group improved from 11.64 to 83.63, a gain of 71.99, while the score for the control group improved from 11.62 to 81.07, a gain of 69.45 (mistakenly listed as 66.45, one of at least two errors in the table of results).

5. Interpretations

The authors state that the results are inconclusive, although the students at each level in the experimental group did "comparatively better in total points of cognitive change" (p. 170) than the students in the control group.

Critical Commentary

The authors must be given credit for acknowledging that the results of their study are inconclusive.

Their report of the study leaves several questions unanswered:

1) When were the students selected for the experimental and control groups? If selection occurred at the beginning of the term, why analyze data obtained for only "five of the last eight" (of how many?) modules in the course?

2) On what basis were the students in each group classified into the four achievement levels? The phrase "prior course performance" is ambiguous.

3) How many students were there at each achievement level? Computing the "collective mean" scores for each group of 30 students as the authors did, by taking the unweighted average of the mean scores at each achievement level, seems to imply that there were 7.5 students at each of the four levels. Such an occurrence is highly unlikely.

4) Presumably the mean pretest and mean posttest scores on each module are average raw scores, although there is no indication of the total possible score on any of the tests. What, then, are the "course" pretest and posttest scores? They are not the total scores on all five modules, although their superficial resemblance to the module scores leaves that impression.

No one can argue with the authors' contention that the contribution which an audiovisual laboratory makes to mathematics instruction is unique. Those who dispute the authors' contention that this contribution is also significant will not be any less convinced after reading this article.

Sigrid Wagner
University of Georgia

STUDY III -- THE EFFECT OF TEST ITEM ORDERING ON THE PERFORMANCE OF CHILDREN OF DIFFERENT TEST ANXIETY LEVELS IN FOUR STUDIES ON ASPECTS OF ASSESSING COMPUTATIONAL PERFORMANCE. Technical Report No. 297. Zeineh, Farid Abu; Pierce, Edward; Silverman, Mark. Report from the Project on Conditions of School Learning and Instructional Strategies. Wisconsin R & D Center for Cognitive Learning, July 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by James H. Vance, University of Victoria

1. Purpose

To investigate the performance of high, intermediate, and low test-anxious fifth-grade pupils on three different orderings of item difficulty (ascending, descending, and mixed) of an arithmetic computation examination.

2. Rationale

Considerable research has been conducted to investigate the effect of test anxiety on pupil performance on arithmetic tests. The literature suggests that children at about the fifth-grade level are particularly test anxious (Jersild, Goldman, and Loftus, 1940; Angelino, Dollins, and Mech, 1956). Castaneda, Palermo, and McCandless (1956) found a significant interaction between anxiety and test complexity in a study involving fifth graders.

It would seem reasonable that the method of ordering test items according to difficulty might have some effect on the performance of children on the test, and that this effect might operate differentially across various levels of anxiety. The construction of various forms of a test for pupils at different anxiety levels would be a step in the direction of accounting for individual differences in the evaluation phase of instruction.

3. Research Design and Procedure

The subjects were 53 students in two fifth-grade classes in a rural elementary school in Albany, Wisconsin. Sarason's Test Anxiety Scale for Children (TASC) (Sarason, Davidson, Lighthall and White, 1958) was administered to each class by one of the investigators. The scores on the TASC were divided into thirds, thus creating three test-anxiety groups: low (L), intermediate (I), and high (H). Table I gives the range of scores and the number of subjects in each group.

TABLE I

CLASSIFICATION OF ANXIETY TYPES

Group	Number of "Yes" responses on TASC (Maximum - 30)	Number of Subjects
L	0 - 7	16
I	8 - 13	18
H	14 - 26	17

One week later the Mathematics Computation part of the 1970 edition of the Metropolitan Achievement Test Battery was administered by the classroom teachers to the students. Three forms of the test, each differing in the ordering of the difficulty, were used. In Form A the questions were ranked according to their p-values in ascending order of difficulty; in Form D the items were placed in descending order of difficulty; and in Form M items were arranged in a mixed order of difficulty. The three forms of the test were randomly assigned to the subjects within each anxiety level.

Fifty pupils took both the TASC and mathematics test. Before analyzing the data, five cases were dropped from four cells in order to use an equal-cell two-way ANOVA computer program.

4. Findings

The grand mean on the mathematics computation test was 24.80 of a possible 40. Cell means, anxiety group means, and means for each form of the test are given in Table II. None of the three F values for anxiety was significant at $\alpha = .05$ ($p < .025$).

TABLE II

CELL MEANS

Mathematics Computation Test Form	Test Anxiety Level			
	High	Intermediate	Low	Total
D	24.00	24.40	24.40	24.20
M	25.00	23.20	30.40	26.20
A	21.40	24.40	26.20	24.00
Total	23.47	24.00	26.93	24.80

65

5. Interpretations

No significant differences were found between the test scores of subjects in the three anxiety groups, or between the scores achieved on the three forms of the test. The interaction between test-anxiety level and the method of ordering of test items was not significant.

The authors noted two trends in the data (Table II):

1. Students who took the mixed form of the test tended to score higher than students taking the other two forms.

2. Low test-anxious students tended to perform better than intermediate or high test-anxious students.

They suggested that the statistically non-significant results may have been due to the small number of students used and recommended that the study be replicated with a larger sample. The authors also expressed concern about the procedure used to classify students into anxiety groups and the positively skewed distribution of scores on the computation test.

Critical Commentary

In their discussion the authors tended to ignore, or not accept, the statistical results obtained. Although it was found that differences in scores between anxiety groups and test forms were due to chance, the authors state: "Low test-anxiety seemed to have a rather facilitating effect on test performance . . . Students who took the mixed form of the examination scored higher than their classmates . . .". A possible explanation of the latter "result" was offered.

The authors rightly expressed concern about the small sample size and the procedure used to determine the anxiety groups. It had been pointed out earlier in the article that in most of his studies with the TASC, Sarason used the top and bottom 15% of scores to establish high- and low-anxious groups. Using this technique rather than comparing thirds, and employing a larger and more widely representative sample would more likely reveal any effect due to anxiety. An estimate of the reliability of the TASC should be provided to assist readers in interpreting the results.

The nature and format of the criterion instrument should also be considered. The test was multiple choice and the 40 computational items appeared (at least on the original) on two pages. It is possible that regardless of the order of items pupils might skip over questions they viewed as difficult (e.g., fractions) or guess at the answers in order to spend time first on the questions with which they felt most competent. Indeed, the Teacher's Handbook for the test suggests that teachers encourage pupils not to get "stuck" on an item but to move on and come back to it later if there is time. Would ordering by item difficulty have more effect on a non-multiple-choice, problem-solving type of test?

James H. Vance
University of Victoria

MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN RESOURCES IN EDUCATION
October - December 1976

ED 123 805 Loucks, Susan F. An Exploration of Levels of Use of an Innovation and the Relationship to Student Achievement. 15p. MF and HC available from EDRS.

ED 123 822 Chow, Stanley H. L.; And Others. Effects of a Mediated Training Course on Teachers and Students in Mainstreaming Programs. 31p. MF and HC available from EDRS.

ED 123 979 Hedges, Larry V.; Majer, Kenneth. Female and Minority Representation in College Majors as a Function of Mathematics Requirements. 14p. MF and HC available from EDRS.

ED 124 200 Lorton, Paul, Jr.; Searle, Barbara W. Using a Linear Regression Model for Item Selection in C. A. I. 23p. MF and HC available from EDRS.

ED 124 416 Hooper, Frank H.; And Others. A Representative Series of Piagetian Concrete Operations Tasks. Theoretical Paper No. 57. 100p. MF and HC available from EDRS.

ED 124 418 Malin, Jane T. Instructing Junior High School Students in Problem Solving Strategies. 36p. MF and HC available from EDRS.

ED 124 926 Bye, M. P. Reading in Mathematics and Cognitive Development. 18p. MF and HC available from EDRS.

ED 125 160 Panel Evaluation of 19 Pre-College Curriculum Development Projects, December 8-12, 1975. National Science Foundation, Washington, D.C., Directorate for Science Education. 620p. MF and HC available from EDRS.

ED 125 609 Searle, Barbara; And Others. Application of Radio to Teaching Elementary Mathematics in a Developing Country. Third Annual Report. 47p. MF and HC available from EDRS.

ED 125 688 Baldwin, James; And Others. Survey of Developmental Mathematics Courses at Colleges in the United States. 64p. MF and HC available from EDRS.

ED 125 689 Roberts, Frank C. A Cost-Effective Analysis and Follow-up Study on a Multi-Level Mathematics Instruction System at Antelope Valley College. 54p. MF and HC available from EDRS.

ED 125 896 Huberty, Carl J.; Perry, Marcia D. Using Discriminant Analysis in the Homogeneous Grouping of High School Mathematics Students. 14p. MF available from EDRS. HC not available from EDRS.

ED 125 907 Good, Thomas L.; Grouws, Douglas A. Process-Product Relationships in Fourth Grade Mathematics Classrooms. 123p. MF and HC available from EDRS.

ED 125 908 The NIE Conference on Basic Mathematical Skills and Learning (Euclid, Ohio, Octover 4-6, 1975). Volume I: Contributed Position Papers. 234p. MF and HC available from EDRS.

ED 125 909 The NIE Conference on Basic Mathematical Skills and Learning (Euclid, Ohio, October 4-6, 1975). Volume II: Reports from the Working Groups. 49p. MF and HC available from EDRS.

ED 125 917 McDaniel, Ernest D.; Guay, Roland B. Spatial Abilities, Mathematics Achievement, and the Sexes. 15p. MF and HC available from EDRS.

ED 126 194 Williams, Barbara Ivory. The Need to Study Multiple Skills. 30p. MF and HC available from EDRS.

ED 126 678 Otani, Yasuteru. The Influence of the Mother Tongue on Achievement in Mathematics: A Contrastive Study of English and Japanese. Workpapers in Teaching English as a Second Language, Vol. 9. 8p. MF and HC available from EDRS.

ED 127 017 Meyer, William J. The Development and Acquisition of the Abstract Concept Roundness. 45p. MF and HC available from EDRS.

ED 127 039 Meyer, William J. Scanning Behavior in a Classification Task. 17p. MF and HC available from EDRS.

ED 127 137 Gregory, John W. Cornell Conditional Reasoning Test as an Adequate Measure of Conditional Moves in Adolescent Verbal Behavior. 36p. MF and HC available from EDRS.

ED 127 138 Chrisman, Gerry L.; Wheatley, Grayson H. Formal Operational Thought and Learning Strategies in Mathematical Structures. 42p. MF available from EDRS. HC not available from EDRS.

ED 127 140 Farrell, Margaret A.; And Others. Field Testing a Systems Analysis Model for Mathematics/Science Teacher Education. 25p. MF available from EDRS. HC not available from EDRS.

ED 127 142 Uprichard, A. Edward; Phillips, E. Ray. An Intraconcept Analysis of Rational Number Subtraction: A Validation Study. 19p. MF available from EDRS. HC not available from EDRS.

ED 127 198 National Assessment of Education Progress. The First National Assessment of Mathematics: An Overview. Education Commission of the States, Denver, Colorado. National Assessment of Educational Progress. 64p. MF and HC available from EDRS.

ED 127 205 Suydam, Marilyn, Comp. Electronic Hand Calculators: The Implications for Pre-College Education. Final Report. 377p. MF and HC available from EDRS.

ED 127 366 Lambert, Nadine M.; Hartsough, Carolyn S. Beginning Teacher Evaluation Study: Phase II, 1973-74, Final Report: Volume III.1. APPLE Observation Variables and Their Relationship to Reading and Mathematics Achievement. 391p. MF and HC available from EDRS.

ED 127 367 Calfee, Robert; Calfee, Kathryn Hoover. Beginning Teacher Evaluation Study: Phase II, 1973-74, Final Report: Volume II. Description of the Field Study and Sample. 113p. MF and HC available from EDRS.

ED 127 374 Elias, Patricia J.; And Others. Beginning Teacher Evaluation Study: Phase II, 1973-74, Final Report: Volume V.5. The Reports of Teachers About Their Mathematics and Reading Instructional Activities. 125p. MF and HC available from EDRS.

ED 127 375 McDonald, Frederick J. Beginning Teacher Evaluation Study: Phase II, 1973-74, Summary Report. 30p. MF and HC available from EDRS.

MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN JOURNALS AS INDEXED BY CURRENT INDEX TO JOURNALS IN EDUCATION
October - December 1976

EJ 138 142 Goodstein, H. A.; And Others. The Achievement of Educable Mentally Retarded on the Key Math Diagnostic Arithmetic Test. Journal of Special Education, v10 n1, pp61-70, Spring 1976.

EJ 138 499 Hornblum, Judith Newman; Overton, Willis F. Area and Volume Conservation Among the Elderly: Assessment and Training. Developmental Psychology, v12 n1, pp68-74, January 1976.

EJ 138 652 Barrera, Manuel, Jr.; Glasgow, Russell E. Design and Evaluation of a Personalized Instruction Course in Behavioral Self-Control. Teaching of Psychology, v3 n2, pp84-88, April 1976.

EJ 140 081 Novillis, Carol F. An Analysis of the Fraction Concept into a Hierarchy of Selected Subconcepts and the Testing of the Hierarchical Dependencies. Journal for Research in Mathematics Education, v7 n3, pp131-144, May 1976.

EJ 140 082 Wheatley, Grayson H. A Comparison of Two Methods of Column Addition. Journal for Research in Mathematics Education, v7 n3, pp145-154, May 1976.

EJ 140 083 Grouws, Douglas A.; Good, Thomas L. Factors Associated with Third- and Fourth-Grade Children's Performance in Solving Multiplication and Division Sentences. Journal for Research in Mathematics Education, v7 n3, pp155-171, May 1976.

EJ 140 084 Kuehls, Ernest A. Effect of Interspersed Questions on Learning From Mathematical Text. Journal for Research in Mathematics Education, v7 n3, pp172-175, May 1976.

EJ 140 085 Barszcz, Edward L.; Gentille, J. Ronald. Retroactive Interference of Similar Methods to Teach Translation of Base Systems in Mathematics. Journal for Research in Mathematics Education, v7 n3, pp176-182, May 1976.

EJ 140 086 Williams, Eddie R.; Mick, Harold W. Measuring the Effectiveness of Using Slide-Tape Lessons in Teaching Basic Algebra to Mathematically Disadvantaged Students. Journal for Research in Mathematics Education, v7 n3, pp183-192, May 1976.

EJ 140 117 Griggs, Shirley A. An Experimental Program for Corrective Mathematics in Schools for the Socially Maladjusted and Emotionally Disturbed. School Science and Mathematics, v76 n5, pp377-380, May-June 1976.

EJ 140 199 Williams, Eddie R.; Mick, Harold W. Research and Development of Synchronized Slide-Tape Units for a Mathematics Laboratory. Two-Year College Mathematics Journal, v7 n2, pp28-33, May 1976.

EJ 140 324 Lumb, David; Child, Dennis. Changing Attitudes to the Subject, and the Teaching of Mathematics Amongst Student Teachers. Educational Studies, v2 n1, pp1-10, March 1976.

EJ 140 532 Stanley, Julian C. Psychology in Action: Test Better Finder of Great Math Talent Than Teachers Are. American Psychologist, v31 n4, pp313-314, April 1976.

EJ 140 777 Crable, Elain; Dinnan, James A. A Question of Teacher Competency. Southern Journal of Educational Research, v10 n1, pp51-58, Winter 1976.

EJ 142 051 Aiken, Lewis R., Jr. Update on Attitudes and Other Affective Variables in Learning Mathematics. Review of Educational Research, v46 n2, pp293-311, Spring 1976.