DOCUMENT RESUME

ED 132 032 SE 021 658

AUTHOR Osborne, Alan R., Ed.
TITLE Investigations in Mathematics Education, Vol. 9, No. 2.
INSTITUTION Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE 76
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; *Cognitive Development; Elementary Secondary Education; Geometry; Instruction; Learning; Logic; *Mathematics Education; Ratios (Mathematics); Research; *Research Reviews (Publications); Subtraction; Teacher Education
IDENTIFIERS Calculators

ABSTRACT

Thirteen different research reports related to mathematics education, plus a set of five studies by one investigator and his associates, are abstracted and analyzed. The set of five studies is primarily concerned with how students deal with ratio. Of the other studies, one focuses on the longitudinal effects of inservice teacher training; five deal with student understanding of topics in mathematics (problem solving, logic, geometry, and subtraction); three are concerned with instructional procedures (college remedial mathematics, CAI, and use of calculators); two look at questions in learning theory; and two focus on sociological concerns. Research related to mathematics education which was reported in RIE and CIJE between January and March 1976 is listed. (DT)

Vol. 9, No. 2 Spring 1976

INVESTIGATIONS IN MATHEMATICS EDUCATION

INVES
IN
MATH
EDUC

Expanded Abstracts and Critical Analyses of Recent Research

Center for Science and Mathematics Education
College of Education
The Ohio State University
in cooperation with the
ERIC Science, Mathematics, and
Environmental Education Clearinghouse

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 9, Number 2 - Spring 1976

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...

Two reports abstracted in this issue of Investigations in Mathematics Education deserve commentary. Each in its own manner suggests the productive potential of collaborative efforts in research in mathematics education.

First, Merlyn Behr provides an abstract of an evaluation report of an extensive inservice education project conducted in California. The evaluation report of the Specialized Teacher Project is notable in that few teacher education programs at either the preservice or inservice levels are designed to allow sound application of principles of evaluation. The natural and reasonable limitations such as size, implementation, and staffing impose constraints on the possible and appropriate for evaluation and research. The Final Report of the Specialized Teacher Project is one of a small number of evaluations that indicates inservice education does make a difference in the mathematical performance of children.

Behr's abstract identifies several significant questions for teacher education that could be productively examined in other such studies. As I read his commentary it was easy to conclude that most of the questions could not be studies without collaborative efforts spanning several institutions of teacher education. Most single institutions do not conduct sufficiently large teacher education programs to conduct thorough research or evaluation studies at either the preservice or inservice level.

Second, Jack Easley and Ken Travers abstract and comment on a set of five studies by Karplus and his associates. The studies are primarily concerned with how children and youth deal with the concept of ratio. Karplus' research has evolved over several years. A longitudinal examination of how the research questions about a single topic evolve and mature is instructive. One can see how ideas coalesce and begin to have a cumulative effect.

The Karplus studies suggest again the wisdom of collaborative efforts in research. The fact of five or more studies focused on a single topic has an additive potential for finding new knowledge about the learning and teaching of mathematics. If researchers could cooperate in identifying questions and conducting research on a topic, then the time to produce a set of related studies like those of Karplus and his associates can effectively be shortened and the development of new knowledge of the teaching and learning processes for mathematics accelerated.

The research community in mathematics education is only beginning to realize the potential of collaborative research efforts. Until recently little evidence of such cooperation has existed and few mechanisms for encouraging collaboration have been utilized by funding agencies, although many of the important problems in mathematics education are large enough to benefit by such efforts. Recently the PMDC Project of The Florida State University tapped the wisdom and talent of several researchers at three institutions to examine the learning of young children. The University of Georgia Center for the Study of the Learning and Teaching of Mathematics held a series of five research workshops in the spring of 1975 with the express goal of stimulating collaborative efforts in research. The workshops were concerned with the topics:

1. Teaching Strategies*
2. Number and Measurement Concepts*
3. Space and Geometry Concepts

4. Models for Learning Mathematics*

5. Problem Solving

From the mutual stimulus and enthusiasm of the workshops, twelve different working groups evolved. Several of the working groups are conducting collaborative research. Each group appears to be attacking problems related to a single topic with the purpose of being systematic. We hope this has payoff in the near future.

Collaborative, cooperative research is a relatively new phenomenon in mathematics education. Few precedents exist in the field for a group of researchers addressing a single problem area. Undoubtedly, the community of scholars in mathematics education will need to learn some things about collaboration. At points this learning may be painful but on the whole it augurs well for the field. Collaborative, cooperative efforts need to be encouraged and sustained.

Alan R. Osborne
Editor

*The proceedings of each of these workshops will be available from ERIC/SMEAC, The Ohio State University, 1200 Chambers Road, Third Floor, Columbus, Ohio 43212. As this issue of I.M.E. goes to press, the proceedings of the workshops marked with an asterisk are already available at $4.00 each. The others will be available later this year.

Spring 1976

Bassler, Otto C.; And Others. Comparison of Two Instructional Strategies for Teaching the Solution of Verbal Problems. Journal for Research in Mathematics Education, v6 n2, pp170-177 May 1975.
Abstracted by STEPHEN S. WILLOUGHBY 1

Corn, J.; Behr, A. A Comparison of Three Methods of Teaching Remedial Mathematics as Measured by Results in a Follow-Up Course. MATYC Journal, v9 n1, pp9-13, Winter 1975.
Abstracted by HAROLD C. TRIMBLE 6

Eisenberg, Theodore A. Negation, Disjunction Syllogisms, and Mathematics Achievement. Journal of Psychology, v90, pp1969-1974, May 1975.
Abstracted by LARS C. JANSSON 8

Fisher, Maurice D.; And Others. Effects of Student Control and Choice on Engagement in a CAI Arithmetic Task in a Low-Income School. Journal of Educational Psychology, v67, pp776-783, December 1975.
Abstracted by KENNETH E. VOS 11

Flaherty, E. G. The Thinking Aloud Technique and Problem Solving Ability. Journal of Educational Research, v68, pp223-225, February 1975.
Abstracted by GERALD A. GOLDIN 15

Geeslin, William E.; Shavelson, Richard J. An Exploratory Analysis of the Representation of a Mathematical Structure in Students' Cognitive Structures. American Educational Research Journal, v12, pp21-39, Winter 1975.
Abstracted by JAMES M. MOSER 19

Genkins, Elaine F. The Concept of Bilateral Symmetry in Young Children. In Children's Mathematical Concepts. Rosskopf, M. F. (Ed.) Teachers College Press, New York, pp5-44.
Abstracted by THOMAS E. KIEREN 23

Hancock, Robert R. Cognitive Factors and Their Interaction with Instructional Mode. Journal for Research in Mathematics Education, v6 n1, pp37-50, January 1975.
Abstracted by OTTO BASSLER 27

Jantz, Richard K.; Sciara, Frank J. Does Living with a Female Head of Household Affect the Arithmetic Achievement of Black Fourth Grade Pupils? Psychology in the Schools, v12, pp468-472, October 1975.
Abstracted by EDWARD M. CARROLL 30

"The Karplus Studies"
Karplus, Elizabeth F.; Karplus, Robert. Intellectual Development Beyond Elementary School I: Deductive Logic. School Science and Mathematics, v70, pp398-406, May 1970.
Karplus, Robert; Peterson, Rita W. Intellectual Development Beyond Elementary School II: Ratio, A Survey. School Science and Mathematics, v70, pp813-820, December 1970.
Karplus, Robert; Karplus, Elizabeth F. Intellectual Development Beyond Elementary School III--Ratio: A Longitudinal Study. School Science and Mathematics, v72, pp735-742, November 1972.
Karplus, Elizabeth F.; And Others. Intellectual Development Beyond Elementary School IV: Ratio, The Influence of Cognitive Style. School Science and Mathematics, v74, pp476-482, October 1974.
Wollman, Warren; Karplus, Robert. Intellectual Development Beyond Elementary School V: Using Ratio in Differing Tasks. School Science and Mathematics, v74, pp573-613, November 1974.
Abstracted by JOHN A. EASLEY, JR. and KENNETH J. TRAVERS 34

Schafer, Pauline; And Others. Calculators in Some Fifth-Grade Classrooms: A Preliminary Look. Elementary School Journal, v76, pp27-31, October 1975.
Abstracted by JOHN E. TARR and JACK D. WILKINSON 44

Scott, Ralph. Shifts in Reading Readiness Profiles During the Past Decade. Journal of Genetic Psychology, v126, pp269-273, June 1975.
Abstracted by THOMAS R. POST 47

State Board of Education, San Diego, California. Final Report, Specialized Teacher Project, 1971-72. Mathematics Improvement Projects.
Abstracted by MERLYN J. BEHR 51

Woods, Shirley S.; And Others. An Experimental Test of Five Process Models for Subtraction. Journal of Educational Psychology, v67, pp17-21, February 1975.
Abstracted by ROBERT D. BECHTEL 55

Mathematics Education Research Studies Reported in Research in Education (January - March 1976) 59

Mathematics Education Research Studies Reported in Journals as Indexed by Current Index to Journals in Education (January - March 1976) 63

8

COMPARISON OF TWO INSTRUCTIONAL STRATEGIES FOR TEACHING THE SOLUTION OF VERBAL PROBLEMS. Bassler, Otto C.; Beers, Morris I.; Richardson, Lloyd I. Journal for Research in Mathematics Education, v6, pp170-177; May 1975.

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

1. Purpose

To assess the relative effects of two strategies of instructing students to solve verbal problems. The two strategies studied are called the "Polya Method (PM)" and the "Dahmus Method (DM)." In PM, the problem solver follows a sequence of steps: (1) read and understand; (2) plan (including identifying unknowns, relations, operations and analogies to other problems); (3) carry out the plan; and (4) examine the obtained solution. In DM the problem solver translates the verbal problem piecemeal into mathematical symbolism.

Hypothesis 1: DM subjects would achieve significantly higher translation scores (i.e., write an appropriate mathematical equation).

Hypothesis 2: PM subjects would achieve significantly higher solution scores (i.e., get the correct answer with or without an equation).

2. Rationale

There have been many studies comparing problem solving strategies and many articles arguing for particular methods. The topic seems important and the results of previous studies seem inconclusive. Since the emphasis in DM is on translation while the emphasis in PM is on understanding the problems, the hypothesis seems reasonable. Although the sample was divided into three ability levels, there seemed to be no reason for hypothesizing interaction between ability and treatment.

3. Research Design and Procedure

Forty-eight students from three ninth-grade algebra classes in a private, Catholic girls' secondary school in Nashville, Tennessee were divided into six cells (3 ability groups x 2 treatment groups) with 8 students per cell. Ability groups were determined by the Orleans-Hanna Algebra Prognosis Test.

An "individualized instruction" technique involving colored slides and taped programs was used to teach students the detailed steps of the particular strategy they were to use. The tape recorder was stopped when individual written work was required. Each program was designed to consume seven 40-minute instructional periods. According to the authors, DM "required somewhat more instruction than PM, but time allowed was roughly equal in both treatments."

Posttests were administered on the day following instruction, and retention tests were administered approximately four weeks later. The tests consisted of ten verbal problems, seven of which were similar to those presented in the instruction and three of which were more complex. Students were instructed to find an equation and solve it, but it was stressed that even if they could not find an equation they should try to find a solution. This procedure resulted in two scores, one for correct equations and one for correct solutions. The retention test items were essentially identical to the posttest items except that numerical values and contexts were changed. The Cronbach alpha reliability estimate for the posttest equation scores was .84 and for solution scores was .78.

A three-factor analysis of variance was used to analyze the data. The factors were ability (high, intermediate, and low as measured by Orleans-Hanna Algebra Test), treatment (PM, DM), and test occasion (post, 4-week retention). Separate analyses were made for equation scores and for solution scores.

4. Findings

Means and standard deviations are reported in Tables 1 and 2.

TABLE 1

Means ($\overline{X}$) and Standard Deviations (SD) of Scores on the Equation Criterion of the Posttest and Retention Test for Each Cell

	Treatment			
	PM		DM	
Ability Level	$\overline{X}$	SD	$\overline{X}$	SD
Posttest				
High	12.5	2.6	9.6	5.6
Intermediate	8.3	5.6	2.1	3.1
Low	4.9	2.3	1.4	1.2
Retention Test				
High	13.1	5.0	8.6	6.3
Intermediate	9.6	4.9	1.9	1.6
Low	3.9	3.0	.4	.7

(n per cell = 8)
(score range 0-20)

TABLE 2

Means ($\overline{X}$) and Standard Deviations (SD) of Scores on the Problem Solution Criterion of the Posttest and Retention Test for Each Cell

	Treatment			
	PM		DM	
Ability Level	$\overline{X}$	SD	$\overline{X}$	SD
Posttest				
High	12.5	4.0	12.5	4.8
Intermediate	9.5	4.9	7.6	3.9
Low	6.8	4.9	5.0	3.2
Retention Test				
High	14.4	4.3	12.0	4.3
Intermediate	9.0	4.5	9.3	5.4
Low	8.6	4.4	7.1	3.4

(n per cell = 8)
(score range 0-20)

At the .05 level of significance the results of the analysis of variance were:

For Equation Criterion:

1. There were no significant interactions of any order.

2. The effect of treatment was significant. PM subjects had a significantly larger mean score than the DM subjects on the equation criterion.

3. There was a significant ability effect. Although individual comparisons were not made, the mean score of the high-ability subjects was greater than that for the intermediate-ability subjects, which, in turn, was greater than for the low-ability subjects.

4. There was no significant difference between the mean posttest score and the mean retention test score.

For Solution Criterion:

1. There were no significant interactions of any order.

2. There was no significant effect due to treatment.

3. There was a significant ability effect. It appears that the mean for the high-ability subjects is greater than that for the middle-ability subjects, which, in turn, is greater than for the low-ability subjects.

4. There was a significant test occasion effect. The mean for the retention test is significantly larger than the mean for the posttest.

Since subjects were encouraged to determine solutions even if they could not write equations, higher scores on the solutions criterion than on the equations criterion were not ruled out, and in fact occurred for all ability groups.

5. Interpretations

Ninth-grade algebra students taught to solve verbal problems using only teaching machines involving colored slides and taped programs apparently are able to write correct equations better if they are taught by the Polya Method (PM) rather than the Dahmus Method (DM). However, the investigators call attention to the fact that the level of attainment of all students was not as high as the investigators had hoped, and suggest that this lack of learning by all may have suppressed differences between groups. Also, the students had all learned to solve some verbal problems by a PM-type technique in earlier grades which may have inhibited the DM learners but reinforced the PM learners.

The authors claim that DM subjects did translate correctly but simply failed to combine the translation into an acceptable equation.

The higher retention scores than posttest scores suggest that problem solving is resistent to forgetting, since the investigators were assured that no direct instruction pertaining to solution of verbal problems occurred between the two tests. Most of the gain occurred on the three problems that were more complex for the problem-solution criterion, while the subjects made almost no gain on problems similar to those on which they had received instruction.

The authors suggest that both methods appear to have benefits and perhaps should be combined in regular instruction, but used as pure strategies for remedial work.

Critical Commentary

Verbal problem solving is an important aspect of mathematics teaching, and the authors have made an interesting contribution to our understanding of learning to solve problems. However, there are limitations to the study which make it hard to generalize to a typical classroom situation. These limitations include:

1. The fact that all teaching was done using machines.

2. The fact that all subjects were girls (from a Catholic Parochial school with better than average students). This may be an especially important limitation in light of the report from the National Assessment indicating sex differences in solving verbal problems.

3. The very small number of subjects in the experiment could account for the lack of statistically significant differences on the solutions criterion.

4. Since it was known that students had received training in previous grades in solving verbal problems, perhaps a pretest on verbal problems similar to the posttest would have been more appropriate than the Orleans-Hanna Algebra Progress test. As things stand, there is the possibility that students knew more about problem solving before the treatment than after.

Stephen S. Willoughby
New York University

A COMPARISON OF THREE METHODS OF TEACHING REMEDIAL MATHEMATICS AS MEASURED BY RESULTS IN A FOLLOW-UP COURSE. Corn, J.; Behr, A. MATYC Journal, v9 n1, pp9-13, Winter 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Harold C. Trimble, The Ohio State University.

1. Purpose

"To measure the success of these [three remedial] methods by student achievement in a follow-up course." The three "methods" compared in this study were "conventional," "modular" and "programmed" as employed in a non-credit remedial mathematics course for students entering the City University of New York.

2. Rationale

Such new methods as modular, programmed, CAI, audio-visual, tutorial, etc. have often been evaluated through end-of-course measures. But "rarely have these methods been measured in terms of their capacity to prepare a student for a follow-up course." These investigators assume that at least one measure of "good" remedial mathematics is performance in a subsequent course in mathematics.

3. Research Design and Procedure.

This was an after-the-fact study. That is, existing records were used at the end of the third term following a remedial course. Students who passed the remedial course under each method were classified as passing, not passing, or not registering in a follow-up course during this time period; grade point averages in follow-up courses in mathematics were computed; overall grade point averages in the subsequent courses were compared; numbers of students originally enrolled in the remedial course (whether or not they passed) were used as a base for comparing numbers passing a subsequent course in mathematics; even the numbers of students still in school were considered. The investigators clearly recognize the limitations imposed by lack of control and a sample too small to permit following students into specific subsequent courses. Yet they recommend the idea of studies of follow-up work for the purpose of evaluating approaches to remedial mathematics.

4. Findings

Among the several statistical comparisons made, only the one concerned with students passing the remedial course and then passing a follow-up course in mathematics yielded a significant difference. The source of the difference was identified as being between conventional and modular classes. The investigators recognized that follow-up courses were taught by the conventional method and, hence, that practice effect in learning by this method may explain this difference. They recognized several other limitations of their study too and suggest ways to minimize these in future research.

5. Interpretations

The suggestion in the data that the conventional method is superior as judged by the performances in a follow-up course of students who passed the remedial course is limited by lack of control. The investigators propose longer range studies, with better defined alternative remedial methods, allowances for differences in initial knowledge of mathematics, etc.

Critical Commentary

The idea of follow-up to assess program effectiveness is, of course, not new. Often it is a difficult one to implement. For example, it is costly to follow graduates from a teacher education program and achieve measures of the relative effectiveness of the program. Changes in the program and limited sample sizes tend to muddy the waters. As applied to courses in remedial mathematics, the idea seems eminently workable. Many schools will have adequate numbers of students to implement the suggestions for further research made by these investigators. This reviewer does agree that "The proof of the pudding is in the eating."

Harold C. Trimble
The Ohio State University

NEGATION, DISJUNCTION SYLLOGISMS, AND MATHEMATICS ACHIEVEMENT. Eisenberg, Theodore A. Journal of Psychology, v90, pp1969-1974, May 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Lars C. Jansson, University of Manitoba.

1. Purpose

To determine if students who are high in mathematical achievement would correctly assess disjunctive syllogisms with positively and negatively worded premise statements more often than students low in mathematical achievement, and if they would be more confident of their answers. A disjunctive syllogism is of the form PVQ, X, where X $\in$ {P, -P, Q, -Q}.

2. Rationale

A number of earlier research studies suggested that students of all ages have more trouble assessing the truth-value of negatively worded sentences than positively worded ones. This is particularly true when negatively worded statements are used as premises in implicational syllogisms. Moreover, the ability to assess implicational syllogisms correctly when they are worded in a negative manner does not increase solely by maturation. Because of the positive correlation between logical reasoning ability and mathematical achievement, it was hypothesized that students of high mathematical achievement would have less trouble than those of low mathematical achievement in assessing disjunctive syllogisms whose component sentences were negatively worded.

3. Research Design and Procedure

A 32-item test incorporating the four types of disjunctive syllogism was used. The types were identified as follows:

I. P V Q, -P / Q ?	III. P V Q, -Q / P ?
II. P V Q, Q / P ?	IV. P V Q, P / Q ?

For each type, the first premise could be worded affirmatively (A) or negatively (N), and thus appeared in four forms with two questions each (AA, AN, NA, NN). To each item the S could answer YES, NO, or MAYBE. Following each such item he was asked, "Are you confident that your answer is correct? ... YES ... NO." This logic test and the SMSG/NLSMA Mathematics Inventory I examination were administered to 550 male and female grade 8/9 algebra students with no prior logical training. High and low mathematical achievement groups were obtained by arbitrarily selecting the top and bottom 150 students according to scores on the Mathematics Inventory.

For each item type and wording form, the mean percentage of the population in high and low math groups answering correctly (according to formal logic) were compared using a t-test (16 comparisons). Similarly, levels of confidence were compared using t-tests.

4. Findings

Percentage of items correct for the high group ranged from 15 to 54, for the low group from 13 to 63. No comparisons showed significant differences and all directional hypotheses were rejected.

5. Interpretations

Low math students unexpectedly outperformed high math students on Type I and Type III items, although high math students expressed more confidence in their responses. Low math students were not reluctant to choose MAYBE on item types II and IV and thus it was not the case that one type of student effectively faced a YES/NO option while the other group had three choices. The noted trend for Types I and III deserves further study as does the high confidence expressed by both groups in their answers, especially in light of generally poor performance.

Critical Commentary

1. While the author refers several times to the fact that Type I and Type III items were better handled by the low math group, the fact remains that no significant differences were observed. The trend, however, deserves notice. The level of significance and whether a one- or two-tailed test was used was not stated.

2. It is noted that high math students had more confidence in their responses (whether justifiably or not) than low math achievers. Again, an interesting trend, but n.s.d.

3. No comparison was made of item types, i.e., either by math achievement groups, or among item types for both groups. Research by this reviewer on implicational syllogisms suggests that type is a far more important contributor to performance variability than the location of negations in the premises.

4. It would be interesting to know if any attempt was made to interpret the data in terms of the inclusive-exclusive or question. Could one or the other of these interpretations by the Ss help to explain the data?

5. While the rationale and directional hypotheses seemed warranted, a rather large sample yielded no significant difference. This provides information of a sort--if the investigator-constructed test is reliable, a point on which no information is given.

6. There is no indication of what effect, if any, the confidence questions may have had on the response pattern to the logic items. Is there any relationship between an individual's response pattern and his degree of confidence? Did any individual answer all items (or all of a particular sort) correctly? If so, was he confident of his correctness?

The issues of concern in this research deserve continuing attention, and a number of directions have been suggested. While there is little in this study of immediate import to curriculum builders, it is a contribution to the growing bodies of research literature in the areas of logical reasoning and psycholinguistics which will in the future have much to say to mathematics educators concerned with logical development and the language of mathematics.

Lars C. Jansson
University of Manitoba

EFFECTS OF STUDENT CONTROL AND CHOICE ON ENGAGEMENT IN A CAI ARITHMETIC TASK IN A LOW-INCOME SCHOOL. Fisher, Maurice D.; Blackwell, Laird R.; Garcia, Angela B.; Green, Jennifer C. Journal of Educational Psychology, v67, pp776-783, December 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Kenneth E. Vos, The College of St. Catherine.

1. Purpose

The purposes are to study:

a) the effects of difficulty level and choice of difficulty level on the "engagement" of students using a computer-assisted instruction (CAI) program on mathematics problems,

b) possible patterns of student choices, and

c) the relationship of locus of control measures to choice of difficulty levels.

2. Rationale

It was assumed that a learner would choose arithmetic problems with difficulty levels that would produce "optimal environmental stimulation". Previous studies had found that those problems which are "optimally stimulating" produce the desirable match between a learner's degree of understanding and the difficulty levels of the curriculum.

3. Research Design and Procedure

The experiment was conducted with 4th and 5th grade students from a low-income and 90% Mexican-American background school in California. From a pool of 100 students, 19 subjects were randomly selected and classified the Choice group. From those remaining in the pool, 19 subjects were selected as the No-choice "yoked control" group. Each subject in the No-choice group was matched to a subject in the Choice group on sex, school grade, age, initial mathematics achievement level (Metropolitan Achievement Test or Stanford Achievement Test), ethnic background, and classroom teacher. Neither group had previous CAI experience.

The CAI program was administered via two teletypewriters located in a separate location from the students' classrooms. Due to previous experiences by the authors, pictures and toys were placed within easy reach of the subject during the experiment in order to counteract the "intensive attracting influence" of CAI.

The basic instructional material was the Stanford CAI Arithmetic Strands Program (Suppes, et al.). Both groups began the first of a total of 15 CAI sessions at grade 1.0 difficulty level. Both groups worked for an initial 15 minutes each session on problems presented by the computer.

After the initial 15 minutes the Choice group were asked by computer if they wanted more problems. If the response was positive, they selected harder, easier, or similar problems. The choice was presented each five-minute period for a maximum of four choices each session. Choice subjects could select difficulty levels ranging from grade 1.0 to grade 7.5. The No-choice subjects did not have the option of selecting a difficulty level, but rather worked the same length of time and at about the same difficulty level as their matched counterpart.

Assessment involved an engagement/disengagement observation and a locus of control measure. The engagement/disengagement observation instrument included ten engagement categories (e.g., touching keys, eyes on paper) and four disengagement categories (e.g., turns toward toys, looks away from keys). The locus of control measure had four dimensions; stable/unstable, control/no control, internal/external, and self/other.

Measures were obtained for length of time each session, correct responses, number of problems completed per five-minute interval, pattern of choice, difficulty level, engagement level, and locus of control dimensions. Data were analyzed by ANOVA and t-test procedures for matched pairs.

4. Findings

(a) Both groups averaged 23 minutes per session. The No-choice group had a significantly higher ($p < .01$) percentage of correct responses (84%) than the Choice group (80%). A significant difference (.01) in the average number of problems completed per five-minute period was detected; the No-choice group averaged 40 and the Choice group averaged 31. The initial mathematics achievement grand mean was 3.7 while the average difficulty level of problems was 2.5 and 2.3 for Choice and No-choice groups respectively. The Choice group's average engagement level was significantly higher ($p < .05$) than the No-choice engagement level. No significant differences were found for disengagement scores. A significant decline ($p < .05$) in engagement and a significant increase ($p < .01$) in disengagement occurred over time. Also a significant decrease ($p < .01$) on the length of time spent on the computer occurred over time.

(b) When patterns of choices by the Choice group were considered, eight subjects, classified as "maximizers," consistently chose easier or similar problems; five subjects, classified as "minimizers," consistently chose harder or similar problems; six subjects, classified as "few-choices," made too few choices to determine any pattern. "Maximizers" had the highest engagement levels with difficulty levels lower than their initial achievement level. "Minimizers" had a higher engagement level when they chose easier problems; neverthess they consistently chose more difficult problems with lower engagement levels.

(c) On the locus of control measure, the Choice group selected significantly ($p < .01$) fewer stable and internal characteristics, and significantly ($p < .05$) more self characteristics than the No-choice group. There was no significant difference on the control/no control dimension.

5. Interpretations

No matter how the data were analyzed, this conclusion was evident: No-choice subjects worked more problems and did so with greater accuracy. Since the engagement levels were consistently higher for the Choice group, this could indicate that there was more interest in the choosing factor than in actually solving the problems. Further research is needed to determine how the choosing factor and engagement levels are related to academic achievement.

Based on the finding that engagement levels decreased as difficulty levels increased, it was stated that "...optimally stimulating problem difficulty levels seem to be related more to the choice patterns of the subjects than to one theoretical motivation curve that can be applied to all learners".

A choice factor for learners could indicate a control over performance by effort and by choice of difficulty, but this was balanced by a lower academic performance and a higher engagement level.

Critical Commentary

Some of the current research related to curricular concerns have an immediate impact for the classroom teacher. This study is not one of them. Most, if not all, of the reported conclusions are evident to observant teachers. In addition, the clarity of the study was clouded by engagement and locus of control measures which yielded little information.

Although the authors were careful with statistical analyses and procedures, several concerns do arise.

(1) There were only two tables. Table 2 was totally unintelligible. In addition the tabulation of data was inefficient and ineffective.

(2) The locus of control measure was of little value as evidenced by the authors' cursory attention in the results and discussions sections. It could have been deleted from the study.

(3) The authors did not discuss certain "interesting" items:

(a) The population of 90% Mexican-American low-income background and its impact on the results was not mentioned. It seemed to be deliberately avoided.

(b) Students ("minimizers") choosing harder problems had an initial mean achievement level of 4.06 and chose a mean difficulty level of 3.22 with 53% correct. Students ("maximizers") choosing easier problems had an initial mean achievement level of 3.39 and chose a mean difficulty level of 1.41 with 91% correct. No mention was made of this result. Does confidence in choosing more difficult problems depend on the student's current mathematics achievement level? A concise study could be designed to answer this particular question.

(c) The authors seemed to be so locked into engagement and control measures that they neglected to pursue why students were choosing difficulty levels two or three grade levels above or below their initial achievement levels.

For this reader the results only reinforced the bias: given a choice, either do not choose, or choose something you know already. Fourteen of 19 subjects in the experiment must also have believed this.

Kenneth E. Vos
College of St. Catherine
St. Paul

THE THINKING ALOUD TECHNIQUE AND PROBLEM SOLVING ABILITY. Flaherty, E. G.
Journal of Educational Research, v68, pp223-225, February 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Gerald A. Goldin, Graduate School of Education, University of Pennsylvania.

1. Purpose

The purposes of the study were:

(a) to study the effects on problem-solving performance of the requirement that subjects "think aloud" during problem solving, including the effect of prior practice in the thinking-aloud technique; and

(b) to develop a coding system for describing subjects' problem-solving processes.

2. Rationale

In the study of problem-solving processes, the researcher must seek to externalize and render observable as much as possible of the problem-solving behavior of each subject. One widely used technique is to require that the subject "think aloud" during problem solving. The experimenter may thus obtain a transcript of the subject's verbal "protocol" for subsequent analysis.

However, it has been suggested that the requirement of verbalization might unduly influence the nature of the phenomena being studied--that it might, for example, improve or impair problem-solving performance, or modify the problem-solving processes that the subjects would otherwise use. Therefore this study compares subjects in a verbalization (thinking-aloud) group with subjects in a non-verbalization group, with respect to (a) the use of correct problem-solving procedures and (b) the use of ten categories of problem-solving processes forming a part of the author's coding scheme.

Prior practice in the thinking-aloud technique is often recommended by researchers, since subjects may be unaccustomed to verbalization during problem solving. The author also examines the effects of such practice on the above-mentioned dependent variables.

Paige and Simon (1966) have classified subjects as either "verbal" or "physical" problem solvers, based on their approaches to problems describing physically impossible situations. "Verbal" problem solvers tend to translate directly from the verbal problem statement to mathematical equations or formulas, while "physical" problem solvers tend to set up some form of internal representation of the physical situation described in the problem. It has been suggested that verbal problem solvers might be impaired more than physical problem solvers by instructions to think aloud, and this possibility is also investigated.

3. Research Design and Procedure

Participating in the study were 100 senior high school students just completing their second year of algebra. The students had been identified by their teachers as "above average" in ability. The author employed a 2x2 factorial design--verbalization (thinking aloud) vs. non-verbalization, and prior practice vs. no prior practice, randomly assigning 25 subjects to each group.

Subjects in the "practice" groups were presented with two algebra word problems prior to the administration of the main problem-solving test. All subjects were then presented with six algebra word problems, including two problems describing physically impossible situations. Subjects in the verbalization groups met individually with the experimenter, who encouraged them to think aloud while working each problem, asking occasional questions to elicit their thinking. These interviews were tape-recorded. Subjects in the non-verbalization group did not meet individually with the experimenter, but submitted written problem solutions.

The six algebra word problems were scored on the basis of whether or not correct procedures were employed, with no credit being taken off for purely computational or arithmetic errors. Two points were allotted for each problem, allowing a maximum score of 12 points. On the "contradictory" problems, full credit was granted if the subject "noted both the grammatical and physical implications of the problem," thus recognizing the contradiction, and partial credit (one point) if the subject either substituted a physically possible problem for the given problem, or rendered a direct translation of the given problem into equation form.

The author's coding system variables are listed below. The written problem solutions submitted by non-verbalization subjects could not be accurately coded for the seven starred categories. Mean scores and standard deviations were obtained for the remaining ten categories; for the verbalization subjects these were based on their recorded problem-solving sessions, and for the non-verbalization subjects these were based on their written problem solutions. These scores were compared for differences of statistical significance.

Coding-System Variables (Flaherty, 1975, from Table 1)

- *1. Misreads problem
- *2. Rewords problem
- 3. Draws diagram
- *4. Indicates familiarity with type of problem
- *5. Notes need for auxiliary information
- *6. Lacks a systematic approach
- 7. Recalls definition or auxiliary information
- 8. Fails to use correct auxiliary cues
- 9. Unsuccessful, adopts new approach
- 10. Fails to retain model of solution
- 11. Makes computational errors
- *12. Indicates concern about method
- *13. Signifies inability to solve problem
- 14. Uses equations
- 15. Uses deduction and arithmetic
- 16. Stops without solution
- 17. Makes structural errors

*Written problem solutions could not be accurately coded for these categories.

Some of the subjects were classified as either verbal or physical problem solvers based on their consistent responses to "grammatical" or "physical" cues in the two contradictory problems. For these subjects scores were computed based on the remaining four (non-contradictory) test problems. Mean scores for the verbal and physical problem solvers were compared for all subjects, and for the thinking-aloud subjects taken separately.

4. Findings

Neither the thinking-aloud requirement nor the prior practice variable had a statistically significant effect on the test scores (use of correct procedures). There was also no significant practice by thinking-aloud interaction.

On nine of the ten coding system variables for which scores were compared, there were no statistically significant differences between the thinking-aloud and the non-verbalization subjects. With respect to the "computational errors" variable, a significant difference ($p < .01$) was found--verbalization subjects tended to make more computational errors.

Of the 100 subjects, 30 were classified as "verbal" and 37 as "physical" problem solvers. For these 67 subjects, there were no significant differences between the two types of problem solvers in their scores on the four non-contradictory problems. Within the thinking-aloud group, there were 15 verbal and 22 physical problem solvers. Here the physical problem solvers did significantly better than the verbal problem solvers ($p < .01$) on the four non-contradictory problems.

5. Interpretations

In general the findings support the use of the thinking-aloud technique as a means of acquiring data on problem solving. The use of this technique had little influence on problem-solving success, and little influence on most of the problem-solving processes examined.

Thinking aloud did seem to have an adverse effect on the verbal problem solvers. This may be a consequence, the author suggests, of the principle that two similar tasks performed simultaneously interfere with each other more than do two dissimilar tasks. For the verbal problem solvers, the thinking-aloud requirement may have been more similar to the problem-solving task itself than it was for the physical problem solvers, and thus may have interfered more with the verbal problem solvers' performance.

Critical Commentary

An investigator must always be alert to the possibility that intervention to take measurements significantly modifies that which is being measured. The author has asked a worthwhile question, in that the extent of such possible modification should certainly be examined for the thinking-aloud technique of acquiring data on problem solving.

Unfortunately the description of the study afforded by this article is inadequate to evaluate the reported findings. No information as to school location, school population characteristics, or subjects' socioeconomic status is given. We do not know the sex of the subjects, nor are any data reported separately for males and females. It is unclear when, how, or by which teachers participating students were identified as of "above average" ability. The procedures used for the "practice" groups are not described; in particular we are not informed how long before the main test the practice problems were given, or whether the practice groups had knowledge of their results on the two practice problems prior to taking the main test. The conditions of administration of the test problems are not described; we do not know, for example, how much time was allotted to each problem.

No raw scores, means, or standard deviations are reported at all. The absence of statistically significant differences between relatively small groups of subjects does not imply that there are no real differences; only that the magnitudes of any differences are not sufficiently large for the investigator to establish their existence with small samples. We are given no information on the size of any of the observed differences, statistically significant or not.

The tests were scored for use of correct procedures, rather than for correct answers, a potentially subtle distinction to define and maintain. There is no discussion of inter-scorer reliability; we do not know if the tests were scored independently by more than one investigator, or if they were scored "blind" by people unaware of the experimental groups to which each subject belonged. Similarly there is no discussion of scoring procedures or inter-scorer reliability for the "verbal" vs. "physical" classification of problem solvers, or for the coding system variables.

In the comparison of the "verbal" with the "physical" problem solvers in the thinking-aloud group, we are not told how many of each group had prior practice and what effect, if any, this may have had.

The author's coding system is suggestive, but many of the categories are vague, and there is no discussion or elaboration provided. A clear need does exist in the field for detailed, reliable, and independently reproducible coding methods for verbal problem-solving protocols.

Gerald A. Goldin
Graduate School of Education
University of Pennsylvania

AN EXPLORATORY ANALYSIS OF THE REPRESENTATION OF A MATHEMATICAL STRUCTURE IN STUDENTS' COGNITIVE STRUCTURES. Geeslin, William E.; Shavelson, Richard J. American Educational Research Journal, v12, pp21-39, Winter 1975.

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

1. Purpose

The purpose of the study was to examine the correspondence between a representation of the structure of a subject matter (content structure) and a representation of this structure in the students' memories (cognitive structures) as a result of instruction in the subject matter.

2. Rationale

Mathematics curriculum revision of the past two decades emphasized an understanding of the structure of mathematics. Yet little empirical work has been done concerning the communication of a structure to students. Structure is defined according to an earlier work by Shavelson; the study used a definition by Begle for mathematical structure. Content structure was defined as a web of concepts and their interrelations in a body of instructional material. This structure was represented by a procedure devised by Shavelson in several earlier studies that uses the theory of directed graphs (digraphs) as described in a work by Harary, Norman, and Cartwright in 1965. Cognitive structure is defined as a hypothetical construct referring to the organization (interrelationships) of concepts in long-term memory. Cognitive structure is examined by a word association technique. The technique is one described by Garskof and Houston, P. E. Johnson, Geeslin, and Shavelson.

3. Research Design and Procedure

The Ss were 87 eighth-grade students from 3 intact classes; most were Anglo-American of average to slightly above average mathematical ability. Ss were randomly assigned to an experimental (E) or control (C) group. The E group (N = 43) received programmed instruction on probability from a text developed by SMSG. The C group studied a programmed text on factors and primes. All Ss received a Word Association (WA) and achievement test on probability as a pretest, posttest (11th or 15th calendar day of study), and retention test (23rd calendar day). In addition, an attitude scale, the Pro-Math Composite developed by NLSMA, was administered to both groups at the inception of the study. The main achievement test consisted of 28 free response and 7 multiple-choice items, the first 30 testing comprehension and the last 5 of a novel, problem-solving nature.

The content structure was analyzed using the digraph procedure. Ten concepts were selected a priori from the probability material: probability, equally likely, outcome, event, experiment, zero, intersection, trial, independent, and mutually exclusive. Then all sentences in the text containing at least two of these concepts were diagrammed using a parsing grammar. From this diagram, a digraph using rules described in an earlier study by Shavelson was constructed. From all digraphs, a "super digraph" was constructed from which a 10x10 similarity matrix was finally derived. This similarity matrix representing content structure was examined using Kruskal's (Psychometrika, 1964) multidimensional scaling procedure. A pictorial model showing the plot of the results was then interpreted as another representation of the structure of the probability text.

The cognitive structure was investigated using a word association technique. Each S was instructed to write as many other mathematical concepts as he/she could think of related to each of the 10 key concepts listed above. These data were then converted into a 10x10 similarity matrix by using the relatedness coefficient (RC) developed by Garskof and Houston (Psychological Review, 1963). This RC matrix was also subjected to Kruskal's scaling procedure.

Comparisons were made between each student's RC matrix and the digraph matrix. One method was a simple visual comparison of the pictorial models. The other method, which produced a score for each S, was to determine the Euclidean distance between matrices. This is obtained by squaring each difference between corresponding elements of the two matrices, summing the squares, taking the square root of the sum and then dividing by 90 (the number of off-diagonal elements in each matrix). A nonparametric analysis of variance was performed on the Euclidean distance at pretest, posttest, and retention points.

Achievement test scores were analyzed by a 2x3 (treatment by test occasion) analysis of variance. Scores from the attitude, achievement, and cognitive structure measures were intercorrelated using Kendall's Tau for the E group.

4. Findings

The data for the achievement test are shown in Table 1.

TABLE 1

Means and Standard Deviations of Scores on the Achievement Test for Each Treatment and Test Occasion

Treatment Group	Pretest	Posttest	Retention Test
Experimental	$\bar{X}$ = 3.65 σ = 2.45 n = 43	$\bar{X}$ = 15.54 σ = 5.74 n = 41	$\bar{X}$ = 16.21 σ = 6.32 n = 43
Control	$\bar{X}$ = 3.00 σ = 1.90 n = 42	$\bar{X}$ = 3.73 σ = 2.46 n = 40	$\bar{X}$ = 4.16 σ = 3.06 n = 43

Visual comparisons between the posttest and retention RC matrices and the content structure digraph suggested strong similarities. The nonparametric analysis of variance on Euclidean distances indicated ($p < .01$) that the cognitive structure of the Ss in the E group correspond more closely to the content structure than that of the C group.

The data for the correlations between attitude, achievement, and cognitive scores are shown in Table 2. Keep in mind that a perfect correlation between achievement and correspondence of structure scores would be -1.0 since the smaller the Euclidean distance, the closer the correspondence between content and cognitive structures.

TABLE 2

Rank Order Correlations (Tau) Between All Measures for Subjects in the Experimental Group

	Achievement			Correspondence of Structure		
	Pre	Post	Retention	Pre	Post	Retention
Attitude	243	186	158	-181	116	042
Achievement						
Pretest		236	285	-198	-132	-144
Posttest			724	051	024	104
Retention Test				-036	-080	027
Correspondence of Structure						
Pretest					097	375
Posttest						372

[As reported in Geeslin and Shavelson, 1975, p36]

5. Interpretations

The study indicated that the analysis of content structure using digraph theory could be applied to a mathematics curriculum. The achievement test data indicated that the instructional material was effective in teaching probability to eighth-grade students and that these students learned and retained the structure of probability. Learning structure and learning to solve problems in probability appear not to be highly related.

Suggestions for further study are to examine instructional variables that lead to a closer correspondence between content and cognitive structure than observed in the present study. Also suggested are studies examining the correspondence between cognitive structure, computational skills, and problem solving.

Critical Commentary

First, some minor points. I really cannot see what benefit to the study was the collection of attitudinal data. Could it be another example of data overkill? Also, I quarrel with the authors' assertion that the

Ss in the E group learned much probability. It seems to me that mean scores of 15.1 and 16.2 out of a possible 35 are not all that great, even though they clearly did much better than the Ss in the C group who had no instruction in probability.

Thus, it is not the least bit surprising that instruction in a subject leads to a closer correspondence between its content structure and the cognitive structure of those receiving instruction than the cognitive structure of those who did not receive the instruction. Unfortunately, the article did not present any data regarding just how close the correspondence was. The visual diagrams did not help because the posttest RC graph was a two-dimensional one (description of the dimensions not given) while the retention RC graph was three-dimensional (description of dimensions again not given).

When viewed as part of the apparently on-going effort by the second author to investigate this particular research methodology, the study is worthwhile. From the subject matter point of view, the value is less apparent. Using hindsight (always better then the other kind!), the concern for learning structure is much less prevalent now than it was 10 or even 5 years ago. To be of more value, I wish that the study had been done with high school students, the place where structure still seems to have relevance for some students.

James M. Moser
University of Wisconsin-Madison

THE CONCEPT OF BILATERAL SYMMETRY IN YOUNG CHILDREN. Genkins, Elaine F. In Children's Mathematical Concepts. Rosskopf, M. F. (Ed.) Teachers College Press, New York, pp5-44.

Also reported in:

A COMPARISON OF TWO METHODS OF TEACHING THE CONCEPT OF BILATERAL SYMMETRY TO YOUNG CHILDREN. Genkins, Elaine Francis, Columbia University, 1971. Dissertation Abstracts, v31A, pp1355-1356, September 1971.

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

1. Purpose

The stated purposes of the study were to:

(1) compare the effectiveness of paper-folding and mirror strategies in teaching the concept of bilateral symmetry to children in kindergarten and grade 2,

(2) compare the development of this concept in children at these grade levels, and

(3) investigate relative success of children in identifying exemplars and non-exemplars of the concept.

2. Rationale

This study was one of six which tried to interrelate mathematics and developmental psychology in illuminating learning and instructional problems. Because of the teaching techniques and the concept studied, the mathematical background, given in detailed elementary definitions, was drawn from the transformation geometry.

The psychological background drew on research from empirical perception work as well as relevant concept development literature. The perception literature was used particularly as a basis for the various orientations of the test and teaching items. The developmental framework supported the notion that children at two age levels might be expected to behave differently on bilateral symmetry tasks.

3. Research Design and Procedure

The objective of understanding of the concept of bilateral symmetry was operationalized in a 48-item test (actually two parallel versions were used to prevent confounding symmetry and orientation). This test contained 24 items on cards illustrating bilaterally symmetric figures (B types) (open, closed, curved and polygonal), 12 each oriented vertically and horizontally (V and H types). There were also 24 items with figures which were not bilaterally symmetric (N type). Of these, 12 were point symmetric (P types) and 12 were asymmetric (A types).

Two instructional sequences were designed using paper-folding and mirror work to develop the concept. For the paper group, 16 figures on tracing paper, using each of the types V, H, P, and A, formed the concrete basis of instruction. Under individual instruction, a child was first shown and then guided to see how paper-folding could determine if a figure was bilaterally symmetric or not, using B type, P type, and A type figures. The child then tried the remaining instructional figures on his own. The child had both to judge symmetry and to give reasons for his conclusions. The mirror sequence was similar but it involved mirror placement on the figures.

To test the validity of this limited approach to motion geometry for young children and the relative efficacy of the two instructional sequences, three groups were formed at each of two grade levels (kindergarten and grade 2). Children for these groups were nearly all of the kindergarten and grade 2 pupils in four independent schools in New York City and came mainly from upper middle class families. There were 91 and 94 children at each grade level respectively.

The teaching-testing sequence was administered individually by the investigator. Children were randomly assigned to either the Paper, Mirror, or Control group. The training period lasted less than 20 minutes in all cases, with grade 2 training proceeding slightly faster than that in kindergarten. Because of the nature of the test, the control group children required a brief expository period with one example of each of the V, H, P, and A types. The training period was followed immediately by testing for children in all groups. This lasted less than 25 minutes in all cases.

The experimenter determined for each child the number of correct responses for V, H, P, A, B, and N types as well as the total number out of 48. For each type and total score, a frequency distribution was tabulated for students from all four schools together at each grade level. (In comparing grade levels, kindergarten classification levels were used.) The chi square statistic was used to analyze data for various types using tables such as:

Kindergarten

	High	Low
Paper	X_{11}	X_{12}
Mirror	X_{21}	X_{22}

4. Findings

Kindergarten level:

The proportion of high scores in the Paper group was significantly higher than the Mirror and Control groups for P, A, and N type figures and exceeded the Control group on the total set.

Grade 2 level:

The proportion of high scores in the Mirror group significantly exceeded that in the Control group for all types of figures and the total set, while the Paper group exceeded the Control group for only the P, N, and total sets. The proportion of high scores in the Mirror group exceeded that in the Paper group for P, A, and N types.

Grade comparisons:

The proportion of high scores in the Control group (using kindergarten scores) at the grade 2 level significantly exceeded that of the kindergarteners for bilaterally symmetric figures and the total set. This result was repeated for the Paper group. For the Mirror group, the proportion of high responses of grade 2 children significantly exceeded that of the kindergarten children for all types of figures.

Figure type:

Control group: At both grade levels V figures were more easily discriminated than H figures and A figures were more easily discriminated than P figures. There were no differences between discriminating N and B figures.

Paper group: At the grade 2 level V figures were more easily discriminated than H figures, A more easily than P, and N more easily than B. At the kindergarten level, only the latter two differences obtained.

Mirror group: Discrimination differences appeared less pronounced for this group, although at the kindergarten level there were differences favoring A over P and N over B. At grade 2, only the A over P difference obtained.

5. Interpretations

(1) The paper folding method is appropriate for kindergarten children particularly in identifying non-exemplars of bilateral symmetry. The mirror method does not appear enjoyable or effective at this level.

(2) The mirror method appears appropriate at the grade 2 level in helping children understand all aspects of bilateral symmetry. Paper folding was enjoyable to this level of child and could be an adjunct instructional method.

(3) Bilateral symmetry is appropriate mathematically and psychologically as an introduction to motion geometry at grade 2 level and even for kindergarten children.

(4) There are interesting possible relationships between congruent parts of a figure and bilateral symmetry. Thus for children at this level, it may be that bilateral, translational, and rotational symmetries would be confounded.

Critical Commentary

(1) The test items and to an extent the instructional procedures are very well done. These should prove useful to other researchers and teachers of young children.

(2) The very interesting findings of this study are clouded by the analysis techniques. Because of the distributions, the median split seems to have a high chance of classification error. Even though group and not individual results were of concern, it seems some criterion score (e.g., 75%) or some other splitting mechanism would have been more meaningful.

(3) There are questions of internal validity (e.g., test reliability, replicability of instructional procedures) and external validity (e.g., sample type) which go unaddressed in this paper.

(4) I wish the author had commented more on the developmental aspect of the study. For example, to what extend did the Control group at grade 2 do better than the treatment groups at kindergarten level? If such differences obtained, is the earlier instruction useful?

(5) The confounding effects of point symmetry (and possibly congruency in general) seems to cast doubt on whether bilateral symmetry is indeed the best entry concept into motion geometry.

Thomas E. Kieren
University of Alberta

COGNITIVE FACTORS AND THEIR INTERACTION WITH INSTRUCTIONAL MODE. Hancock, Robert R. Journal for Research in Mathematics Education, v6 n1, pp37-50, January 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Otto Bassler, George Peabody College for Teachers.

1. Purpose

The purpose of this study was to investigate (1) the main effects of mode of presentation (verbal or figural) and sex, and (2) the interactions between personological variables and mode of presentation as ninth-grade students are learning concepts and principles associated with a contrived mathematical relation.

2. Rationale

One way to individualize instruction is to make differential assignments on the basis of some aptitude-variable score. To do this effectively individual learning preferences for content or mode of instruction must be investigated. One attack on such a problem is through aptitude-treatment interaction which seeks to accommodate the cognitive preferences of different people.

Previous research with college students indicated that (1) a group that studied a verbal program did better on three of the four criterion measures than did a group that studied a figural program; (2) there was no evidence of a disordinal interaction between personological variables and mode of instruction; and (3) female subjects scored significantly higher on the subtest that measured achievement at the highest cognitive level. These findings may be unique to college-level students and different results may be obtained with a sample from a younger age level.

3. Research Design and Procedure

Subjects were 119 ninth-grade students enrolled in either general mathematics, regular algebra, or accelerated algebra. These students were assigned randomly by sex to one of two treatment groups where, for three 43-minute class periods, they studied one of two programmed instructional units. The content of the two programs dealt with selected concepts and principles associated with a linear order relation. One program used a verbal mode of presentation while the other presented the material by means of a figural mode. On the day following completion of the instructional programs, subjects were given a 33-item multiple-choice test. This instrument measured achievement at two different cognitive levels as well as assessing total achievement. Four weeks later a similar test of retention was administered.

The independent variables in this study consisted of IQ scores on the Otis-Lennon Mental Ability Test - Form J; Standard Achievement Scores in Mathematics from the Iowa Test of Mental Ability; and scores on a battery of nine tests developed by Guilford et al., purporting to measure cognitive ability dealing with figural, semantic and symbolic content. In all there were 14 independent variables.

Main effects for each of the six dependent variables (post-learning and retention subscores and totals) were tested using _t_-tests. These tests compared the mean scores of (1) the verbal and figural groups, (2) the male and female groups, (3) males who studied verbal material and males who studied figural material, and (4) females who studied verbal material and females who studied figural material. Interaction effects of independent variables with treatments were studied using regression analysis. Simple linear regression equations were determined and tested for goodness of fit. If both regression lines had a significant linear component, analysis was continued. Next, if the regression lines intersected within the range of the observed scores on the independent variable, the difference between the regression coefficients was tested for significance. If this difference was significant, then it was concluded that for the dependent variable a disordinal interaction existed between the independent variable and the two treatment modes. Significance levels (α) were selected at a higher level than usual in an effort to secure statistical power in the .7 to .8 range.

4. Findings

The significant findings were: on the retention test, cognitive level I, male subjects who studied the verbal program did significantly better than male subjects who studied the figural program; female subjects scored higher than male subjects on all measures except the retention test, cognitive level II. All other comparisons were non-significant.

There were 17 significant disordinal interactions. Of these interactions thirteen occurred on tests of retention and four on tests of immediate learning. Another breakdown of these interactions indicates that one involved IQ, eleven involved semantic factors, four involved symbolic factors, and one involved a figural factor. Finally, of those interactions that were related to sex differences, eight of the nine involved male subjects.

5. Interpretations

The lack of significant findings due to mode of presentation suggests that perhaps these ninth-graders had not yet developed a cognitive preference—or at least an adaptability to material that is verbally oriented. The overwhelming superiority of female subjects was unexpected and, hence, continues to be a variable of interest.

The preponderance of disordinal interactions on tests of retention was interpreted as possible support for the statement that students who study an instructional program consistent with their cognitive preference are better able to assimilate the material into their cognitive structure than those who do not. It also has the pedagogical implications that instructional modes which may be adequate for initial learning may not be adequate for retention.

In terms of variables that may be productive sources of independent variables in future aptitude treatment interaction studies, the results of this investigation would suggest memory variables and semantic subcategory variables.

Critical Commentary

As the author points out, ATI research is in the formative stages and any interactions obtained in this study should be interpreted with caution. With this in mind the study provides some interesting and possibly fruitful variables to investigate in future research.

The study was well designed and carefully conducted. There are, however, some questions about the analysis. It would have been preferable to conduct ANOVAs rather than t-tests to test for significance of the main effects of mode of instruction and sex. This analysis would have permitted investigation of interaction effects between these variables. In the absence of any significant interaction, there is then no need to test the simple effects. It is also debatable if tests of significance should be conducted both on subscales measuring two different cognitive levels and on the total test score. These techniques can perhaps be rationalized since this is a study searching for experimental hypotheses rather than one attempting to provide definitive results.

The report of this research would have been strengthened by an example of the figural and verbal modes of presentation. Presentation of both modes via programmed instruction may have resulted in the comparison of one level of verbal presentation with another. This was impossible to determine through reading the report. It would also have been useful if the cognitive levels used in the dependent variable had been defined.

Studies such as this, which repeat with modification previously conducted studies, are needed in mathematics education. Comparison of the results of this study using ninth-grade students with those of the college-level students of a previous study provide valuable information about learning differences as well as give rise to speculation about the effectiveness of the treatments with other age levels.

Otto Bassler
George Peabody College for Teachers

DOES LIVING WITH A FEMALE HEAD OF HOUSEHOLD AFFECT THE ARITHMETIC ACHIEVEMENT OF BLACK FOURTH GRADE PUPILS? Jantz, Richard K.; Sciara, Frank J. Psychology in the Schools, v12, pp468-472, October 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Edward M. Carroll, New York University.

1. Purpose

The purpose of this study was "...to examine the effects of living with a male or female head-of-household upon the arithmetic achievement test scores of Black fourth-grade pupils" (1975, p. 469). A subsidiary purpose was to determine the influence of the factors of sex and intelligence and their interactions with the factor of head-of-household upon arithmetic achievement.

2. Rationale

The authors cited government census figures for 1972 which indicated that an increasing number of households were headed by females and that approximately 37% of the households of Black families in large urban centers were headed by females. Some research on the matricentric family system and the factors as related to Black families revealed that this system made it extremely difficult for the low-income Black father to assume his parental role. Further research investigations revealed that father-absent children performed more poorly than father-present children on some intelligence tests, academic tests, and overall achievement. The paucity of research on the effect of father-absence upon the arithmetic performance of Black elementary school pupils led to the null hypothesis:

> That there is no effect of head-of-household, sex, intelligence, and their interactions on the arithmetic scores of Black fourth grade pupils.

3. Research Design and Procedure

The population for the study included all Black fourth-grade pupils enrolled in eight model cities schools of a large midwestern school corporation during the school years 1970-71 and 1971-72. All data relating to sex, intelligence test scores, arithmetic achievement test scores, and family status (whether pupil lived with a female head-of-household or a male head-of-household) were collected from the permanent record cards of the pupils. Approximately 20 percent of the population was rejected because of incomplete entries on the permanent record cards, resulting in a sample of 1,073 pupils included in the study. Of these, 300 pupils were living with a female head-of-household and 773 pupils with a male head-of-household. The Metropolitan Achievement Test was used to measure arithmetic performance and the Lorge Thorndike Intelligence Test was utilized as the measure of intelligence. These tests were administered in the Spring term of the school years 1970-71 and 1971-72 as a part of the regularly scheduled test program by the testing bureau of the school corporation.

The data were analyzed employing a 2x2x3 factorial design with 3 factors for intelligence utilizing a multivariate program. The scores of pupils were divided according to the criteria that IQ test scores were: higher than 100, from 84 to 100, and less than 84. The Newman-Keuls procedure for unequal Ns was used to test for significance within the interactions of sex and head-of-household, and intelligence and head-of-household, once an overall significance had been determined by ANOVA.

3. Findings

a. There were no significant differences in mean arithmetic scores between the male and female Black fourth-grade pupils.

b. There were significant differences in mean arithmetic scores favoring pupils with higher IQ scores over those with lower IQ scores.

c. There were significant differences in mean arithmetic scores favoring those living with male heads-of-households over those living with female heads-of-households.

d. There were significant differences in mean arithmetic scores favoring female pupils living with male heads-of-households over the other interactions of sex and heads-of-household.

e. There were significant differences in mean arithmetic scores favoring those pupils with IQ scores greater than 100 living with male heads-of-households over all other interactions of the factors of intelligence and heads-of-households.

5. Interpretations

The authors acknowledged that there is a variety of factors which may explain the difference in arithmetic performance of those children living with a male head-of-household over those living with a female head-of-household. For example, they cited results of previously reported research on father-absence indicating:

a) that children from fatherless homes had lower IQ scores by the time they reached fifth grade than did children from intact homes (Deutsch, 1965).

b) that lower IQ scores of children in father-absent homes might be attributed to additional stress and loss of male identity (Lessing, Zagorow, and Nelson 1970).

c) that socio-cultural backgrounds, peer interactions, and length of time of father-absence are important to pupils' performance (Biller, 1971).

d) that when a father leaves and the type of absence are both important to cognitive development (Biller and Bahm, 1971).

e) that the effect of the father's absence varies with the age of the child (Hartnagel, 1970).

The authors concluded that the effect of living with a male or female head-of-household upon arithmetic performance is complex and should not be viewed in simple terms of cause-and-effect relationships. However, the effect might be considered an indication of potential difficulty for some students. With an increasing number of Black students living with female heads-of-household, more male models might be needed in the elementary schools. The schools cannot replace the family, but they can develop a set of experiences not now available at home. Additional research is recommended to identify some of these effective school experiences.

Critical Commentary

The abstractor agrees with the conclusions reached by the authors "that the effect of living with a male or female head-of-household upon arithmetic performance is complex and should not be viewed in simple terms of cause-and-effect relationships" and that there is a strong indication that "more male (teacher) models might be needed in the elementary schools." Also, additional research is certainly recommended to identify some effective school experiences with children when more male teachers are on the elementary-school staff.

Since this was a post hoc study, the data related to socioeconomic educational background of the Ss may not have been available to the investigators; however, it would have been helpful to the reader. It has been demonstrated that (1) the more educated the parent, the more educated the child; (2) a child's academic performance is likely to be related to his parents' social class; and (3) children whose parents are more educated begin their school careers with higher achievement scores than do their less fortunate peers. These data would have assisted the reader in understanding the findings of this study.

Other questions raised by the report were:

1. What were the age range of the Ss in the sample? Were any of the pupils repeaters, etc.?

2. What is a mean score of 3.56? Is it a grade level or a stanine? Does this score represent only Test 1 Arithmetic Computation, Test 2 Arithmetic Problem Solving and Concepts, or a combined score for both tests in the MAT Elementary Arithmetic Test Series for Grades 3 and 4? What were the exact form and title of the tests?

3. What criteria was used to factor IQ scores into the three regions below 84, from 84 to 100, and 100 plus? When the generally accepted division of below 90, from 90 to 110,

and 110 and above are not used to denote "below average", "average" and "above average", respectively, very few cross comparisons can be made with other studies of this nature.

4. If the factors of heads-of-households such as mothers-only, fathers and mothers, and fathers-only were used in the design, would the findings be affected?

Edward M. Carroll
New York University

References

Biller, H. B. The Mother-Child Relationship and the Father-Absent Boy's Personality Development. Merrill-Palmer Quarterly of Behavior and Development, v17, pp227-241, 1971.

Biller, H. B.; Bahm, R. M. Father Absence, Perceived Maternal Behavior, and Masculinity of Self-Concept Among Junior High School Boys. Developmental Psychology, v4, pp178-181, 1971.

Deutsch, M. The Role of Social Class in Language Development and Cognition. American Journal of Orthopsychiatry, v35, pp78-88, 1965.

Hartnagel, T. F. Father Absence and Self Conception Among Lower Class White and Negro Boys. Social Problems, v18, pp152-163, 1970.

Lessing, E. E.; Zagorow, S. W.; Nelson, D. WISC Subtest and IQ Score Correlates of Father Absence. Journal of Genetic Psychology, v117, pp181-195, 1970.

INTELLECTUAL DEVELOPMENT BEYOND ELEMENTARY SCHOOL I: DEDUCTIVE LOGIC. Karplus, Elizabeth F.; Karplus, Robert. School Science and Mathematics, v70, pp398-406, May 1970.

INTELLECTUAL DEVELOPMENT BEYOND ELEMENTARY SCHOOL II: RATIO, A SURVEY. Karplus, Robert; Peterson, Rita W. School Science and Mathematics, v70, pp813-820, December 1970.

INTELLECTUAL DEVELOPMENT BEYOND ELEMENTARY SCHOOL III--RATIO: A LONGITUDINAL STUDY. Karplus, Robert; Karplus, Elizabeth F. School Science and Mathematics, v72, pp735-742, November 1972.

INTELLECTUAL DEVELOPMENT BEYOND ELEMENTARY SCHOOL IV: RATIO, THE INFLUENCE OF COGNITIVE STYLE. Karplus, Elizabeth F.; Karplus, Robert; Wollman, Warren. School Science and Mathematics, v74, pp476-482, October 1974.

INTELLECTUAL DEVELOPMENT BEYOND ELEMENTARY SCHOOL V: USING RATIO IN DIFFERING TASKS. Wollman, Warren; Karplus, Robert. School Science and Mathematics, v74, pp593-613, November 1974.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by John A. Easley, Jr. and Kenneth J. Travers, Committee on Culture and Cognition, University of Illinois-Urbana.

1. Purpose

This series of five studies, by Karplus and his associates at the University of California-Berkeley, investigated the conceptual development of children in the upper elementary grades and high school, and (in one study) of adults. Using the group test, the experimenters classified responses according to a hierarchical schema, with categories ranging from concrete to abstract levels of thought. The studies were intended to reveal information about how well certain concepts, primarily that of ratio, function at various ages and to explore conditions under which learning of these concepts might be facilitated. Implications of the findings for mathematics and science curricula were also discussed.

2. Rationale

The concepts of ratio and proportion are essential to an understanding of quantitative relations in science, as for example in many mathematical models. Hence, it is important that educators and researchers learn methods for determining the extent to which concepts, such as ratio, have been attained. Furthermore, curriculum developers ought to have at their disposal methods for identifying areas of need among students at various age and grade levels. The studies also sought to shed light on what tasks or techniques appear to promote concept acquisition and what learnings have accrued from various curricula. The possibility of cross-national comparisons of concept learning was also suggested. The use of group testing techniques (as opposed to clinical interviews) for assessing knowledge would make possible large-scale investigations of the questions addressed by Karplus' research.

The research appeared to have its point of departure in that of Inhelder and Piaget (1958), which defines a stage of formal operations and links proportional thinking to this stage. [However, these studies of intellectual development also relate to work by Lunzer and Pumfrey (1966), Lovell and Butterworth (1966), and Dienes (1973). A review of these and other studies porporting to deal with formal operations was presented by Lunzer (to appear) at the 1973 meeting of the Jean Piaget Society (see Critical Commentary below).]

3. Research Design and Procedure

The distinguishing features of the five studies are summarized in Table 1. Each study is subsequently referred to by the Roman numeral appearing in the title (e.g., "III" refers to Karplus and Karplus, 1972).

The five studies had similar designs in that each consisted of the administration of a task, an analysis and classification of Ss' responses to the task, a tallying of the responses in the various categories, and a discussion of the resulting data.

3a. Brief Description of Tasks Used

Island Puzzle: This puzzle involved four islands in an ocean and the possibilities of travel between them. The task was designed to assess abstract reasoning ability. Individual responses were written.

Paper Clips, Form A: Ss each were given a sheet of paper on which a stick figure 7 9/16 inches tall was drawn, together with a chain of from seven to ten No. 1 "Gem" paper clips. Ss were shown a display chart with the same figure as they had on their sheets, with a scaled up-version on the back side of the chart. The experimenter had a chain of eight jumbo paper clips. Ss were shown that the small figure was four "biggies" (jumbo paper clips) tall and the large figure was six "biggies" tall. Ss were to measure the small figure on their sheets with "smallies" (small paper clips) and to predict the tall figure's height in "smallies." Ss wrote their responses.

Paper Clips, Form B: This was similar to Form A, with the important difference that Ss did not ever see the tall figure. Ss were told that the small figure was four buttons tall and the large figure was six buttons tall. They were to measure the small figure in paper clips and predict the height of the large figure in paper clips.

Candy: Ss were given written information concerning a person who had many bags of two kinds of candy, each bag containing a certain number of reds and a different certain number of yellows. All bags were alike. Ss were given written information relevant to predicting the number of candies taken from the bags. One task (a) involved application of the ratio 5:3; the other (b) involved the ratio 2:1.

TABLE 1

Tasks Used in the Five Studies

Study	I	II	III	IV	V
Name of task and required completion time	Island Puzzle (10-15 min.)	Paper Clips Form A (15 min.)	Paper Clips Form A (15 min.)	Paper Clips Form B (15 min.)	*Paper Clips Form B (15 min.) *Candy (15 min.) **Ruler (2-4 min.) **Pulley (4-8 min.) *Workbook (20 min.) geometry arithmetic
How administered	Group	Group	Group	Group	*Group **Individual
Subjects used	Six groups (N = 449) from grade 5 through adult (teachers of science)	Six groups (N = 727) from grades 4-12, urban and suburban	Grade 6, 8, and 11 suburban students tested in study II at grade 4, 6, and 9 (N = 155) plus new group of grade 8 students (N = 141)	Grades 4-9 urban and suburban (N = 616)	Grades 7-8 students suburban (N = 450)

Ruler: An individually administered task was used, involving an unmarked rod and a ruler marked in both inches and centimeters. Ss were shown that a displacement of 2 inches on the rod was equivalent to a displacement of about 5 cm. Ss were then asked how many centimeters would be equivalent to a displacement of 8 inches.

Pulley: Individuals were interviewed using a mechanism which involved two pulleys (diameter ratio 3:2) fixed to same shaft, a string attached to each pulley, and a meter stick. S and experimenter worked together on the mechanism and S was shown that a displacement of 10 cm on his or her string corresponded to a displacement of 15 cm on the experimenter's string. S was then asked how far the experimenter's string would move when S's string was displaced 6 cm and to explain why.

Workbook: Two tasks were designed to be similar to those found in textbooks. One geometrical task was intended to assess the S's ability to recognize a fraction of a whole and to represent this faction pictorially and numerically. The other task, numerical, required the S to apply proportional reasoning with no circumstantial clues.

Categories Used to Analyze Tasks

Study I:

N (no explanation)
I (prelogical)
II_a (transition to concrete)
II_b (concrete)
III_a (transition to abstract logic)
III_b (abstract logic)

Studies II-V:

N (no explanation or statement given)
I (intuition): estimates, guesses without reference to data
IC (intuitive computation): data used haphazardly or illogically
A (addition): uses difference rather than ratio
S (scaling): uses change of scale, not related to scale inherent in data
AS (addition and scaling): use of difference and scaling
IP (incomplete proportion): uses one ratio only
P (proportional reasoning: uses properties of proportionality
3 subcategories of P:
PC (proportion, concrete)
AP (addition and proportion)
R (application of ratio)

4. Findings

All five studies revealed a tendency for the median frequency of tasks to move from the lower level categories (such as I and S) to the higher categories (such as P) as the grade level of the respondents increased. Wide fluctuations in performance within grade levels were found in II. In one urban class, 90% of the responses were placed in categories N, I, or IC, while in another class, categories AS and P were more numerous than in any one suburban class. In III, a longitudinal design produced data suggesting that the categories may be representative of developmental levels. Of the 153 students involved in the two-year study, 28% moved into P or AS, while only one student moved out from P to A. At the lower levels, 7% moved from other categories into I or IC, while during the same two-year period, 65% moved out. It was also found, however, that 40% of the students remained in their categories for both testings.

Table II, which combines the categories into three levels, summarizes the classification of the Ss in study III over 2 years (1969 and 1971).

TABLE II

Matrix Comparing Students in 1969 and 1971 by Levels
(Number of Students)

	1971			
1969	Level I	Level II	Level III	Total
Level I = I + IC	19	22	10	51
Level II = S + A	7	45	26	78
Level III = AS + P	0	1	23	24
Total	26	68	59	153

(From Karplus and Karplus, 1972, p. 739)

Study IV used a new form of the Paper Clips Task. It was more abstract than Form A, in that the S did not ever see the figure whose height was to be predicted using a proportion. A dramatic reduction was found in the number of Ss' (scaling) responses to Form A (30%) when Form B (4%) was used, with a corresponding increase in categories IC, IP, A, and P. The latter categories require, the researchers assert, conceptual processing of the data by the Ss (IV, p. 480).

The variety of tasks used in V proved to be interesting. Results on Form B of the Paper Clips Task were similar to those obtained by the 8th graders in IV. Sex differences in responses were not appreciable. It was found that the value of the ratio used is a factor to be taken into account.

In the Candy Task, application of an integral ratio (2:1) was interpreted as not indicative of formal reasoning, since use of the ratio did not correlate with proportional reasoning in a more complicated task. (V, pp. 597-598). The Ruler Task was found to be easy for the junior high school students, since 87% responded with proportional reasoning. The Pulley Task was more difficult and the geometrical and numerical items were most abstract (V, p. 604). It was noted that a perfect score on the numerical items was a good predictor of success on the geometric items, while the reverse did not hold.

5. Interpretations

The researchers, overall, were disturbed by the implications of their findings. In study I, intellectual development, as assessed by their taxonomy, reached a "disappointingly low level" in the high school age group and did not progress much further (I, p. 403). In study II, it was found that successful proportional reasoning was not reached until the last years in high school. This concern was reiterated in III, where evidence was found that many students did not advance to more abstract categories of thought during the intervening two years of that longitudinal study. Another disturbing implication, prompted by the data of studies II and III, was evidence of apparent obstacles to learning which may be inadvertently set up by "mathematics courses, by teachers, and by the children's cultural environment" (II, p. 817). Of particular notice was the dramatic contrast between the responses of urban and suburban 11th and 12th grade students. It was found that 80% of the suburban students, but only 9% of the urban students, were classified at the highest level P.

The findings of study IV placed emphasis on the context in which the problem was presented. Since Form B of the Paper Clips Task was more abstract than Form A, the researchers concluded that Form B compelled the students to make use of the data. Indeed, they concluded, Categories I and S under Form A (studies II and III) may reveal an attitude toward handling of the data rather than the respondent's cognitive level of competence (IV, p. 480).

The variety of tasks used in V led the researchers to generalize about the influence of a task upon the S's response. Tasks tending toward concreteness (Ruler, Paper Clips) led to more correct responses than did the abstract tasks. The lack of applicability of proportional reasoning to physical relationships raised questions about the appropriateness of many instructional strategies, particularly at the junior high school level, where only about 15% of the subjects were found to have reached the highest level. The researchers speculated that one source of the problem may be that ratios are introduced as fractions and proportions as equivalent fractions. "Curricula make little effort to interpret ratio, proportion, and the related division process in terms of ... corespondences of measurements. In this use of division, the concept of remainder has no place" (V, p. 610).

Critical Commentary

It should be clearly noted that Karplus and his colleagues are investigating a different problem from that studied by Piaget. It is easy to lump these together with all kinds of studies involving tasks that can be called "formal reasoning," by the ordinary meaning of that term. Lunzer (1973), for example, made no notice of the peculiar use Piaget and the Geneva School make of that term, but made a broad survey of problem-solving studies in which logic or other formal rules or procedures were used and made it the basis for a critique of Piaget's theory of the stage of formal operations. Karplus and his collaborators are careful not to identify what they call proportional reasoning in studies II-V and what they call "abstract logic" in study I, with Piaget's Formal Operations. At least, they are quite open to the possibility that these may turn out to be different things. The situation is reminiscent of one involving the notion of relative motion.

When Piaget argued that understanding relative motion required the INRC group, and hence was a formal level task, Easley (1964) attempted to provide an operational definition of the INRC group in the snail board problem Piaget had used in his studies of relative motion. However, Piaget made it quite clear that young children who could perform in a way that satisfied that operational definition would not thereby automatically be credited with having achieved the formal level (personal communication). Here in the studies under review, we have evidence from high school seniors (and even many adults) who do not employ "abstract logic" or "proportional reasoning" in these problems, but we are not entitled to infer that they have not achieved the stage of formal operations, as Piaget defines it, even though Inhelder and Piaget state that proportional reasoning is only attained at the stage of formal operations.

Piaget's theory, briefly, is that four cognitive structures called logical operations (identity, negation, reciprocity, and correlation), usually between the ages of 10 and 14, unite to form a single structure, the INRC group. This greatly increases the power of the individual over the intuitive feelings of physical quantity operations. Thus, Piaget writes:

> It seems evident, in the instance of weight, that the difficulties of dynamic interpretation presented by this notion play a big role in the delay of its operational structuration, because of the contradictions that must be overcome between the demands of structuration and the diversity of objective causal situations. The same applies to volume, the delayed logicalization of which seems to be linked to geometric problems of internal continuum...going beyond the realm of concrete operation (1974, p. 3).

Second, the group composition of the individual operations now permits operations on operations, which supports the development of proportionality. Thus, Piaget writes:

> But we have also seen how the subject succeeds (first) in constructing by reflexive abstraction his multiplicative operations as additive operations at the second power, then his structures of proportionality by equalization of relationships (therefore, again, by relationships of relationships or relationships at the second power)...(1974, p. 67).

To discover whether a subject has attained the stage of formal operations for proportionality operations in a given context, Piaget and his colleagues employ the clinical interview with probing (at times, prompting the subject first in one direction and then in another to separate any efforts to please the interviewer from what the subject genuinely believes) and continuing the probing until satisfied that the most advanced level of which the subject is capable has been demonstrated. They thereby incur the criticism of behaviorists that they are leading and prompting their subjects. However, Piaget argued (1929) that without such methods they cannot discover intellectual structures.

On the other hand, Karplus and his colleagues are interested in the primary or spontaneous level of thought (defined in terms of their categories) that subjects employ when solving a paper-and-pencil test. Although they state that they have used interviews to check the levels of performance they get on paper-and-pencil tests, it is clear that they have avoided the clinical interview with its probing and prompting. This difference in purpose and procedure explains in part the differences in age distribution found between the two groups of studies. A second major contribution to these differences, which the research to date cannot isolate from the first, is the phenomenon Piaget calls decalage (separation or displacement). This refers to the delay in development from the first case in which a subject can use a given form of thought (in this case, formal operations) to a more difficult application that is different in content. While Piaget's theory of decalage is not very well developed, the essay quoted above (Piaget, 1974) indicates that the developmental relationships between various kinesthetic structures and logical ones is complex. Another difference is that Karplus et al. (study V) employ a concept of types of reasoning which depends on the external static form of arguments. This contrasts sharply with Piaget's interest in internal dynamic processes (see Easley, 1973).

The gulf between the narrow conceptual and methodological traditions that characterize these two groups of studies has not yet been bridged. Both groups have legitimate research interests and both have practical applications to education. It is interesting, for example, that the Elementary Science Study and the Science Curriculum Improvement Study, the latter of which is directed by Karplus, have generally been interpreted as attempting to challenge the highest intellectual competence of children and not merely tap their typical performance in a situation. It might seem then that Piaget's approach to research on cognitive development could be more relevant to Karplus' elementary-school project than Karplus' own research approach is, or perhaps that, in preparation for the development of new secondary-school curricula, a more external form is thought to be required. Perhaps there is another reason we do not understand for the maintenance of this separation.

The issues raised by this series of studies are important ones. Other mathematics educators will surely share the investigators' concerns over the appropriateness of the current curriculum (and its typical method of implementation) for developing applications of mathematics to real-world problems. If it can be agreed that the mathematics curriculum should indeed promote creative use of numbers to describe physical objects (as is required in the sciences), then a research base for curriculum development in integrated mathematics and science programs is an important emphasis for future study. Also, despite the difference between Piagetian clinical interviews and Karplus' paper-and-pencil tasks, both offer alternatives to the conventional fascination with standardized (group-normed) tests of aptitude or achievement.

Karplus and his associates also raise the matter of cross-national comparisons of educational programs through an examination of intellectual development. Such a topic would appear to lend itself readily to consideration for inclusion in the proposed second round of school-subject surveys by the International Association for the Evaluation of Educational Achievement. Information is needed on the relationship between different curricular emphases found in various countries and the attendant differences in conceptual development. Might it be the case that some countries are indeed much more successful than others in promoting the attainment of proportional reasoning for the majority of students during the junior high school or early high school years? In terms of Piagetian methods, recent suggestions that the rate of intellectual development might be uniform across cultures (Kamara and Easley, 1967) is not well supported beyond the stage of concrete operations. Further research on this front is needed also.

John A. Easley, Jr.
Kenneth J. Travers
Committee on Culture and Cognition
University of Illinois-Urbana

References

Dienes, Z. Mathematics Through the Senses, Games, Dance, and Art. Slough, Bucks, England: National Foundation for Educational Research Publishing Company, Ltd., 1973.

Easley, J. A., Jr. Comments on the INRC Group. Journal of Research in Science Teaching, v2, pp233-235, 1964.

Easley, J. A., Jr. Symbol Manipulation Re-examined--An Approach to Bridging a Chasm. Paper presented at the Third Annual Meeting of the Jean Piaget Society, May 22, 1973.

Inhelder, B.; Piaget, J. The Growth of Logical Thinking from Childhood to Adolescence. New York: Basic Books, Inc. 1958.

Kamara, A. I.; Easley, J. A., Jr. Is the Rate of Cognitive Development Uniform Across Cultures?--A Methodological Critique. In Piagetian Psychology: A Cross-Cultural Perspective. Dasen, P. (Ed.). New York: Gardner Press, in press.

Lunzer, E. A. Problems of Formal Reasoning in Test Situations. In European Research in Cognitive Development. Mussen, P. (Ed.) Monograph of the Society for Research in Cognitive Development, v30, pp19-41, 1965.

Lunzer, E. A. Formal Reasoning: A Re-appraisal. Address given to the Third Annual Meeting of the Jean Piaget Society, May 22, 1973.

Lunzer, E. A.; Pumfrey, P. D. Understanding Proportionality. Mathematics Teaching, v34, pp7-12, 1966.

Lovell, K.; Butterworth, I. B. Abilities Underlying the Understanding of Proportionality. Mathematics Teaching, v35, pp5-9, 1966.

Piaget, J. The Child's Conception of the World. London: Routledge and Kegan Paul, 1929.

Piaget, J. Causality and Operations. In Understanding Causality. Piaget, J.; Garcia, R. (Eds.). New York: W. W. Morton and Company, 1974.

CALCULATORS IN SOME FIFTH-GRADE CLASSROOMS: A PRELIMINARY LOOK. Schafer, Pauline; Bell, Max S.; Crown, Warren D. Elementary School Journal, v76, pp27-31, October 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by John E. Tarr and Jack D. Wilkinson, University of Northern Iowa.

1. Purpose

The primary hypothesis investigated was that pupils who had briefly explored calculators would do better on an arithmetic achievement test than pupils who had not explored calculators.

2. Rationale

The study was intended to begin exploration in the use of electronic calculators in the classrooms, to make informal classroom observations, and to generate some hypotheses (especially on achievement testing). The writers report that thus far there are few research-based answers to questions related to classroom use of calculators.

3. Research Design and Procedure.

The study compared arithmetic achievement scores of two groups of pupils--an experimental group and a control group. The inquiry was conducted in April 1974, in five fifth-grade classrooms of the University of Chicago Laboratory School, where the pupils are predominantly from middle- and upper-middle-class families and generally score above the national norms on standardized tests. Three classes (69 pupils) served as the experimental group; two classes (46 pupils) served as the control group.

Pupils in the experimental group were given calculators to explore for fifty minutes on each of two days. They were given problems to do and were encouraged to ask questions about the calculators.

The Mathematics Computation Test (distributed by the Educational Testing Service) was used as both a pretest and a posttest. Each item was categorized as either a calculator or a non-calculator example. Thus, an administration of the test yielded three scores: (1) the whole score; i.e., the number of examples correct on the entire test; (2) the calculator score; i.e., the number of examples correct that required either the use of some additional information or a two- or three-step computation; (3). the non-calculator score; i.e., the examples not scored for the calculator score.

Form A of the test was given in February 1974, and used as the pretest. Form B of the test was given as a posttest about a week after the experimental group had its two-day calculator experience in April 1974. The pretests showed no significant differences between the control and experimental groups on any of the three raw scores. The posttest data were summarized and a t-test used to examine the differences in the means.

4. Findings

The posttest results for calculator and non-calculator examples are given in the following table.

Type of Examples	Group	Number of Pupils	Mean Score	Standard Deviation	t
Calculator	Experimental	69	22.91	1.78	4.204*
	Control	46	20.96	3.20	
Non-calculator	Experimental	69	17.71	5.61	1.269+
	Control	46	18.98	1.27	

*Significant at the .001 level
+Not significant

There was no difference reported between the groups on the pretest whole score.

The partial scores for the groups show a highly significant difference in favor of the experimental group in the performance on calculator examples. On non-calculator examples the performance of the experimental group was not statistically different from that of the control group.

5. Interpretations

Pupils using calculators answered more of the calculator examples than they would have without them. The use of calculators may help on examples where calculation is the main issue.

There may be some loss from trying to use calculators when they are not appropriate. The performance of the experimental group was poorer than that of the control group. Perhaps the pupils in the experimental group depended too much on the calculators.

Pupils made few attempts to estimate answers, even to the proper order of magnitude. This skill is almost essential if calculators are to be used effectively.

Curiosity ran high and interest in learning additional mathematical content was keen. In the classes that were introduced to calculators motivation and interest were boosted substantially and pupils generated many questions that could easily have been exploited to begin a series of explorations about mathematics.

Critical Commentary

There is little question as to the need for action and developmental research dealing with the role and use of the hand-held calculator in teaching mathematics. This article provides both direction for future researchers and questions for current practitioners.

The fact that pretest data were not reported created some question in the way the data dealing with non-calculator examples were interpreted.

The writers infer that this non-significant difference may be interpreted to mean that "perhaps the pupils in the experimental group depend too much on calculators." Later they state that, "there may be some loss from trying to use calculators when they are not appropriate." How reasonable is it to make these inferences when the treatment and control groups may have varied that much on the pretest?

The nature of the treatment was not clear. The writers state that the experimental group was given calculators to explore and that "children were given problems to do and were encouraged to ask questions about the machines." Some additional information regarding instruction would have been helpful. Would the nature of the instruction, type of examples, and problems be the most important variable in any study of this sort?

The study suggests implications for further research. One slight variation of the study would be to consider four groups: (1) Calculator experience; pencil and paper on test; (2) Calculator experience; calculator used on test; (3) No calculator experience; pencil and paper on test; (4) No calculator experience; calculator used on test. This study considered groups (2) and (3); another study could consider all four. Other questions for further research include: If the use of calculators were more than simply a two-day exploratory experience, but rather a one-week, structured experience using materials written specifically for the calculator, would even greater differences in scores be found? If materials were used with calculators in which common difficulties were encountered and pupils were sensitized to these difficulties, would they then perform better on the non-calculator examples? Are the high-interest aspects of calculator-usage lasting effects or short-lived effects? Perhaps of greatest importance is the question, do pupils with calculators better learn mathematical concepts and skill?

John E. Tarr
Jack D. Wilkinson
University of Northern Iowa

SHIFTS IN READING READINESS PROFILES DURING THE PAST DECADE. Scott, Ralph. Journal of Genetic Psychology, v126, pp269-273, June 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Thomas R. Post, University of Minnesota.

1. Purpose

To determine whether there have been, within the last decade, changes in the readiness profiles of American children as measured by the tests of Primary Mental Abilities: PMA: K-1. "This instrument yields not only a total I.Q. but also four sub-test scores which provide a general but educationally useful readiness profile."

2. Rationale

"It has been known for some time that social changes may produce a shift in the cognitive patterns of children. A number of selected longitudinal studies have indicated that the intellectual structures of children in Western societies have been changing during the past several decades. For teachers, the important question is not so much whether the pattern of readiness skills has changed, but the implications this may hold with respect to teaching strategies.

"Since the 1962 revisions of the PMA: K-1, substantial changes have occurred in American society which might substantially modify the readiness profiles of children entering first grade. The average child now watches TV about 50% more than the time he is in school or an annual average of 1,340 hours. Observation suggests that the typical child today spends less time in motor-expressive types of behavior (such as sports, fishing, helping parents in the kitchen, barn, or garage) than did his pre-TV counterpart. Voluntary reading is another learning-related activity which involves individual decision making, and this too seems to be on the wane: from 1959 to 1971 the number of children who voluntarily read at least one hour per week declined from 82% to 67%.

"It is entirely possible that changes in children's free time activities have altered the readiness profiles of first grade youngsters.

"The present study examined the possibility that some primary or basic readiness abilities, upon which Thurstone's theory of primary mental abilities is constructed, have been modified more by recent social influences than have other abilities."

3. Research Design and Procedure

The Control Subjects (comparative base) for this survey consisted of 10,035 first-grade children who participated in the 1962 norming of the SRA tests of Primary Mental Abilities (PMA).

Experimental subjects were selected from six schools within Midwestern school districts. Schools were chosen by local administrators because they "were representative of lower or upper-class neighborhoods." All 232 children attending first grade in these six schools (124 lower socioeconomic class, 108 middle socioeconomic class) constituted the Experimental Subjects. In 1972 Experimental S's were given the same version of the PMA: K-1 as administered to the norming group in 1962.

The PMA was comprised of four sub-scales:

(1) Verbal Meaning (VM) - purported to measure receptive language.

(2) Perceptual Speed (PS) - requires Ss to find similar objects.

(3) Number Facility (NF) - no explanation given. (The reviewer infers this refers to numerical recognition, simple addition and subtraction facts and similar type items).

(4) Spatial Relations (SR) - requires child to reproduce a geometric design and identify missing geometric figures from partially completed squares.

Each subscale of the 1962 version of the PMA: K-1 had a mean of 100 and a standard deviation of 16.

Four "types" or categories of IQ were considered; each corresponding to one of the subtest categories. Two levels of Socioeconomic Status (SES) were identified: Lower and Middle. Contrasts of 1962 and 1972 results on the PMA: K-1 were obtained by use of t-test statistics. Eight separate analyses were conducted (one for each of four categories of IQ, and two levels of SES).

4. Findings

Table 1 summarizes the results of the aforementioned analyses.

TABLE 1

Mean 1972 Primary Mental Abilities (PMA): K-1 Subtest IQ Attainments and Resulting z Scores when Compared with 1962 PMA Norms***

Socioeconomic class	Verbal meaning (VM)	Perceptual speed (PS)	Number facility (NF)	Spatial relations (SR)
Lower (N = 124)				
IQ	104.6	111.8	111.5	99.7
t	2.3*	6.0**	5.8**	.2
Middle (N = 108)				
IQ	107.3	111.4	107.7	100.1
t	3.7**	5.2**	3.6**	.3

* $p < .05$
** $p < .01$
*** 1962 norms: $\bar{X} = 100$, s.d. = 16

As can be seen, children enrolled in first grade during 1972 scored consistently higher on three of four subtests than those enrolled in 1962. Six of eight results were statistically significant at the .05 or .01 level. No significant differences were reported for the category labeled Spatial Relations (SR). The author speculates that "the shift in cognitive profiles suggests that social changes, such as the infusion of vast amounts of TV stimulation may produce a deficiency of sensory-motor experiences upon which more abstract forms of reading and comprehension are increasingly dependent as children progress through school." The term Sensory-motor experiences was used by the author interchangeably with the term Spatial Relations (SR), the fourth category of IQ measured by the PMA: K-1.

Critical Commentary

Surveys of this type provide useful landmark information regarding where we have been and how far we have come. They do not, however, provide reasons for observed pre-post differences. This point is made because the author on several occasions has referred to this research as an experiment, which it clearly is not.

The major question about this research concerns the selection and definition of the 1972 population. Three other concerns are briefly referred to below:

1. "Experimental" Ss in the 1972 population consisted of all first graders (N = 232) in six volunteer schools in a "medium-scope Midwestern school district." "Control" Ss consisted of 10,000+ first-grade children who participated in the 1962 norming. No descriptive population information was provided by the author. This fact coupled with the lack of randomization casts doubt as to whether these two populations were in fact comparable.

2. Given the rapidly changing social conditions referred to in the paper, it seems possible that the 1962 version of the PMA may no longer be a relevant yardstick from which meaningful changes in readiness profiles can be deduced.

3. As can be noted, Table 1 (which has been reproduced exactly) refers to "t" scores and "z" scores simultaneously. The reviewer found other reference to "z" scores in the body of the paper confusing. It is not clear how "z" scores are relevant to the analysis reported.

Changes in readiness profiles of children entering the educational system must have implications for the reconsideration and possible redesign of school curricula. Persons and/or groups responsible for the development of curricula should be aware of results such as those reported in this study and more importantly should take seriously the task of detailing specific implications for curricular redesign.

It seems obvious (to the reviewer) that this point also applies to the area of mathematics education. Curricular change in mathematics, at least as applied to the methodological concerns of mathematics instruction, has been agonizingly slow. With a few notable exceptions, the elementary school programs used by today's schools are not substantively different from those used several decades ago. Clearly more attention must be paid to academic capabilities as well as a variety of other pupil characteristics in the development and implementation of educational experiences for children. The author's final admonition is to be taken seriously: "If classroom instruction is to keep pace with the impact of social change, it appears that there is a need for periodic reassessment of children's readiness profiles."

Thomas R. Post
University of Minnesota

FINAL REPORT, SPECIALIZED TEACHER PROJECT, 1971-72. State Board of Education, San Diego, California. Mathematics Improvement Programs. Superintendent of Schools, Department of Education, San Diego County, 6401 Linda Vista Road, San Diego, California 92111.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Merlyn J. Behr, Florida State University.

1. Purpose

The Specialized Teacher Project is one of four Mathematics Improvement Programs established by the 1967 session of the California State Legislature. The Project provided in-service training for elementary teachers who subsequently taught mathematics to their own and one other class. This Final Report describes a 1971-72 study which focused on the longitudinal effects of the in-service training.

2. Rationale

As background for the 1971-72 study were studies conducted in 1968, 1969, and 1970. The focus of these studies was to investigate the questions of whether: (1) in-service training of teachers is effective as measured by their pupils' achievement, (2) one format for conducting the training is more efficient or economical than others, and (3) the exchange of classes is a desirable feature.

Data for these studies were obtained as a result of teacher-training workshops within the Specialized Teacher Project. The training program of the workshops was to acquaint teachers with the techniques of creating in their classrooms a climate in which pupils might more effectively learn mathematics. Instruction reflected the thinking of Piaget and other theorists who describe a basic hierarchy in learning, from manipulation, to representation, to abstraction. It is believed that children learn mathematics, first by responding to and manipulating objects in the environment and then by reacting to representations of these physical situations before operating with the abstractions that emerge from these earlier experiences.

"Engaging in the same kinds of activities, enjoying the same kinds of individualized learning experiences, and making the same kinds of discoveries as subsequently their pupils would, the teachers were trained to" use manipulative materials, mathematics-oriented games, measuring devices, and other aids in their instructional programs. Provision was made for participants to experience a wide variety of activities, in which manipulative materials (attribute games, pattern blocks, rods, geoboards, balances, unifix apparatus, et cetera) were used to discover mathematical concepts.

Emphasis was placed on the content defined by the Mathematics Framework for California Public Schools (The Second Strands Report): Numbers and Operations, Geometry, Measurement, Application of Mathematics, Statistics and Probability, Sets, Functions and Graphs, Logical Thinking, and Problem Solving. Classes were in session each day during the ten-day workshops from 8:30 A.M. to 3:30 P.M. The faculty consisted primarily of elementary school teachers, uniquely trained, who were successfully practicing an active learning approach to the teaching of mathematics in their own classrooms.

Significantly greater growth was observed at the second-grade level for pupils whose teachers were involved in the workshops than for other pupils, on measures of both comprehension and computation on the Modern Math Understanding Test. As measured by scales of items selected from the National Longitudinal Study of Mathematical Ability (NLSMA), fifth-grade pupils performed significantly better on nine of ten scales than did pupils whose teachers did not attend the workshops. The scales included items on computation as well as understanding of such topics as operations, geometry, probability, and graphing.

Pupil growth was found to be independent of the way in which in-service training was organized. The achievement of pupils whose teachers participated for two summers was greater than that of pupils whose teachers attended only one two-week session. Pupils who received instruction for two years from Specialized Teachers continued to grow at nearly the same rate as pupils who received such instruction for the first year. The Project was equally effective for teachers of children from various socioeconomic areas.

3. Research Design and Procedure

For the 1971-72 study the performances of pupils whose teachers attended one or two two-week summer workshops were again compared, and the impact of the training on teachers who participated in 1969 and 1970 but not in Summer 1971 was examined. Four categories of classes were defined:

1. Teachers who attended only the 1971 two-week summer workshop

2. Teachers who attended the 1970 and 1971 two-week summer workshops

3. Teachers who attended the 1969 and 1970 two-week summer workshops but did not participate in 1971

4. Teachers who attended none of the workshops

The teachers in the fourth category, i.e., those who received no in-service training, applied to attend the Summer 1971 workshops but could not be accommodated. Improved versions of the tests used in previous studies were administered to approximately 4500 pupils. Seventy-five second-grade and 74 fifth-grade classes participated in the study. The five workshops and the separate geographical regions served by each were represented proportionally in each of the cells.

4. Findings

Statistically significant differences were produced at the second-grade level on four of the five scales (Computation with Whole Numbers, Pattern Recognition, Whole-Number Operations, and Intuitive Geometry) and at grade five on two of six scales (Word Problems and Whole-Number Operations). To adjust for the differences that existed between the groups prior to the treatment and to permit the making of unbiased comparisons between the effects, an analysis of covariance procedure was employed.

The classes of second-grade teachers who attended two-week in-service training workshops during the summers of both 1970 and 1971 performed better on all scales than the teachers who attended only the 1971 session. Although the differences on individual scales were not statistically significant, there is a significant overall difference between the groups if the results of the individual scales are combined.

The classes of fifth-grade teachers who attended both the 1970 and 1971 workshops performed better than the classes of teachers who participated only in 1971 on five of the six scales. The findings thus corroborate the conclusion of the 1970-71 study that a second summer's training produces a substantial improvement in pupil performance, particularly on measures of mathematical comprehension.

It would appear that participation in the workshops had a lasting beneficial effect on pupil performance during the 1971-72 school year. Classes whose teachers participated in 1969 and 1970, even though their teachers received no reinforcement at a workshop in 1971, performed measurably better than pupils whose teachers did not attend a workshop. Statistically significant differences were observed on all five scales (Computation with Whole Numbers, Pattern Recognition, Whole-Number Operations, Intuitive Geometry, and Word Problems) at grade two, and on two of the six scales (Word Problems and Pattern Recognition) at grade five.

In addition, there were no significant differences in performance during the 1971-72 school year between the classes of teachers who attended the 1969 and 1970 summer workshops and the classes of teachers who attended the 1970 and 1971 sessions. Pupils whose teachers participated in 1969 and 1970 but not in 1971 performed as well as pupils of those teachers whose training was more recent.

5. Interpretations

The conclusion of the previous studies that the in-service training provided by the Specialized Teacher Project produces a significant improvement in pupil performance in mathematics, particularly in the area of mathematical comprehension, was verified by the 1971-72 investigation.

Critical Commentary

The study reported in this Final Report together with results of the previous studies have important implications to in-service elementary teacher education. A significant philosophy which was apparent in the conduct of the workshops was that teachers teach in the way they are taught. This philosophy was not explicitly stated but was clearly exemplified by a statement describing the program: "Engaging in the same kinds of activities, enjoying the same kinds of individualized learning experiences, and making the same kinds of discoveries as subsequently their pupils would, the teachers were trained to..." Although this report does not give a comprehensive and detailed report of the particular activities in which the teachers were engaged, The Second Strands Report is cited as a probable source for this information.

The study involved 16-21 classes per cell. This is a relatively large number; thus it is apparent that considerable confidence can be placed in the statistical results. On the other hand, since the teacher population involved was entirely in-service teachers who are more sensitive to both mathematical and pedagogical issues of the elementary school than pre-service teachers, care must be exercised in generalizing the results to pre-service education. Although one might like to believe that a similar philosophy applied to pre-service teacher education would result in similar pupil gains, this study does not give information about this important question. A study of comparable magnitude applied to pre-service education could make a significant contribution.

The report could have been improved by observing more specifically the areas in which there were greater gains for second-grade pupils of teachers who participated in the workshop than for others. Note is made here of the fact that when pupil scores of teachers participating in summer 1971 vs. no in-service are compared, greater gains were observed for Intuitive Geometry and Word Problems but not on Whole Number Computation or Whole-Number Operations. When results for teachers participating in summers 1970 and 1971 vs. no in-service are compared, greater gains were observed for pupils of participating teachers in the areas of Pattern Recognition and Intuitive Geometry and not the other three. Do these results suggest simply that as a result of the workshops teachers began to teach these concepts whereas their counterparts avoided them? This does not necessarily argue that the workshops were not successful in producing greater pupil gains, but it does cast the success into a different frame of reference. To what does one attribute the success of the training? To the fact that teachers learned and used new pedagogical techniques through the process of being taught as they "should" teach, or to the fact that they learned and began teaching mathematics they were previously uncomfortable with and had avoided? Some measures of this would have been a useful addition to the study. Significantly greater gains were not observed in the important skill areas of whole number computation and operations. Similar observations can be made about the fifth-grade data.

Merlyn J. Behr
Florida State University

AN EXPERIMENTAL TEST OF FIVE PROCESS MODELS FOR SUBTRACTION. Woods, Shirley S.; Resnick, Lauren B.; Groen, Guy J. Journal of Educational Psychology, v67, pp17-21, February 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Robert D. Bechtel, Purdue University-Calumet Campus.

1. Purpose

To show that the observed reaction times of second- and fourth-grade students in responding to "single-digit" subtraction exercises (not problems) can be made to fit one of five proposed models.

2. Rationale

Groen and Parkman (1972) investigated the processes used by children in responding to "simple" addition exercises. The simple counting models used in that investigation were adapted to this study. Five counting methods were proposed for responding to a subtraction exercise of the type m - n = ☐:

"1. The counter is set to 0; it is then incremented m times and is then decremented n times. The solution is the final value in the counter.

2. The counter is set to m and then decremented n times. The solution is the final value in the counter.

3. The counter is set to n and then incremented (m - n) times, that is, until m is reached. The solution is the number of times the counter has been incremented.

4. The counter is set to 0; it is incremented n times and is then incremented until m is reached. The solution is the number of times the counter has been incremented after n is reached.

5. Either Process 2 or Process 3 is used, depending on which requires fewer operations."

3. Research Design and Procedure

The five counting models were evaluated by fitting observed reaction times to regression lines. The subjects were 40 second-grade students and 20 fourth-grade students. A device measured the reaction time of a student to the nearest hundredth of a second.

Fifty-four exercises were presented individually to each subject for five successive days. The 54 exercises were all possible exercises of the form

m - n = ☐

where m and n are integers, $m \geq n$, $0 \leq m \leq 9$ and $0 \leq n \leq 8$. [Abstractor's Note: Why was 9 - 9 = ☐ omitted?]

4. Findings

"A series of regression analyses showed that 30 of the 40 second graders were best fitted by Model 5, with F values ranging from 13 to 36, slopes .22 to .59, and R^2s, .20 to .45. Of the remaining 10 subjects, 6 were fitted best by Model 2, with Fs ranging from 40 to 62, slopes, .45 to .55, and R^2s, .35 to .51. Two subjects were maximally significant on Model 4 (with F values of only 8 and 9, respectively). All of these F values were significant at the .01 level (with 1 and 54 degrees of freedom). Two subjects were not fitted by any of the models. The 20 fourth-grade subjects were all fitted best by Model 5, with F values from 10 to 31; slopes, .11 to .29; and R^2s, .11 to .36. Very few subjects obtained a significant F (at the .01 level) on more than one model. For example, of the 30 subjects in second grade best fitted by Model 5, only 4 were significantly fitted by any other model."

5. Interpretations

The reaction times for most second-grade students and all fourth-grade students best fit a model of counting down from the larger number (m) or counting up from the smaller number (n), using whichever required the fewer steps. Six second graders consistently counted down. Also, "Older children took only about half as much time for each increment or decrement as the younger."

Critical Commentary

This study investigated processes used in responding to 54 of the 100 basic subtraction facts. Since mastery of these facts is required of most second-grade students and every third-grade student, this study is relevant. The study reaffirmed that subtraction exercises of the form

$$\begin{array}{r} m \\ -\ n \\ \hline \end{array}$$

where n is small or the difference is small are much "easier" than those with n approximately half of m. An extension of this study could include the other 46 basic subtraction facts.

The finding that suggested some second-grade students consistently counted down while all fourth-grade students counted either up or down might be related to the real-world applications used in the instructional program. A "take-away" setting is normally the first object model for subtraction, followed by comparison applications, and finally by the "How many more are needed" (missing addend) application. The "missing addend" application is the basis for the "counting-up" process. Correlations between these settings and the computational processes could be explored.

The experimenters cited making several deletions from their data. Lengthy reaction times were omitted, as were response times for incorrect answers. These cases might prove very interesting in their own right and could help launch further research. Also, throwing away "irrelevant" data helped insure a better fit.

I assume that the references to Grade 1 (page 17) and Grade 5 (page 21) were careless errors, and should be read "Grade 2" and "Grade 4", in order to refer to subjects of the study.

Robert D. Bechtel
Purdue University-Calumet Campus

MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN RESEARCH IN EDUCATION
January - March 1976

ED 111 441 Laudato, Nicholas C.; Roman, Richard A. Computer-Assisted Instruction in Word Problems: Design and Assessment. 38p. MF and HC available from EDRS.

ED 111 668 Matthews, Paul G.; Atkinson, Richard C. Verification of Algebra Step Problems: A Chronometric Study of Human Problem Solving. Technical Report No. 253. Psychology and Education Series. 59p. MF and HC available from EDRS.

ED 111 670 Stanley, Julian C. Special Fast Mathematics Classes During School: Algebra Taught Quickly by College Professors to Fourth Through Seventh Graders. 53p. MF available from EDRS. HC not available from EDRS.

ED 111 676 Loiseau, Roger A. A Study of Coordination Between Mathematics and Chemistry in the Pre-Technical Program. 52p. Available (on loan) from ERIC/SMEAC, The Ohio State University, 1200 Chambers Road, Columbus, OH 43212.

ED 111 677 Hoile, Margaret Ann. Factor Label Method of Units Analysis: A Learning Package for Secondary Mathematics Students. 77p. MF and HC available from EDRS.

ED 111 678 Robinson, Mary L. Attitudes and Achievement: A Complex Relationship. 13p. MF available from EDRS. HC not available from EDRS.

ED 111 681 Steffe, Leslie P. (Ed.) Research on Mathematical Thinking of Young Children: Six Empirical Studies. 207p. MF available from EDRS. HC not available from EDRS.

ED 111 684 Rotman, Jack W. College Algebra and the Math Laboratory. 28p. MF and HC available from EDRS.

ED 111 688 Abramowitz, Susan. Adolescent Understanding of Proportionality: The Effects of Task Characteristics. 31p. MF and HC available from EDRS.

ED 111 689 Abramowitz, Susan. Proportionality: As Seen by Psychologists and Teachers. 14p. MF and HC available from EDRS.

ED 111 690 Abramowitz, Susan. Adolescent Understanding of Proportionality: Skills Necessary for Its Understanding. 42p. MF and HC available from EDRS.

ED 111 691 Kulm, Gerald. The Effects of Practicum Experience on the Opinions of Secondary Mathematics Teachers. 18p. MF and HC available from EDRS.

ED 111 692 Anderson, Edwin R. The Certainty of Information in Instructional Decision Making. Report No. 76-4. 25p. MF and HC available from EDRS.

ED 111 696 Consumer Math: Selected Results from the First National Assessment of Mathematics. Education Commission of the States, Denver, Colorado. National Assessment of Educational Progress. 49p. MF and HC available from EDRS.

ED 112 372 Marzano, Robert J.; Osband, Fran. Mathematics Achievement Tests and Reading. 8p. MF and HC available from EDRS.

ED 112 574 Clements, J. E.,; Tracy, D. B. The Effects of Tactile vs. Verbal Reinforcement on Attention to Task and Arithmetic Accuracy of Emotionally Disturbed Boys. 17p. MF and HC available from EDRS.

ED 112 580 Fox, Lynn H. Sex Differences: Implications for Program Planning. 26p. MF and HC available from EDRS.

ED 113 021 Wilkinson, Alexander. Counting Strategies and Semantic Analysis as Applied to Class Inclusion. Report No. 61. 49p. MF and HC available from EDRS.

ED 113 123 Jones, Richard John. Comparative Career Orientations of Beginning Male Mathematics, Science, Social Studies, and English Teachers. 101p. Available from University Microfilms (72-[illegible]877).

ED 113 152 Suydam, Marilyn N.; Osborne, Alan R. (Eds.) Algorithmic Learning. 194p. MF and HC available from EDRS.

ED 113 158 Allen, Layman E.; Ross, Joan. Instructional Gaming as a Means to Achieve Skill in Selecting Ideas Relevant for Solving a Problem. 17p. MF and HC available from EDRS.

ED 113 159 Allen, Layman E.; Main, Dana B. The Effect of Instructional Gaming upon Absenteeism: The First Step. 26p. MF and HC available from EDRS.

ED 113 160 Main, Dana B.; Plant, Mark. Resource Allocation Games as the Environment for Evaluation. 18p. MF and HC available from EDRS.

ED 113 161 Main, Dana B.; Ross, Joan. Experimental Games as an Unobtrusive Probe in Playing EQUATIONS. 5p. Available from ERIC/SMEAC, The Ohio State University, 1200 Chambers Road, Columbus, OH 43212.

ED 113 163 Allen, Layman E.; Ross, Joan. Improving Skill in Applying Mathematical Ideas: A Preliminary Report on the Instructional Gaming Program at Pelham Middle School in Detroit. 13p. MF and HC available from EDRS.

ED 113 169 Montgomery, Mary E.; Whitaker, Donald R. Report of the Coordinators' Training for Large Scale Field Testing of Developing Mathematical Processes. Technical Report No. 296. 135p. MF and HC available from EDRS.

ED 113 170 Waters, William Meade, Jr. Report on a Study to Determine the Effect of Knowledge of Recursive Definitions upon Subsequent Application of the Principle of Mathematical Induction. 61p. MF and HC available from EDRS.

ED 113 188 Trent, John H. Need for In-Service and Pre-Service Metric Education. 16p. MF and HC available from EDRS.

ED 113 189 Trent, John H. Metric Education in Mathematics Methods Classes. 4p. MF and HC available from EDRS.

ED 113 194 Harris, Linda F. Title I Individualizad Mathematics Program, Grades 4-8, Final Evaluation Report. 71p. MF and HC available from EDRS.

ED 113 195 Wenzelburger, Elfriede. Verbal Mediators in Mathematics for Transfer of Learning. Technical Report No. 3. 17p. MF and HC available from EDRS.

ED 113 196 Scandura, Joseph M.; And Others. Higher Order Rule Characterization of Heuristics for Compass and Straight Edge Construction in Geometry. 29p. MF and HC available from EDRS.

ED 113 197 Ronshausen, Nina Lorraine. A Comparison of the Effects on Achievement and Attitude of Two Methods of Tutoring First-Grade Mathematics in the Inner City: Programed vs. Directed. 225p. Available from University Microfilms (72-6682).

ED 113 198 Test Factors, Instructional Programs, and Socio-Cultural-Economic Factors Related to Mathematics Achievement of Chicano Students: A Review of the Literature. SMESG Working Paper No. 10. Stanford Univ., California School Mathematics Study Group. 33p. MF and HC available from EDRS.

ED 113 203 Hubbard, W. Donald; Buchanan, Anne E. Developing Mathematical Processes: 1972-73 Field Test Report. Technical Report No. 324. 67p. MF and HC available from EDRS.

ED 113 204 Abernatha, Evelyn; Wiles, Clyde A. A Three Month Trial of Developing Mathematical Processes (DMP) with Ten Educable Mentally Retarded Children. Technical Report No. 336. 31p. MF and HC available from EDRS.

ED 113 212 Roman, Richard A. The Word Problem Program: Summative Evaluation. 14p. MF and HC available from EDRS.

ED 113 703 McKinney, John A. The Development and Implementation of a Tutorial Program for Parents to Improve the Reading and Mathematics Achievement of Their Children. 133p. MF and HC available from EDRS.

ED 114 241 Therkelsen, Edward Robert. The Basic Mathematics of Astronomy: A Sourcebook for Science Teachers. 192p. Available from University Microfilms (70-26,400).

ED 114 374 Mohr, Paul (Ed.) College of Education, Florida A&M University, Tallahassee, Florida. Research Monograph. 42p. MF and HC available from EDRS.

MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN JOURNALS AS INDEXED BY CURRENT INDEX TO JOURNALS IN EDUCATION
January - March 1976

EJ 122 045 Applebee, Bernice L. An Inquiry into the Relationship Between Math Growth and Correction Patterns. Illinois School Research, v11 n3, pp43-46, Spring 1975.

EJ 123 010 Suppes, Patrick; And Others. Performance Models of American Indian Students on Computer-Assisted Instruction in Elementary Mathematics. Instructional Science, v4 n3/4, pp303-314, October 1975.

EJ 123 169 Belsom, C. G. H.; Elton, L. R. B. The Effect of Syllabus on Mathematical Knowledge. Physics Education, v9 n7, pp462-463, November 1974.

EJ 123 960 Dimitrovsky, Lilly; Almy, Millie. Early Conservation as a Predictor of Arithmetic Achievement. Journal of Psychology, v91, pp65-70, September 1975.

EJ 124 740 Grouws, Douglas A.; Reys, Robert E. Division Involving Zero: An Experimental Study and Its Implications. Arithmetic Teacher, v22 n1, pp74-80, January 1975.

EJ 124 754 Freudenthal, H. Soviet Research on Teaching Algebra at the Lower Grades of the Elementary School. Educational Studies in Mathematics, v5 n4, pp391-412. December 1974.

EJ 124 758 Jansson, Lars C. Structural and Linguistic Variables that Contribute to Difficulty in the Judgment of Simple Verbal Deductive Arguments. Educational Studies in Mathematics, v5 n4, pp493-505, December 1974.

EJ 124 842 Selkirk, Jack. An Enquiry Into Adverse Attitudes Towards Advanced Level Mathematics. International Journal of Mathematical Education in Science and Technology, v6 n2, pp181-186, May 1975.

EJ 126 462 Dodd, Carol A.; And Others. Diagnosis and Remediation of Pupil Errors: An Exploratory Study. School Science and Mathematics, n75, pp270-276, March 1975.

EJ 126 717 Burrell, K.; And Others. "Do You Mean I've Got it Wrong?" Mathematics Teaching, n31, pp53-57, 1975.

EJ 126 720 Roberge, James J. Development of Comprehension of Logical Connectives in Symbolic or Verbal Form. Educational Studies in Mathematics, v6 n2, pp207-212, July 1975.

EJ 126 724 Branca, Nicholas A. Learning a Structure by Mathematics Teachers. Australian Mathematics Teacher, v30 n6, pp213-217, December 1974.

EJ 126 725 Saunderson, A. Effect of Immediate Knowledge of Results on Learning. Australian Mathematics Teacher, v30 n6, pp218-221, December 1974.

EJ 126 751 Lincoln, Eugene A.; And Others. Parents Make a Difference in Teaching in Urban Schools. Arithmetic Teacher, v22 n6, pp460-463, October 1975.

EJ 126 797 Bright, George W. Identification of Embedded Figures by Primary Grade Children. School Science and Mathematics, v75 n6, pp535-541, October 1975.

EJ 126 834 Pavlick, Frank M. The Attitudinal Effect of Using the Computer in an Elementary Statistics Course. International Journal of Mathematical Education in Science and Technology, v6 n3, pp353-360, August 1975.