DOCUMENT RESUME

ED 175 732 SE 028 859

AUTHOR Suydam, Marilyn N., Ed.
TITLE Investigations in Mathematics Education, Vol. 12, No. 3.
INSTITUTION Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE 79
NOTE 71p.
AVAILABLE FROM Informat. Reference Center (ERIC/IRC), The Ohic State University, 1200 Chambers Rd., 3rd Floor, Columbus, Ohio 43212 (subscription $6.00, $1.75 single copy)
JOURNAL CIT Investigations in Mathematics Education: v12 n3 Sum 1979

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Abstracts: Anxiety: *Concept Formation: *Effective Teaching: Elementary Secondary Education: Learning: *Mathematics Education: Research Reviews (Publications): *Sex Differences: *Student Ability: *Thought Processes

ABSTRACT

Thirteen research reports related to mathematics education are abstracted and analyzed. Four of the reports deal with learning or thought processes, two each with concept formation, sex differences, teacher effects and student abilities, and one report deals with anxiety. An editor's note lists various other sources of reviews of research in mathematics education. (PK)

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 Clearinghouse for Science, Mathematics and Environmental Education

Volume 12, Number 3 - Summer 1979

Subscription Price: $6.00 per year. Single Copy Price: $1.75
Add 50¢ for Canadian mailings and $1.00 for foreign mailings..

A note from the editor . . .

IME has presented abstracts and commentaries on a wide variety of reports in mathematics education. It is obvious, however, that all of the research has not been reviewed in this journal. Your attention is called to the fact that the Journal for Research in Mathematics Education added a "Telegraphic Reviews" feature beginning with the May 1978 issue; longer reviews have also appeared in JRME since January 1977. These reviews have included the following:

Callahan, Leroy G. and Glennon, Vincent J. Elementary School Mathematics: A Guide to Current Research. Washington: Association for Supervision and Curriculum Development, 1975. [JRME 8: 151-152; March 1977.]

Cronbach, L. J. and Snow, R. E. Aptitudes and Instructional Methods: A Handbook for Research on Interactions. New York: Irvington Publishers, 1977. [JRME 9: 390-395; November 1978.]

Fox, L. H.; Fennema, E.; and Sherman, J. Women in Mathematics: Research Perspectives for Change. NIE Papers in Education and Work, No. 8. Washington: National Institute of Education. [JRME 9: 387-390; November 1978.]

Krutetskii, V. A. The Psychology of Mathematical Abilities in Schoolchildren. (Translated by Joan Teller; edited by Jeremy Kilpatrick and Izaak Wirzup.) Chicago: University of Chicago Press, 1976. [JRME 8: 237-238; May 1977.]

Stanley, J. C.; George, W. C.; and Solano, C. H. The Gifted and the Creative: A Fifty-Year Perspective. Baltimore: The Johns Hopkins University Press, 1977. [JRME 10: 73-78; January 1979.]

Stanley, Julian C.; Keating, Daniel P.; and Fox, Lynn H. (editors). Mathematical Talent: Discovery, Description, and Development. Baltimore: The Johns Hopkins University Press, 1974. [JRME 8: 148-151; March 1977.]

Thompson, S. and DeLeonibus, N. Guidelines for Improving SAT Scores. Reston, Virginia: National Association of Secondary School Principals, 1978. [JRME 9: 384-386; November 1978.]

National Advisory Committee on Mathematical Education. Overview and Analysis of School Mathematics Grades K-12, The NACOME Report. Washington: Conference Board of the Mathematical Sciences, 1975. [JRME 8: 68-73; January 1977.]

On Further Examination: Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline. New York: College Entrance Examination Board, 1977.
[JRME 9: 155-160; March 1978.]

Telegraphic Reviews have appeared in the following issues:
JRME 9: 238-239; May 1978.
JRME 9: 396-397; November 1978.
JRME 10: 79-80; January 1979.
JRME 10: 159-160; March 1979.
JRME 10: 239-240; May 1979.

Other journals have also published reviews of research reports in mathematics education; thus, the Arithmetic Teacher, the Mathematics Teacher, and the American Mathematical Monthly are occasional sources. Another recent review is the following:

Gelman, R. and Galistell, C. H. The Child's Understanding of Number. Harvard University Press, 1978.
[Educational Researcher 8: 16-18; May 1979.]

The Calculator Information Center has developed two documents containing reviews of research on calculators, using a format parallel to that of IME. Dissertations comprise a large proportion of the 44 reports analyzed in these volumes:

Investigations with Calculators: Abstracts and Critical Analyses of Research (edited by Marilyn N. Suydam). Columbus, Ohio: Calculator Information Center, January 1979.

Supplement, June 1979.

If you know of other reviews of research reports in mathematics education, please let me know.

INVESTIGATIONS IN MATHEMATICS EDUCATION

Summer 1979

Babad, Elisha Y. and Bashi, Joseph. ON NARROWING THE PERFORMANCE GAP IN MATHEMATICAL THINKING BETWEEN ADVANTAGED AND DISADVANTAGED CHILDREN. Journal for Research in Mathematics Education 9: 323-333; November 1978.

Abstract and comments prepared for I.M.E. by ROBERT D. BECHTEL, Purdue University Calumet.

1. Purpose

The purpose of this study was to narrow the performance gap between "advantaged" and "disadvantaged" students in mathematical thinking. Mathematical thinking was assessed by means of a special curriculum in cryptarithmetic designed for this project. (In cryptarithmetic each digit in a whole number addition exercise is replaced by a letter. The task is "to break the code"; i.e., to write the exercise using standard numerals.)

2. Rationale

Appreciable gaps occur between disadvantaged and advantaged students in socioeconomic status (SES), IQ, and achievement in mathematics. (As defined in the study, school children in Jerusalem, Israel, whose parents immigrated into Israel from Asian and African countries and who suffer various degrees of environmental and educational deficits are labeled disadvantaged.) The major hypothesis was that after the instructional program of the study, the performance gap in in cryptarithmetic between advantaged and disadvantaged students would be smaller than the gaps in SES, IQ, and mathematics achievement, and similar gaps observed in Learning Potential (LP) tests. (The authors cite their previous work with LP tests. Nonverbal reasoning tasks are administered in the LP test in a "test-coach-test" sequence.)

3. Research Design and Procedures

Subjects and Teachers. The subjects were 162 students in five seventh-grade classes in five Jerusalem schools, selected to represent a wide range of home background and achievement. All subjects in the sample had mastered the prerequisites for the cryptarithmetic program, a mastery of the four basic arithmetical operations. The cryptarithmetic program was taught by three experienced teachers in the regular context of the classroom with 30-35 students. The teachers did not know the students prior to the study and were not given information regarding academic standings of the students.

Procedure. The Milta Verbal Intelligence Test, a mathematics achievement test (based on the mathematics curriculum of the Israeli Ministry of Education) and the Series Learning Potential Test (Babad and Budoff), pretest then posttest, were administered in this listed order within a period of two weeks before the beginning of the cryptarithmetic program. The cryptarithmetic program was taught in all classes during hours scheduled for mathematics and over a period of two months. The cryptarithmetic test was given immediately upon completion of the program. The sample was divided into high and low SES groups according to the father's level of education; 8 or more years placed the subject in the high-SES group.

4. Findings

The table below shows the major findings.

Means and Standard Deviations of the Measured Variables in the Total Sample and the Low-SES and High-SES Subsamples

Variable	Total Sample $N = 162$		Disadvantaged (Low-SES) Fathers' Education 5.8 Years $N = 94$		Advantaged (High-SES) Fathers' Education 11.8 Years $N = 68$		Difference Between High-SES and Low-SES
	$\bar{X}$	*SD*	$\bar{X}$	*SD*	$\bar{X}$	*SD*	D
Achievement in Mathematics	15.31	5.9	14.29	5.6	17.16	5.6	2.87*
Verbal IQ	97.55	12.7	96.30	11.1	101.93	12.0	5.63*
LP Pretest	55.94	11.4	54.53	11.5	58.47	11.0	3.94*
LP Posttest	70.06	9.9	69.51	9.3	71.30	10.4	1.79n.s.
Cryptarithmetic	15.65	5.4	15.33	4.9	16.58	5.6	1.25n.s.

*t-test for the difference between the means of the high-SES and low-SES groups significant, $p < .05$.

The two SES groups differed in performance on the conventional IQ and mathematics achievement tests, and did not differ in performance on the LP posttest and the cryptarithmetic test.

5. Interpretations

The authors state that their earlier studies "showed that disadvantaged students can be taught problem-solving strategies and can reach a level of mastery in conceptual thinking that is not expected from their performance on one-shot IQ tests." Then the authors assert:

> This study was an attempt to enlarge the scope of this approach from the domain of nonverbal analogies to mathematical thinking, which is a crucial component of the school curriculum in mathematics.
>
> The results of this project give support to the hope that the cultural gap can be somewhat narrowed--both in mental assessment and in school achievement--without compromising the required level of conceptual thinking. To achieve that, mastery of the prerequisites must be guaranteed, and material cannot be taught before the students can perform all the operations required for entering that unit of learning. Special attention must be given to disadvantaged slow learners by supporting them as they enter a unit that they have a feasible chance of mastering, and by preventing deficits in moving from one unit to another. This calls for utmost care in designing the learning environment and the learning process; a strict set of basic principles should be observed.

Abstractor's Comments

This study shows that a program which is well-planned throughout, beginning with learner prerequisites and ending with terminal skill behaviors, and which is carefully taught by teachers who interact with their students, will be effective. Of course, this is well known; the study by E. Y. Babad and J. Bashi is a replication study. So-called "disadvantaged" students are students, hence this principle applies to them. To illustrate this point, read the following statement of the authors twice, omitting the word "disadvantaged" in the second reading.

The learning environment should be planned so as to reduce the pressures on the disadvantaged student by providing continuous support, enhancing the student's sense of competence through ample opportunity for successful practice, and maintaining the student's interest and motivation throughout the program.

A few other comments:

1. It seems a shame that the cryptarithmetic program in the study used a considerable amount of time--valuable time for the students in their study of mathematics. Cryptarithmetic is not a mainstream topic in the mathematics curriculum.

2. One may conduct studies ad infinitum to show that "disadvantaged" students can learn "like other people." Such studies and, in fact, all reasonable arguments, will have no effect on the actions of biased, prejudiced people.

3. The conduct of studies in which people are classified as "disadvantaged" in pseudo-scientific ways subtly reinforces class consciousness and prejudice.

Christoplos, Florence and Borden, JoAnn. SEXISM IN ELEMENTARY-SCHOOL MATHEMATICS. Elementary School Journal 78: 275-277; March 1978.

Abstract and comments prepared for I.M.E. by SIGRID WAGNER, University of Georgia.

1. Purpose

The purpose of this study was to investigate the effect of sex-stereotyped content in verbal problem-solving mathematical questions on the mathematics scores of a group of first-grade boys and girls.

2. Rationale

It has been shown that girls perform better on tests of mathematical problem-solving ability when (1) their attitude toward mathematics improves or (2) the content of the problems has to do with stereotypical female interests. One reason that boys do better than girls in mathematics may be that the content of verbal problems in elementary-school mathematics textbooks is often oriented toward stereotypical male interests.

3. Research Design and Procedures

The subjects were 27 first-grade students, 12 girls and 15 boys, in an intact classroom. The instrument consisted of ten pairs of verbal problems, one in each pair being of stereotypically female content, one of stereotypically male content. All 20 problems are recorded verbatim in the article. The first two problems were these: (1) If you had one hammer and your daddy brought you another one, how many hammers would you have? (2) If you had one doll and your daddy brought you another one, how many dolls would you have? Four pairs of problems involved basic addition facts, four involved basic subtraction facts, and two involved the combinations 12 + 6 and 11 + 9.

The 20 problems were ordered in consecutive pairs with the "male" problem first in 6 of the 10 pairs. The problems were read aloud to the students in an individual interview setting. Each student was given two scores: the number of correct responses to the female-oriented questions and the number of correct responses to the male-oriented questions.

4. Findings

Mean scores for boys and girls on the two types of questions are shown below:

	Girls	Boys
Female-oriented questions	5.00	4.00
Male-oriented questions	3.70	5.67

Using a one-tailed t-test for dependent scores, it was found that the girls' mean score was significantly higher on the female-oriented questions than on the male-oriented questions ($t = 4.45$, $p < .05$). No girl scored better on the male-oriented questions than on the female-oriented questions; similarly, no boy scored better on the female-oriented questions than on the male-oriented questions. For total scores, differences between boys and girls were not significant.

5. Interpretations

The authors state that the results "warrant respect" even though they "cannot legitimately be generalized" (p. 276). The authors also note that the differences between the two test scores were greater for boys than for girls; they conjecture that it may be harder for boys to break away from self than it is for girls, who more often have to do so. The authors conclude that educators have a responsibility to reduce as much as possible the effects that sex-biased text materials may have upon student achievement.

Abstractor's Comments

The authors chose an interesting question to investigate. It is unfortunate that the results of this particular study are confounded by a weakness in the design of the test instrument. Although the authors were careful to use pairs of exactly parallel questions, it is not clear why they chose to ask matching questions in consecutive order. Many children participating in the study must have noticed this pattern rather quickly and may have reacted in either of two ways: (1) Some students may have realized that the answers to matching questions had to be the same; these students may, therefore, have seemed less affected by the sex-stereotyped content than they otherwise might; (2) Other students may have decided their first answers were incorrect and so changed them upon being given a "second chance." The effect of this latter reaction was partially controlled by the fact that in 60 percent (why not exactly half?) of the pairs of questions the male-oriented question was asked first. Nevertheless, data on (at least) the split-half reliability of the instrument should have been provided. Better yet, the order of the questions should have been randomized.

For the purposes of this article, the authors obviously made no attempt to do an exhaustive review of the literature. However, their brief discussion of the rationale for the study leaves the reader with the distinct impression that males at any age outperform females in mathematics. In fact, most available data suggest that significant sex-related differences in mathematics achievement usually do not appear until around the middle school years and rarely as early as first grade. Since mathematics textbooks have in the past tended to be biased toward stereotypically male content, it would seem that, if sex-biased content does affect achievement, it is a long-term effect.

On the other hand, this study provides some data that suggest that sex-biasing may have an immediate effect and, therefore, that girls in the early grades have done as well as boys in mathematics only by overcoming the handicap of male-oriented mathematics content. These girls may do even better in the future, now that sex-biasing has been largely eliminated in most mathematics textbooks. More ironic is the authors' finding that boys were more affected by sex-stereotyped content than girls were. One implication of this result is that the elimination of sex bias in mathematics textbooks may be hurting boys more than it is helping girls. If boys in the early grades have needed the advantage of mathematics materials oriented to their interests in order to do as well as girls, it will be interesting to see what long-term effects sex-balanced materials may have.

Lest all this speculation become too enticing, a sobering fact asserts itself: one of the reasons the authors of this article found statistically significant differences in the performance of girls and boys on these mathematics problems is that neither the boys nor the girls did very well on the test. No one would argue with the authors' contention that educators should strive to reduce the detrimental effects of sex stereotyping, but surely the first priority of mathematics educators must be to improve the mathematics achievement of all students. By improving the performance of both boys and girls, many statistically significant sex-related differences may be found to be really insignificant.

DeWolf, Virginia A. HIGH SCHOOL MATHEMATICS PREPARATION AND SEX DIFFERENCES IN QUANTITATIVE ABILITIES. Seattle: Educational Assessment Center, University of Washington, 1977. ERIC: ED 149 968.

Abstract and comments prepared for I.M.E. by GERALD KULM, Purdue University.

1. Purpose

The study investigated sex differences in quantitative ability. It was hypothesized that sex differences on quantitative test scores would not emerge when subjects were grouped.

2. Rationale

The work of Fennema and Sherman (1977) provided the motivation for this study. The differences in scores on achievement tests can be explained by the different patterns in coursework between males and females. Males enroll in more mathematics-related courses in school. When this difference in backgrounds is controlled statistically, many quantitative and spatial ability differences disappear. This rationale was extended in the present study to include differences in mechanical reasoning ability which may be accounted for by a background in courses such as physics.

3. Research Design and Procedures

The subjects were high school juniors who had taken the Washington Pre-college (WPC) battery. Every tenth person was selected producing 2093 (962 male, 1131 female) subjects. Tests selected from WPC were: Quantitative Skills (computational relationships), Applied Mathematics (algebra word problems), Mathematics Achievement (algebra and geometry knowledge), Spatial Ability (visualize transformations in three dimensions), and Mechanical Reasoning (physical principles of mechanical devices). A Quantitative Composite score on the first three tests was also used. Finally, each subject's grades and credits in general mathematics, algebra, geometry and advanced mathematics, and physics were obtained. For each sex, subjects were grouped according to the number of semesters of mathematics and physics taken. Five major coursework patterns were obtained when these groups were formed.

For the total groups, t-tests were computed for the mean semesters taken in each subject, overall mathematics GPA, and the WPC test scores. Next, in each of the five coursework groups having 30 or more subjects, t-tests were computed for overall mathematics GPA and the WPC scores. Finally, multiple regression analyses were performed using the total sample. The WPC scores were predicted from coursework alone and from coursework plus sex. On one criteria for which sex was a significant additional predictor, actual and predicted scores (without sex) were compared for each sex separately. A similar regression procedure was carried out using mathematics GPA as an additional predictor.

4. Findings

The t-tests revealed that males took more semesters of mathematics and physics ($p < .001$) (except general mathematics) and scored higher on all of the WPC tests ($p < .001$). The females had a significantly higher mean mathematics GPA ($p < .05$).

When groups with similar coursework were compared, a different pattern was observed. Males scored significantly higher on mechanical reasoning in all five groups and females had higher mathematics GPA's in four of five groups. In none of the groups were there differences in mathematics achievement or spatial ability scores. Finally, as the groups gained mathematical background, male performance on quantitative tests and applied mathematics became more superior.

The regression analyses using the total sample confirmed the t-test results. Sex was a significant added predictor for quantitative, applied mathematics, and mechanical reasoning scores. Similar results were obtained when mathematics GPA was included. Females' scores were overpredicted and males' scores underpredicted when compared with actual WPC scores.

5. Interpretations

Sex differences in quantitative and spatial ability are partially reflections of different mathematics training. Controlling algebra and and geometry coursework explains the disappearance of differences on mathematics achievement but not for spatial ability and, to some extent, quantitative ability. The differences in quantitative ability which remain may be due to: (1) sex bias in item content, (2) lack of control for relevant coursework in applied mathematics (e.g., woodworking, electronics), or (3) lack of control for quantitative aptitude prior to high school.

The type of mathematics coursework taken appeared in this study to reflect mathematical aptitude. With this interpretation, it seems important in future studies to consider both coursework and mathematics aptitude in interpreting sex differences in quantitative ability. Aptitude may play an important role in sex differences in mathematical ability.

Abstractor's Comments

This study provides further evidence that presumed male superiority in mathematics can be explained largely by the differences in previous coursework. The method of random selection from a large population sample makes the results of the study worth noting. The technique of considering specific coursework patterns revealed interesting results which ought to be investigated further. The author's hypothesis that mathematics aptitude is involved seems to be worth following up.

The author's discussion did not include the spatial ability and mechanical reasoning results. The latter seems to be explained by the non-inclusion of coursework or other prior experience with mechanical ideas. The spatial ability result suggests that previous coursework in mathematics does provide practice with spatial concepts and accounts for differences found between sexes.

Little doubt remains by now that sex differences are in reality cultural rather than cognitive. There are still a few complexities to be explained, such as the ones raised in this study about quantitative and applied mathematics. However, the important work still lies ahead in finding ways to encourage females to study science and mathematics.

Good, Thomas L. and Grouws, Douglas A. TEACHING EFFECTS: A PROCESS-PRODUCT STUDY IN FOURTH-GRADE MATHEMATICS CLASSROOMS. Journal of Teacher Education 28: 49-54; May-June 1977.

Abstract and comments prepared for I.M.E. by LINDA R. JENSEN, Temple University.

1. Purpose

The purpose of this study was to identify teachers who were consistently either relatively effective or relatively ineffective in terms of student scores on the Iowa Test of Basic Skills and to identify differences in classroom behavior in these teachers.

2. Rationale

The authors contend that systematic research on observed teaching behaviors is a relatively new field of research. Few studies have included observational measures designed to measure teacher functions as an independent variable used to influence student achievement. This study builds on suggestions from a major forum on recent research on teacher effects presented in the Spring 1976 issue of the Journal of Teacher Education.

3. Research Design and Procedures

Nine teachers were identified as relatively stable and effective and nine teachers were identified as relatively stable and ineffective, based upon total mathematics residual gain scores on the Iowa Tests of Educational Development. These teachers were in the top one-third and the bottom one-third of the sample, respectively, across two consecutive years. These teachers also maintained their relative position during the third year, the year of the study.

Subjects were selected from a school district which skirted the core of a large urban city, and students were mostly middle class, although a number of students from high and low income homes were included. The same textbook was used in all classes and the mean aptitude for students in the district had been 100-103 for over a decade.

Two trained observers studied 41 teachers during mathematics instruction periods in October, November, and December. (Forty-one teachers were studied to protect the identities of the previously chosen nine effective and nine ineffective teachers.) Each teacher was observed 6 to 7 times, with each coder making roughly one-half of the observations in a given classroom. The coders did not know the teacher's levels of effectiveness.

Data were collected on the dependent product variable, the total mathematics score on the Iowa Test of Basic Skills, in October and April. Data were collected on the independent process variable in four areas:

a. The distribution of time during mathematics instruction.
b. Teacher-student interaction: low inference descriptions from the Brophy-Good Dyadic system.
c. Teacher-student interaction: high inference descriptions from the work of Emmer & Kounin.
d. Checklists to describe materials and homework.

Classroom mean residual scores on total mathematics were computed and associated with the classroom process (observational) data. Residual scores were used to classify teachers as high-effective or low-effective.

Classroom process data were analyzed using a one-way analysis of variance to see if the high effective teachers differed from the low effective teachers. Significant and near significant results are reported in a table for only the teacher-student interactions. Results are not reported statistically on the distribution of time or the checklists describing materials and homework.

4. Findings

Due to space limitations and the feeling that a teaching act cannot be meaningfully isolated, findings were reported in an integrative manner with the emphasis on positive findings. Variables discussed are ones on which high and low effective teachers are significantly different and ones on which individual teachers within a high or low group consistently show more or less of than teachers in the other group. Although teachers were similar on many factors, these are not discussed, except for the fact that all nine high and all nine low effective teachers utilized whole class instruction.

Teaching effectiveness in terms of distinguishing high and low teachers seemed to be most associated with the following teacher behaviors:

a. Clarity. High teachers generally introduced and explained material more clearly than lows.

b. Questioning. The discussion states that highs asked more product questions (a question that demands a single correct answer), and lows were more likely to ask process questions (a question that demands integration of facts, explanation). This seems to be misleading when compared to the table which reports more direct questions than process questions for both highs and lows and more questions of both types from the lows.

Variables	p value	$\bar{X}$ High	$\bar{X}$ Low
Direct Question	.0113	14.07	28.26
Process Question	.0131	2.72	7.53

c. Expectations. The discussion reports that highs demanded more work and achievement from students. (This is not included statistically on the table.)

d. Feedback. Highs tended to give more immediate, nonevaluative, task-relevant and process feedback than did lows. High teachers were approached by their students more often than were low teachers. Low teachers initiated the contact more frequently than did high teachers. Low teachers also used both praise and criticism more often than highs. The discussion also states that low teachers initiated more alerting and accountability messages than did highs, but this does not seem to be supported by the table which reports:

Variables	p value	$\bar{X}$ High	$\bar{X}$ Low
Average Accountability	.0424	3.46	3.15
Average Alerting	.0350	3.90	3.59

e. Discipline. Lows appear to have more managerial problems than do highs, but discipline questions are significant at only the .0656 level.

5. <u>Interpretations</u>

The authors draw conclusions based on the following limitations:

a. Teacher effectiveness is defined in terms of student progress on a standardized achievement test. Studies linking classroom process to other operational definitions of effectiveness are needed.

b. Only linear relationships between process and product were described. Analysis should be extended to cover any non-linear relations in the data.

c. No attempt was made to distinguish between teacher/student interaction of high and low achievers. The data describes teacher behavior towards students in general.

Given these limitations, the following conclusions and implications were drawn:

a. Too much or inappropriate praise may not facilitate learning.

b. In a highly structured subject like mathematics, it might be better to ask relatively more product than process questions.

c. Teaching to the whole class may not be a bad instructional strategy.

d. Successful teaching is based on a large number of variables, not just one or two critical factors.

Abstractor's Comments

Good and Grouws are to be commended for their observational research on teacher effectiveness. This type of study is very important to teachers and teacher educators as they attempt to understand and implement effective teaching strategies. This work needs to be expanded to include other subject areas with other process and product measures.

A careful study of the table and discussion suggests the following questions and comments:

1. The use of residual gain scores on the Iowa Tests of Basic Skills as the product measure is questionable. There are several reasons for this:

a. The median year-to-year correlations across residual gain scores on all Iowa subtests was .20. This is quite low, even though the

18 teachers who were classified as either high or low were in the top one-third or the bottom one-third for three consecutive years. Teachers as a whole were inconsistent and these 18 may have been chosen by a fluke.

b. Many factors other than the teacher may affect student performance; these are not mentioned. These include such things as socioeconomic status and aptitude. The authors admit that a relatively stable context is needed because residualizing will adjust statistically but not correct for real differences in teaching conditions. However, no attempt is made to show that the classrooms studied were equivalent in these areas. It was reported that the district was largely middle class but a number of students from high and low income homes attended the schools. Were low income students attending the classes of low effective teachers and high income students attending the classes of high effective teachers? The mean aptitude for students in the district was 100-103, but no mean aptitudes were reported for individual schools or teachers. They could have varied widely. Stability in a district does not ensure equality across schools or teachers. Some schools may have practiced ability grouping. These things should have been investigated and reported.

c. As the authors noted, the type of measure used may have an effect on the type of teacher selected. A standardized test may emphasize computational rather than conceptual skills. This might be enhanced by product rather than process questions. Affective scores might favor different classrooms. It would be good to follow up with studies in these areas.

2. Although time measures were taken for descriptive purposes and to test hypotheses on time variables in mathematics instruction such as the ratio of time spent in development vs. practice activities, these results were not reported. The same is true of the checklists to describe materials and homework assignments. This information would have added to the article.

3. A one-way analysis of variance limits the reported results. No interactions could be detected. As the authors note, it eliminates

any analysis of aptitude-treatment interaction. It also may be that it is not suitable to the analysis of the time measures or checklists, and that may be one reason they were not reported. Further analysis is needed.

4. Parts of the table are confusing and some entries seem to be in conflict with the discussion. Discrepancies on accountability, alerting, and questioning were mentioned earlier. Four entries which appear to give ratios of questions, contacts and responses are confusing. This section reads:

Variables	p Value	$\bar{X}$ High	$\bar{X}$ Low
Direct Questions / Total Response Opportunities	.1089	28.13	36.54
Total Teacher Initiated Contacts / Total Student Initiated Contacts	.0058	54.10	116.41
Process Questions / Total Questions	.0518	7.44	14.56
Correct Responses / Total Responses	.0051	82.80	76.17

From other information given in the table, it appears as though this should be:

Variables	$\bar{X}$ High	$\bar{X}$ Low
Direct Questions / Total Questions	$\frac{14.07}{16.79} = .83$	$\frac{28.26}{35.79} = .79$
Total Teacher Initiated Contacts / Total Student Initiated Contacts	$\frac{7.23}{25.35} = .16$	$\frac{11.83}{13.41} = .88$
Process Questions / Total Questions	$\frac{2.72}{16.79} = .16$	$\frac{7.53}{35.79} = .21$
Correct Responses / Total Responses	$\frac{38.70}{44.09} = .88$	$\frac{50.98}{62.37} = .82$

This section is quite confusing and is not clarified by the discussion.

5. No level of significance is set and "significant or near significant process variables" are reported in the table. This leaves the readers to set their own levels of significance. Levels of significance on the table range from .001 to .1089. Since variables which are not "near significant" are not reported, it is questionable what cut-off was chosen as "near significant."

6. It is interesting that the most significant variable, the number of students (with a p Value of .0001), is not discussed. It could be because the high teachers have an average class size of 26.70 and the low teachers have an average class size of 21.34. Although cause and effect is attributed to other variables on the list, the authors are seemingly reluctant to conclude that larger class size contributes to higher test scores. Perhaps, they should be equally reluctant to make statements such as "for producing gains in a highly structured subject like mathematics it may be better instructional strategy for teachers to ask relatively more product than process questions." The proportion of direct questions to total response opportunities is reported "near significant" at the .1089 level in favor of the high teachers. This is a strong conclusion based on weak evidence. Even if differences are shown in the data, this does not necessarily imply cause and effect.

The authors have suggested a rich avenue of research. This area should continue to be expanded.

Greabell, Leon C. THE EFFECT OF STIMULI INPUT ON THE ACQUISITION OF INTRODUCTORY GEOMETRIC CONCEPTS BY ELEMENTARY SCHOOL CHILDREN. School Science and Mathematics 78: 320-326; April 1978.

Abstract and comments prepared for I.M.E. by J. FRED WEAVER, The University of Wisconsin-Madison.

1. Purpose

The investigation was "designed to determine if planned exposure to a greater number of stimuli will effect [sic] achievement of subjects in an introductory unit in Geometry."

2. Rationale

"From the evidence available [from several cited studies] it is evident that stimuli effect [sic] children in different ways but there appears to be no research done in the area of how these stimuli effect [sic] the child's school achievement directly." It would be desirable to have such information.

3. Research Design and Procedures

The subjects consisted of 108 seven-, eight-, and nine-year-old children (59 male, 49 female) from a lower middle class SES area of a large city in the southeast U.S. who attended a six-week summer program (largely "because both parents worked and the program was convenient to their neighborhood"), randomly assigned to self-contr'..ed classrooms in two elementary schools.

Two instructional treatments were devised, each consisting of nine 45-minute lessons on "Geometry" (to be taught on consecutive school days during the second and third weeks of the summer program), followed by a posttest (to be administered on the tenth day of the two-week period) of 30 items. The lesson plans (to be taught by "fifth quarter interns at the University of South Florida") were "designed using a widely used children's [sic] mathematics text as a basis for planning." One treatment, Low Stimuli Group (LSG), consisted of instruction based upon the lesson plans as designed. The other treatment, High Stimuli Group (HSG), consisted of instruction based upon the same lesson plans augmented by "additional stimuli... systematically incorporated into each lesson as defined by F.A.C.T.," a Functional Analysis of Classroom Tasks, an "observation system" in which stimuli are categorized in accord with the nine-cell scheme of Figure 1.

		Cognitional		
		Concrete	Representative	Abstract
Sensory	Visual			
	Auditory			
	Tactile			

Figure 1. Stimulus combinations for the Functional Ar lysis of Classroom Tasks (F.A.C.T.).

Treatments were administered to 11 groups (7-14 children per group) by 11 teachers. The LSG (low-stimulus) treatment was used with five of the groups (51 subjects); the HSG (high-stimulus) treatment, with six groups (57 students). "Each classroom unit had only one type of group (HSG or LSG)."

The same posttest was given (read) to each of the 11 groups. No pretest was used (presumably as a consequence of accepting the following assumptions: "Based on the normal school curriculum used in this county, no child has had formal instruction in geometry. The children selected for the study represent a random school population.").

4. <u>Findings</u>

a. Based upon posttest scores pooled by treatment across instructional groups, the following were reported (verbatim):

Table 2. Comparison of Post-Test Scores Between A High Stimuli Group and A Low Stimuli Group

Experimental Groups	N	Mean	Variance	Range	df	t
Group (LSG)	51	21.19	21.00	9.30	107	-2.709*
Group (HSG)	57	23.33	13.48	14.30		

*($p < .01$)

"As the results of Table 2 indicate, the difference between the two groups is significant.

"In view of the above it appears that systematically planning and increasing stimuli in geometry at this level does have an effect on achievement of children."

b. Posttest scores for the two treatments were compared independently for subjects at each of three IQ levels--below 90, 90-110, above 110--with the following reported results (verbatim):

Table 3. Analysis of Variance for Sub-groups Using I.Q. Scores

Group	df	within	df	between	Ftest	P
IQ 90 -	14	308.87	1	41.68	.134	NS
IQ 90-110	62	15.85	1	35.67	2.25	NS
IQ 110+	26	15.94	1	40.78	2.55	NS

"As the results of Table 3 indicate, there were no significant differences between the sub-groups. While no conclusion can be drawn from this data it would tentatively appear that the cumulative effect of increased stimuli has a positive effect on total group performance but not on the performance of selected sub-groups."

5. Interpretations

"Considering ... prior research findings with the findings presented in this paper, the FACT systems [sic] use as a planning aid for the improvement of instruction passes tests of both practicability [sic] and usefulness for the classroom teacher."

Abstractor's Comments

This is presented in two sections: (1) the published report, and (2) the investigation.

1. The published report.

The incidence of grammatical (and typographical?) errors is *appalling*.

In Table 2, should not df = 106 rather than 107? Is this a typo, or is it a statistical error?

Table 2 includes helpful descriptive information: means and variances for the treatment groups (although additional information could have been included to advantage; e.g., the standard error of the difference between the treatment means). Why was not similar information included in Table 3?

In Table 2 the reported significance of the difference between treatment means was based upon a t-test, but in Table 3 each reported (non)significance of the difference between treatment means was based upon an F-test. Why were different significance tests reported for $\overline{LSG}$ vs. $\overline{HSG}$ comparisons in the two tables?

Why was virtually no information reported regarding the content of the "Geometry" unit? An inkling of this content may be gleaned from the three illustrative (representative?) posttest items (Figure 2), but this is far from adequate. (In passing, how suitable is such content for seven-, eight-, and nine-year-old children?)

4) Name all the line segments you see in this figure.

11) Construct a right angle using your compass and ruler.

15) Put an X on the pictures of simple closed curves.

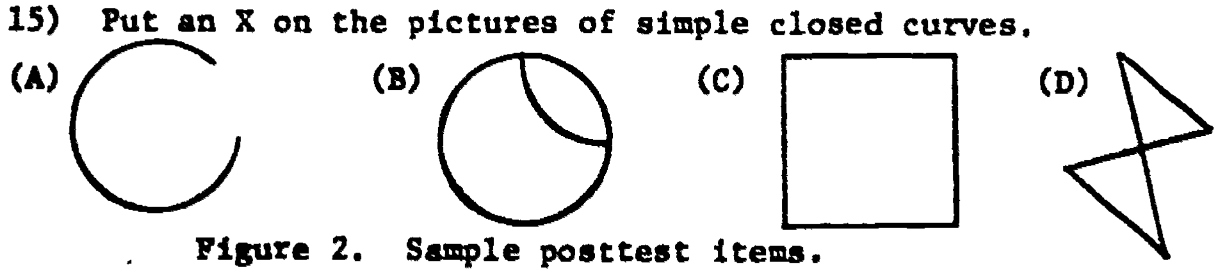


Figure 2. Sample posttest items.

2. <u>The investigation</u>

Many questions could be raised about the sample, the design of the investigation, and the statistical treatment of the data. I shall refrain from doing so, preferring to focus upon an issue pertaining to the conceptualization of the investigation.

Figure 3 was used by the researcher to inform the reader of principal distinctions between the HSG and LSG treatments.

Lesson

	1		2		3		4		5		6		7		8		9	
STIMULI	HSG	LSG	HSG	LSG	HSG	LSG	HSG	LSG	HSG	LSG	HSG	LSG	HSG	LSG	HSG	LSG	HSG	LSG
VC					X		X											
VR	X	X			X		X		X		X	X	X				X	
VA	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
AC																		
AR	X				X	X	X				X	X			X		X	
AA	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
TC																		
TR	X		X	X	X	X	X		X		X	X	X	X			X	X
TA																		
N	5	3	3	3	6	4	6	2	4	2	5	5	4	3	3	2	5	3

Figure 3. Comparison of Stimuli Used in Each Lesson by the HSG and LSG.

It appears that there was *no* treatment distinction in the case of lessons 2 and 6. Why (or why not)?

Treatment distinctions in all other lessons appear to be attributable to HSG lessons having one or more *stimulus combinations* not included in corresponding LSG lessons. In no LSG lesson was there a stimulus combination that was not also included within the companion HSG lesson.

If increased stimuli truly influence posttest performance, is the effect due to the presence of "new" stimulus combinations, or simply to the presence of a greater number of stimuli (regardless of the stimulus combination(s) involved)? The investigation as designed provides no answer to this question, which I consider to be a nontrivial one.

It was obvious in more ways than one that the researcher was far more interested in support for the F.A.C.T. scheme than in "the acquisition of introductory geometric concepts by elementary school children."

Gullen, George E. SET COMPARISON TACTICS AND STRATEGIES OF CHILDREN IN KINDERGARTEN, FIRST GRADE, AND SECOND GRADE. Journal for Research in Mathematics Education 9: 349-360; November 1978.

Abstract and comments prepared for I.M.E. by CHRISTIAN R. HIRSCH, Western Michigan University.

1. Purpose

The primary purpose of this study was to identify the tactics and strategies used by children in grades K-2 when making a sequence of set comparisons in terms of same or greater cardinal number. The study also examined relationships between the frequencies of tactics and strategies used and such variables as grade level, quantification ability, and the magnitude of the cardinal numbers of the compared sets.

2. Rationale

Research by Brownell (1928, 1941, 1963) has demonstrated the importance of investigating tactics and strategies used by young children in dealing with number situations. Most of the reported studies of tactics and strategies associated with the cardinal number behaviors of children have focused on quantification behaviors. Previous research studies on children's set comparison behaviors have identified the use of length, number, correspondence, density, proximity, and guessing tactics. There is a paucity of research with respect to set comparison strategies. Gullen's study is an attempt to fill this void.

3. Research Design and Procedures

The sample consisted of 30 children randomly selected from each of the grades K-2 in the Dearborn Public Schools. The results of an investigator-developed pretest indicated that all 90 subjects understood the terms that were to be used in the administration of the set comparison tasks.

Three versions of a set comparison test (SCT), each consisting of 12 display cards of pairs of sets, were constructed by the experimenter. Each card displayed a pair of uniformly-spaced parallel rows of paper dots. The rows were depicted with unequal lengths and unequal densities.

Thin lines indicating a one-to-one correspondence between elements of the two rows or between the elements of the less numerous row and a proper subset of the other set were shown on each card. The three versions of the SCT differed only with respect to the magnitude of the cardinal numbers of the sets being compared. Version 1 displayed every possible pairing of sets with 3, 4, or 5 elements; Version 2 displayed corresponding pairings of sets with 8, 9, or 10 elements; and Version 3 displayed all possible pairings of sets with 13, 14, or 15 elements. Equivalent pairs of sets were displayed twice in each version.

Ten children from each grade level were selected at random and given one of the three versions of the SCT. Each subject was shown the 12 cards of the assigned version one at a time and asked by the same interviewer to compare the sets in terms of "more" or "same number" of circles and describe the tactic ("How did you know?") used for each comparison.

To test the dependence of correct set comparison and correct quantification, an investigator-developed Quantification Test (QT) was used. Each subject was asked to quantify nine sets (rows of dots) that had the same cardinal numbers as the rows of dots in the three versions of the SCT.

Chi-square tests were used to analyze total frequencies of set comparison tactics, frequencies of tactics by version of the SCT, total frequencies by grade level, and frequencies of correct set comparisons on the SCT by ability correctly to quantify the corresponding two sets on the QT.

Set comparison strategies (patterns of tactics within individual subjects) were identified using factor analysis (Q-technique) as well as joining, ditto, and K-means clustering algorithms. Chi-square tests were used to analyze frequencies of set comparison strategies by grade level and by version of the SCT.

4. Findings

Analysis of the data indicated that the subjects employed number, correspondence, length, density, and guess tactics when making set comparisons. Total frequencies of tactics used were significantly different from one another ($p < .005$) and the frequencies of the use of a particular tactic were found to be dependent on both the magnitude of the cardinal numbers of the compared sets ($p < .005$) and on the school grade level ($p < .005$). To compare sets with 3, 4, or 5 elements, subjects used the number tactic most frequently. The correspondence tactic was used most frequently to compare sets with cardinality 8, 9, or 10. Subjects in grades 1 and 2 used the correspondence tactic more often than all other tactics on all three versions of the SCT. Kindergarteners used the number tactic most frequently when comparing sets with cardinality 3, 4, or 5, but preferred the length tactic when making other set comparisons. The numbers of correct set comparisons were found to be related ($p < .005$) to the subjects' abilities correctly to quantify the sets being compared.

Each of five factor analyses yielded four factors that identified strategies consisting of "nearly exclusive use of a single set comparison tactic." These strategies were correspondence, number, length, and density. Application of the clustering algorithms to the data resulted in 12 clusters within each of which the patterns of tactics were similar. The description of each cluster pattern was used to define a set comparison strategy. Frequencies of the four most correct strategies were combined for each grade level and for each version of the SCT as were the frequencies of the remaining eight strategies. These frequencies were found to be related to both grade level ($p < .005$) and the cardinality of the compared sets ($p < .01$). Kindergarteners used an L-strategy (length tactic used nine times and no other tactic used more than twice) most frequently. First and second graders used a C-strategy (correspondence tactic used on at least 10 of 12 comparisons) most often.

5. Interpretations

First and second graders relied primarily on the correspondence tactic and consequently the C-strategy when making a sequence of set comparisons. On the other hand, kindergarteners used the correspondence tactic in only about 7 percent of the set comparisons. Not one kindergartener used the C-strategy. This suggests that the lines of correspondence on the display cards had little meaning for these children. The frequency of the L-strategy increased with the magnitude of the cardinal numbers of the compared sets. More than 40 percent of all subjects used the unreliable L-strategy when comparing sets whose cardinal numbers were 13, 14, or 15.

Further research could examine the relationships of set comparison tactics and strategies to such variables as levels of quantitative thinking and language development, cognitive style, affective factors, or differences in children's mathematical education.

Abstractor's Comments

This status study is valuable in that it contributes further to our knowledge of the cardinal number behaviors of young children. The study appears well-conceived and carefully executed. The statistical techniques employed seem appropriate and the results are clearly communicated. The author was, however, extremely cautious about interpreting results. Hypotheses for the observed differences were not offered.

It is not clear from the report exactly when during the school year the subjects were interviewed. If it was at the beginning of the year, kindergarteners may have had no experiential basis for the correspondence tactic. Moreover, since children at this stage operate on the basis of the "amount of space occupied", it is not surprising that the length tactic was the most popular unreliable tactic at each grade level. While the author states that the display cards in the SCT were not biased against the use of any one tactic, it appears that length and density offered stronger perceptual cues than did the "thin" correspondence lines.

It is unfortunate that the nature of the previous mathematical experiences of the first and second graders was not investigated. Were the first graders provided with appropriate pre-number experiences in kindergarten? Other questions that may be beneficial to further investigation include:

1) Would vertical or oblique arrangements of parallel rows of dots affect the set comparison tactics (strategies) used?
2) Would the nature of the objects in the sets to be compared affect the tactics (strategies) employed?

Hobbs, Charles E. AN ANALYSIS OF FIFTH-GRADE STUDENTS' PERFORMANCE WHEN SOLVING SELECTED OPEN DISTRIBUTIVE SENTENCES. Technical Report No. 397, Wisconsin Research and Development Center for Cognitive Learning, September 1976.

Abstract and comments prepared for I.M.E. by CAROL A. SMITH and CHARLES E. LAMB, The University of Texas at Austin.

1. Purpose

The purpose of this study was to focus on the performance of fifth-grade students as they solved selected open distributive sentences in relation to three factors (open sentence types, context, number size). Solution methods attempted, as well as errors made, were then identified and classified.

2. Rationale

Assuming that use of the distributive principle is helpful in the study of mathematical topics such as multiplication, the author did a sequential review of selected research to give a foundation for the present study. A number of investigations (Crawford, 1964; Gray, 1965; Schell, 1964) had noted the occurrence of a variety of events which indicate that elementary school students have difficulty in working with the distributive principle. Other investigations (Weaver, 1971, 1973a; Grouws, 1971, 1972) had reported information concerning the performance of young children in solving selected open sentences. Using these studies as a base, the author perceived a need for study of children's performance in solving open distributive sentences. To be included was consideration of solution methods attempted and errors made.

3. Research Design and Procedures

Performance on a recognition level test was used to assign 301 fifth-grade students to one of four levels of recognition of the distributive principle. These students were selected from three elementary schools and all students were judged to be of middle-class socioeconomic status. Twenty students from each of the four levels of recognition were selected as subjects for the investigation. Each student was presented four open distributive sentences in one of the four context levels [symbolic (S); symbolic and pictorial (SP); symbolic and verbal (SV); symbolic, pictoral, and verbal (SPV)], with each sentence reflecting a different combination of the four possible combinations of two levels of Open-Sentence Type [left-horizontal (L-H); right-horizontal (R-H)] and two levels of Number Size [large (L); Small (T)]. Performance on the four-item interview test along with an individual interview provided data for the study.

The experimental design was a split-plot design with repeated measures on the factor levels of open sentence type and number size.

4. Findings

(a) No significant difference was found between boys' and girls' overall performance. Therefore, the data were pooled across sexes.

(b) Solution methods used by students during the investigation were identified and classified. There were two correct and six incorrect methods listed. Correct solution methods were selected 76 times out of a possible 320 items. Incorrect solution methods were selected 244 times. The most frequently selected incorrect solution method (method 3) was picked 198 times.

[Solution Method 3:

$$(3 \times 9) + (3 \times 5) = (\underline{27} \times \underline{15})$$

The two indicated products were used as a solution pair.]

There were 23 additional errors in solution; they were caused by misuse of the distributive principle and computational errors.

(c) There was no indication of significant differential student performance due to open sentence type or to context where the dependent variable was number of correct solution, use of the distributive principle, or use of solution method 3.

(d) There was no indication of significant differential student performance due to number size, when the dependent variable was number of correct solution or use of the distributive principle.

(e) There was an indication of differential performance due to number size when the dependent variable was use of solution method 3. More students used solution method 3 when small numbers were used.

5. Interpretations

(a) The author states, "Due to the low reliability of the Recognition Level Test and due to students' lack of using the distributive principle to any great extent, the relationships between students' use of the distributive principle could not be determined with any reasonable confidence."

(b) Two major conclusions/interpretations were stated:

(1) Results do not indicate whether or not elementary school students have an informal "idea" concerning distributivity and cannot recognize and/or use this principle in the formal ways presented in this study.

(2) The elementary school may not be the proper place for teaching the distributive principle. Maybe it should be postponed until a higher level of mathematics maturity is reached. If this postponement occurred, a restructuring of the presentation of the multiplication algorithm would need to be considered.

Abstractor's Comments

1. In general, the author was very thorough in reporting details relevant to the study. However, in discussing Solution Method 3, the author describes an attempt to determine why students chose this method so frequently. The discussion at this point is confusing and vague. Why do subjects pick this incorrect solution method?
2. The apprehension expressed concerning the reliability of the Recognition Level Test is well founded. The author's discussion of this topic is rather hazy. Perhaps more items should have been used.
3. The question of using principles such as distributivity in elementary school mathematics is an important one. Considering the responses of children to Hobbs' items can provide insights to the

interested reader. Consistency of Fifth-Grade Students' Solution Methods for Certain Open Distributive Sentences: Hobbs' Data Revisited. Project Paper 76-5 by J. Fred Weaver from the Wisconsin Research and Development Center should also be of interest.

4. The fact that a majority of children using the distributive principle came from a particular school suggests a closer look at the situation. More students in the sample came from that school and that school used an "individualized" program. Does this explain what happened?

Mayer, Richard E. EFFECTS OF PRIOR TESTLIKE EVENTS AND MEANINGFULNESS OF INFORMATION ON NUMERIC AND COMPARATIVE REASONING. Journal of Educational Psychology 70: 29-38; February 1978.

Abstract aud comments prepared for I.M.E. by WALTER SZETELA, University of British Columbia.

1. Purpose

The purpose of the study was to investigate differences in the ability of university students to respond to numeric and comparative type questions after they had been conditioned to respond to numeric or comparative type questions where all subjects were presented with quantitative information in either equation format or story format.

2. Rationale

It is suggested that two factors may influence the storage of information needed to solve problems: (1) expectation of whether the data will be used to make numeric judgments or comparative judgments and (2) the context or meaningfulness of the data. It was hypothesized that subjects expecting numeric questions build cognitive structures that include more detail. Such subjects would excel on questions involving retention. Contrastingly, subjects expecting comparative questions build cognitive structures with less detail but more integration of information. Such subjects would excel on questions involving inference. With respect to meaningfulness of the information, it was hypothesized that it is easier to make inferences from information presented in story form than in equation form. The investigator suggested that meaningful information is assimilated to the subject's past experience and is restructured and integrated, whereas information in equation form is added to memory rather than restructured to expedite inferences.

3. Research Design and Procedures

Two separate experiments were done. In each experiment were 80 university students from an introductory psychology course. The design of both experiments was identical. The 2x2 factorial design with 20 students in each cell included the factors of problem set (numeric vs.

comparative) and information format (story vs. equation). Subjects were presented with five four-term linear orderings (e.g., $A > B > C > D$) which incorporated quantitative relationships in the two formats exemplified as follows:

Story: In a certain forest the animals are voting for their leader. The frog gets twice as many votes as the hawk. The hawk gets five times as many votes as the rabbit. The rabbit gets four times as many votes as the bear.

Equation: $F = 2 \times H$, $H = 5 \times R$, $R = 4 \times B$.

After each of the first four linear-ordered passages, each subject received either 12 numeric questions or 12 comparative questions, but after the fifth (target) passage all subjects received 12 numeric and 12 comparative questions. Subjects were randomly assigned to treatment and tested in groups of up to eight per session. Each subject was seated in a booth with a response box controlled by an IBM 1800 computer. Subjects pressed a button to see a passage, pressed again when finished reading, and pressed again to answer a question "YES" or "NO." After each answer a new question appeared until the set was finished, and each subject then read his or her score in percentages on the screen. For each passage, two sets of 24 questions were constructed, one set expressed in words and one set expressed in letters from the equations. The questions were based on a 2x2x2x3 factorial design with factors (a) numeric vs. comparative question; (b) correct answer "YES" or "NO"; (c) question about adjacent links such as A-B, B-C, or C-D or about remote links such as A-C, B-D, or A-d; and (d) which particular link was involved. The second experiment differed from the first by presenting the information in a more complex way using remote pairs only, so that attention to quantitative relations was needed to make even nonquantitative inferences. An example of a story question for the fifth (target) passage is as follows:

Comparative-Positive-Remote (B-D): Does the hawk get more votes than the bear?

4. <u>Findings</u>

For each experiment the proportion of correct responses to comparative/numeric, yes/no, and retention/inference questions were reported for the Comparative-Story, Numeric-Story, Comparative-Equation, and Numeric-Equation groups.

In both experiments:

(a) Performance on comparative questions was better overall than on numeric questions, $p < .001$, $p < .005$.

(b) There was no difference in inference ability among the four treatment groups.

(c) Information in story form gave better overall performance than in equation form, $p < .005$, $p < .001$.

(d) Numeric questions elicited more "no" answers than comparative questions, $p < .001$, $p < .001$.

(e) There was an interaction between problem-solving set and format in which the numeric-equation group performed worse than the comparative-story group, $p < .05$, $p < .05$.

(f) Subjects who read stories did better on comparative questions than on numeric questions, $p < .05$, $p < .005$.

Among different results in Experiments 1 and 2 it was found that:

(g) Overall performance in Experiment 2 was inferior to that in Experiment 1, $p < .001$.

(h) There was a higher tendency to answer "no" in Experiment 2.

(i) In Experiment 2, performance on inference questions was better than on retention questions with respect to the same comparison in Experiment 1, $p < .001$.

5. <u>Interpretations</u>

The author concludes that:

(a) Making inferences with numeric and comparative information are distinctly different processes rather than extensions of one another.

(b) Subjects expecting to answer numeric questions encode information in more detail than subjects expecting comparative problems.

(c) There was no support for the hypothesis that different problem solving sets lead to acquisition of qualitatively different structures. For example, in Experiment 1 both problem-solving sets produced an effect similar to the finding of Potts (1972), in which inference questions were at least as easy as retention questions. The author suggests that all subjects acquire a spatially arranged linear ordering, but that numeric-set subjects include tags indicating quantitative distance.

(d) There was support for Mayer and Greeno's (1975) idea that equations are more likely to be stored in a way to retain original presentation detail, while stories are assimilated and integrated. For example, equation subjects, especially with numeric set, performed better on problems requiring retention of exact quantities.

(e) There is evidence that equation and story subjects store information in qualitatively different ways. In the complex situation of Experiment 2, story groups performed better on inference equations than equation groups.

(f) Presentation organization (e.g., simple linear in Experiment 1 vs. complex in Experiment 2) influences the way information is structured in memory. Complex organization leads to information which is more likely to be integrated in less detail. It is suggested that the Potts (1972) effect may be limited to nonquantitative, simple material.

(g) In mathematics instruction these results suggest that past experience with solving only calculation problems, especially as presented in low-meaning contexts such as equations, results in a rigid, specific encoding strategy and poorer transfer to more interpretive problems.

Abstractor's Comments

This complex study was carefully designed and executed. Two similar experiments differing only in the complexity of the presentation of information to the subjects were carried out. The number of subjects and procedures were otherwise identical. Subjects were randomly assigned to treatments and all cell sizes were equal. Unfortunately, the complexity of the study, a combination of a 2x2 factorial design with four treatment groups coupled with the 2x2x2x3 factorial design of the questions, presents the reader with a maze of data, difficult to assimilate. Furthermore, there were two separate question forms for the 2x2x2x3 factorial question design, which exacerbates the difficulty of comprehension. The reader struggles to separate comparisons between comparative and numeric set subjects doing comparative and numeric problems in story or equation form, and of inference or retention type. If the study were not so well designed, comprehension of the results would be even more difficult. Notwithstanding the careful design and execution, there are some concerns about the study.

(a) With so many results, at least 33 reported, it seems presumptuous to mention a marginally significant three-way interaction among set, format, and answer, $p < .06$. It is even more presumptuous to say that subjects using equations showed higher probability of saying "no" than subjects using story format, $p < .10$. The author omitted the probability level for the Experiment by Answer (yes/no) interaction.

(b) It is inappropriate to refer to the two formats as "meaningful" stories and "nonsense" equations.

(c) No reliability data are given for the measurement instruments.

(d) The author comments that high error rates complicated the interpretability of response time data so they were not treated, but that several analyses were performed on various partitions of the proportion-correct data. Could these partitions have been selected more appropriately?

(e) It is not clear what the author means when he says that "making inferences with numeric information and comparative information are distinctly different processes." All the information was quantitative whether presented in equation or story form. The subsequent questions were either numeric or comparative, not the information.

(f) One wonders why the author did not simplify such a complex study. The 2x2x2x3 factorial design for questions is ponderous. Was the inclusion of a yes/no answer analysis worthwhile? Even more questionable was the inclusion of the three types of remote link questions. The author did not (thankfully) include results and analysis of this factor.

As this study was conducted with university students, the relevance of the voluminous results to mathematics education is reduced. It is not surprising that subjects with a numeric or comparative problem-solving set did better on problems to which they had been conditioned. The result that subjects seemed to assimilate information in story form better than in equation form is important. Yet it is paradoxical, as any elementary or secondary school teacher knows, that students have difficulty with story problems. The author's statement that "solving only calculation problems presented in low meaning contexts like equations results in rigid encoding strategy and poorer transfer to interpretive problems" deserves serious study. Are students poor problem solvers because they have been conditioned to so many pure calculation problems with little meaning? To what extent have students developed a negative set toward story problems which may interfere with reasonable efforts to assimilate information to solve problems? Replication of this study in greatly simplified form with elementary and secondary school students with the focus on the story/equation factor might well provide useful problem-solving information.

References

Mayer, R. E. and Greeno, J. C. Effects of meaningfulness and organization of problem solving and computational judgments. Memory and Cognition, 1975, 3, 356-362.

Potts, G. R. Information processing strategies used in the encoding of linear ordering. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 727-740.

Pike, Ruth and Olson, David R. A QUESTION OF MORE OR LESS. Child Development 48: 579-586; June 1977.

Abstract and comments prepared for I.M.E. by LOYE Y. HOLLIS, University of Houston.

1. Purpose

The purposes of the study were to:

a. Investigate the mental representations that children assign to events in which addition or subtraction operations occur and examine the relationship between these representations and their linguistic and cognitive competence.

b. Examine the childrens' performance when the representations were biased by the experimenter (p. 579).

2. Rationale

Recent studies suggest there are four stages in the process of answering a question or verifying a sentence: setting up a representation of the event; setting up a representation of the sentence or question; a series of comparison operations including truth index charges or recording operations in the event of a mismatch; and making a response. It is well known that children's performance on tasks involving sentence comprehension varies considerably from that of adults. An analysis of the comprehension process should indicate whether the differences between more and less developmentally advanced children lie in the original representation assigned to events and/or questions or in the different set of procedures for comparing those representations. The study examined these alternatives in regard to children's judgments of more and less (p. 579).

3. Research Design and Procedures

Experiment 1

The subjects were 36 English-speaking children composed of 17 kindergarten children with an average age of 5-2 (8 male and 9 female) and 19 second-grade children with an average age of 7-2 (10 male and 9 female). Subjects were tested individually in the order as follows:

a. The language production test. The child was shown a drawing of two glasses containing unequal amounts of brown liquid. Subjects were asked which they would choose if the glasses contained "Coke or better tasting stuff." A justification for the choice was requested. The conversation was taped and analyzed for occurrences of more and less.

b. The question-answering task. The materials consisted of two opaque beakers and 33 marbles. When the subject agreed that both beakers contained an equal number of marbles, the experimenter added an extra marble or subtracted one from the collection and asked, "Who has more (less)?" Four problem types were generalized by combining the operations of addition and subtraction (manipulation) with the questions of more and less. The experimenter recorded answers and latency of response, timed from the last word of the question.

c. The operation test. The materials were those used in the question-answering test with the addition of a lid for the beakers. The child was reminded of the prior test and the marbles were put in equal collections. The experimenters then placed a lid on the beakers on which the child had previously performed the operation and asked, "Now if you can't touch that pile, what can you do to still make it so that mine (yours) has more (less)?" If the child was not successful, it was suggested that he or she think about what the experimenter had done before. When no attempt was forthcoming, the experimenter continued with the next trial. No corrections were made. Subjects had one trial on each of the four instructions and all responses were recorded.

Experiment 2

In this experiment an attempt was made to influence the children in the fail-operations group to set up a representation of a quantitative relationship that was equivalent to that presumably assigned by the pass-operations children. All children were given the following test about one week after the test in Experiment 1.

Simple time drawings which showed a boy and a girl each holding identical containers with unequal amounts of liquid were shown to the children. The children were given biasing instructions prior to the presentation of the picture: "Pay attention to who has more (less)." The four problem types were presented with all four pictures.

Following practice trials, children were given four trials for each of the problem types. No correction procedure was used for the experimental trials. The answers and latency of response were recorded as in Experiment 1.

4. Findings

Experiment 1

The language production test showed that the relationship between grade and results on the test was not significant.

The operations test showed that success on this test was related to grade level. Also, the operations test was shown to be significantly more difficult than the language production test.

Children in grade 2 responded more quickly than kindergarten children, and responses to tasks involving the operation of addition were faster than responses to tasks involving subtraction.

"More" questions were easier than "less" questions and problems involving subtraction.

The comprehensive process involved in the question-answering task used in this experiment was fundamentally the same for all children.

The childrens' performance on the operations test lends further substantiation to the theory that the results are related to the way in which children encode the events.

Experiment 2

The pass-operations children were faster overall than other children, but otherwise their pattern of performances was the same.

When provided with an appropriate representation, the fail-operations children performed exactly like the pass-operations children.

5. Interpretations

Developmental differences seem to reflect differences in the representations children assign to events rather than the subsequent comparison procedures they carry out in the derivation of answers to questions.

The nature of the representations seems to be related more closely to the cognitive than to the linguistic competence of the child.

Whereas the younger children define their terms and represent their knowledge in terms of extensional properties, so that "more" is specified in terms of a perceptual property such as length or on action such as adding, the older children define their terms and represent their knowledge in terms of intentional properties, so that "more" contrasts with "less" and with "same."

Abstractor's Comments

The article reported a study that was well designed and executed. Careful thought seems to have been given to the analysis of the data. For the most part, the authors are to be complimented.

The problem with the article (for me as a mathematics educator) is, "What does it mean with respect to the teaching/learning of arithmetic?" The authors provide little insight into that question and that is unfortunate. The study would appear to provide suggestions concerning sequence of content, presentation of content, and time expectations for mastery of content. In my view, the authors should have dealt with these rather than leaving them to the fantasies of the reader.

Sagiv, Abraham. HIERARCHICAL STRUCTURE OF LEARNER ABILITIES IN VERBAL COMPUTATION PROBLEM-SOLVING RELATED TO STRENGTH OF MATERIAL. International Journal of Mathematical Education in Science and Technology 9: 451-456; November 1978.

Abstract and comments prepared for I.M.E. by FRANK K. LESTER, JR., and JOE GAROFALO, Indiana University.

1. Purpose

The purpose was to establish theoretical and empirical support for the existence of a hierarchical structure of abilities in solving verbal computational problems.

2. Rationale

Among the many variables related to the ability to solve mathematics problems, those which appear to be most strongly related to ability are variables associated with understanding a problem and using data to solve a problem. Particular attention has been given in the psychology literature to variables which can be considered as prerequisites to problem solving. Notable in this literature are the hierarchical models of Gagné, Glaser, and Ausubel. These models suggest that the main factors influencing the ability to solve verbal computational problems are the problem prerequisites (components). With these models as a guide, it is reasonable to make the general hypothesis "... that the score of a stage j (S_j) [of the hierarchy of problem components] is equal to the sum of the products of its component scores s_i ($i = 1,2, \ldots k$) and their respective path coefficients (P_{ij})." That is, for k influencing components, $S_j + \sum_{i=1}^{k} P_{ij} \cdot s_i$

3. Research Design and Procedures

Three problem components were considered as being prerequisites to verbal computational problem-solving ability: (1) comprehension of concepts and laws; (2) ability to derive an equation for solving a one variable problem; and (3) mathematical ability (defined to be the average of algebra, numerical computation, and units computation abilities). The study sought to determine the amount of variance in problem solving which could be accounted for by these three components.

The sample contained 266 tenth- and eleventh-grade intermediate-ability students in 14 classes in 14 schools throughout Israel. Students were tested in mathematical content which they had recently completed studying (viz., "strength of material"--not described in the report). Achievement test items were randomly chosen from an item bank which had been developed to cover this particular content. Tests for each sub-area of the topic were devised. Each test contained four sections: items to test comprehension of concepts and laws, algebraic exercises, computational exercises, and verbal computational problems.

Path analysis was used to test the significance of the following hypotheses derived from the general hypothesis:

H_1: The components for ability to derive a formula (FORM D) are mathematical ability (MATH) and comprehension of concepts and laws (COMP). That is,

$$\text{FORM D} = P_{\text{Math, Form D}} \cdot \text{MATH} + P_{\text{Comp, Form D}} \cdot \text{COMP}$$

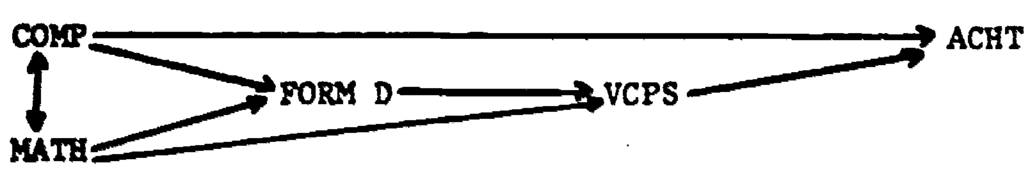
H_2: The components for ability to solve a verbal computational problem (VCPS) are mathematical ability (MATH) and the ability to derive a formula (FORM D). That is,

$$\text{VCPS} = P_{\text{Math, vcps}} \cdot \text{MATH} + P_{\text{Form D, vcps}} \cdot \text{FORM D}$$

H_3: The components for the achievement test (ACHT) are ability to solve a verbal computational problem (VCPS) and comprehension of concepts and laws (COMP). That is,

$$\text{ACHT} = P_{\text{vcps, acht}} \cdot \text{VCPS} + P_{\text{comp, acht}} \cdot \text{COMP}$$

The relationship among these variables is described diagrammatically as follows:



("$a \rightarrow b$" is interpreted as "*a* may cause *b* but *b* does not cause *a*"; the bi-directional arrow represents unanalyzed correlation.)

4. Findings

The author suggests that the principal factors which influence ability to solve verbal computational problems in "strength of material" were determined and were organized into a hierarchical causal structure. Three hypothetical relations were found to be significant.

(a) Mathematical ability and comprehension of concepts and laws account for 67 percent of the variance in the ability to derive a formula for the solution of one variable.

(b) Mathematical ability and the ability to derive a formula account for 89 percent of the variance in verbal computational problem solving.

(c) Verbal computational problem-solving ability and comprehension of concepts and laws account for 92 percent of the variance in the overall achievement test score.

5. Interpretations

The fact that ability to derive a formula is based on mathematical ability suggests that special emphasis should be placed on in-class practice in formula derivation while reserving the actual numerical calculations for homework. It appears that there are components in addition to mathematical ability and comprehension of concepts and laws which influence ability to derive a formula. However, the hypothesized predictor components for verbal computational problem solving and overall achievement do account for the major portion of the variance.

Abstractor's Comments

An attempt to test the validity of theoretical models which hypothesize a hierarchical structure for the abilities involved in problem solving is commendable. Also, the use of path analysis, a popular statistical technique in the social sciences but infrequently used in education, has potential for providing valuable information about possible linear relationships among a set of variables. Thus, path analysis is especially well suited to testing hierarchical models.

Unfortunately, there is little else positive that can be said about this research. The most serious weakness lies with the very sketchy and vague manner in which the research is reported. This sketchiness and lack of clarity affect the quality of the report in two areas.

First, non-standard terms are neither defined nor described and the methods and procedures of the study are very unclear in places. For example, the author claims the content covered by the achievement test "... was divided into sub-areas according to the basic deformations." No indication is provided as to what these "basic deformations" are or how many of them there are.

Also of concern is the statistical analysis. Path analysis is not a method for showing causality, but rather a procedure for examining hypotheses concerning a logically pre-determined weak causal ordering among a set of variables. One assumption needed in this regard is that of no reciprocal causal relationships between variables. In this study, it is not clear whether or not the causal ordering was hypothesized a priori nor is it apparent that the variables can satisfy the assumption of non-reciprocity.

The test description seems to indicate that the predictor tests are comprised of subsets of items from the total achievement test. The lack of adequate description of the predictor tests raises the question of whether or not these subsets are disjoint (i.e., does the computational component of the mathematical ability score represent the computational performance on the VCPS items as well as the performance on the computation items?) If the criterion tests (VCPS, ACHT) are indeed contaminated with items from their predictors, then both the causal ordering and the magnitude of the path coefficients are suspect. These points should be kept in mind when interpreting these coefficients.

Sepie, A. C. and Keeling, B. THE RELATIONSHIP BETWEEN TYPES OF ANXIETY AND UNDER-ACHIEVEMENT IN MATHEMATICS. Journal of Educational Research 72: 15-19; September/October 1978.

Abstract and comments prepared for I.M.E. by FRANK F. MATTHEWS, University of Houston CLC.

1. Purpose

The intent of the study was to determine if under-achievers in mathematics are differentiated from their peers on measures of general anxiety, test anxiety, and mathematics anxiety. It was hypothesized that under-achievers would score highest on measured anxiety.

2. Rationale

Moderate to high levels of anxiety are known to have a debilitating effect on achievement. In addition, some investigators have reported significant negative correlations between measures of mathematics anxiety or test anxiety and achievement, especially at post-secondary levels. It is reasonable to propose a relationship between such anxiety and over/under-achievement i. mathematics, and to investigate such a relationship at younger ages.

3. Research Design and Procedures

The study first classified students into three categories: under-achiever, achiever, and over-achiever. On each of three anxiety measures, the investigators studied whether the scores differed between the groups.

Sample. The sample included 246 children who are approximately 11+ years old. They comprise the population of one grade (Form I) in a New Zealand intermediate school.

Instruments. Ability was measured using the Otis Test of Mental Ability - Intermediate Level -- Form A. This instrument was administered at the end of the preceding school year (grade 5 equivalent). [This test appears to be the Otis-Lennon Mental Ability Test which does come in two forms (J & K) at the intermediate level. However, this does not appear to be the proper level for the age group involved. Given the possibility of a foreign publisher, the exact instrument is in doubt.]

Achievement was measured by sections of the Progressive Achievement Tests - Mathematics, which were developed in New Zealand. [This is a high-quality instrument with a good variety of items at different levels Although its utility is limited in that it was developed particularly for the New Zealand curriculum, the general approach is excellent.]

Three measures were obtained for various types of anxiety. Two were by Sarason et al., the General Anxiety Scale for Children (GASC) (34 yes/no responses) and the Test Anxiety Scale for Children (TASC) (30 yes/no responses). The third was developed by the investigators (MASC) (20 yes/no responses) for which two sample items were listed. Beyond a split-half reliability coefficient of 0.90 for a subset of 50 students, no further information is provided on the instrument.

Classification. Following Thorndike's and Farquhar & Payne's Stage II model for identifying achievement levels, regression equations were developed to predict achievement (PAT) score ($\tilde{Y}$) separately for each sex. Each individual was then identified as an overachiever if $(Y - \tilde{Y}) > S_{yx}$ (where S_{yx} is the standard error of estimate); an achiever if $|Y - \tilde{Y}| \leq S_{yx}$; and an under-achiever if $(\tilde{Y} - Y) > S_{yx}$. This placed 15 percent of the sample in each extreme category.

4. Findings

Correlations between the variables and three univariate (achievement level x sex) analyses of variance were run, one with each of the anxiety scores as the dependent variable. There were no significant interaction effects. Main effects were found only for level on the MASC and sex on the GASC. The level results for the GASC and TASC were in the predicted direction, but were not statistically significant.

5. Interpretations

The results tend to confirm that mathematics anxiety is related to achievement in mathematics. MASC was the only variable to provide negative correlation for both sexes and it is significant in both cases (-0.28 and -0.30). In addition, MASC was the only one of the three tests showing an ANOVA main effect for level. One anomaly of interest is that despite the strong MASC-TASC relationship (r's of 0.73 and 0.72), the one has an effect and the other does not.

Abstractor's Comments

This investigation offers a good introductory look into an aspect of the relationship between anxiety and achievement. This abstractor believes that the general field is one of promise and that this investigation is helpful in developing a view of the problem. As with any introductory study, however, this one suffers from a number of problems. Some of them would have been avoidable, but, unfortunately for all of us, some appear to be inherent in the subject.

The first, and to the abstractor the most unfortunate, was avoidable. The investigators created yet another unpublished mathematics anxiety instrument lacking in psychometric data and even lacking a list of items. A full presentation of the instrument could have been made in less space than is wasted by the presentation format for Table 1. The abstractor recognizes that the investigators desired comparability with the GASC and TASC scales by Sarason et al., but wishes that at least the instrument had been given. The split-half reliability is useful but provides only minimal information. The field of mathematics anxiety in general suffers from a lack of proper instrumentation which is only now beginning to be remedied.

The ability test (if it is the version of the Otis-Lennon mentioned above) is intended for grades 7-9. Yet it was administered at what the investigators give as the end of grade 5. While this may be justified based upon differences between the New Zealand and U.S. educational systems, this is a problem which the investigators should have addressed. It is not necessarily a serious problem given the mean IQ of 98.7 for the group. Another difficulty is the impossibility of positively identifying the instrument from among those currently available. The different names available, the problem of age level, and the apparently invalid form designation are disconcerting despite the abstractor's belief in the identification.

The measurement of achievement was by a single instrument used once. Yet, both of the procedural references cited by the investigators strongly recommend the use of a variety of measures administered over time. The above are specific problems which probably arose from constraints imposed by the situation.

A more serious problem is the use of a g measure for IQ. Many current references measuring intelligence prefer a multifactor approach. This is supported by studies which do not show a single factor, hierarchical structure for instruments which have been investigated. (No factor analytic data is available for the Otis-Lennon.) Yet, even if a multifactor exam had been used, there are still the two choices of using a mathematics scale where the scores are almost certainly already contaminated by anxiety, or of using other scales which the structure implies should be independent of mathematical ability. This dilemma with two unacceptable choices leaves nothing but the use of a g measure where the problem is at least suitably buried, although still serious. It can only be cured by the development of a test of mathematical ability with no perceptable connection with mathematics, an unlikely event.

The use of predicted achievement a la Thorndike and Farquhar and Payne is probably the most appropriate approach to the determination of achievement level. This study does suffer, however, by the deletion of the Stage I screening which Farquhar and Payne used to reduce the incidence of false discrepancies. While this deletion is required by the use of a single achievement score, it does reduce the efficacy of the technique.

While the authors comment on sex differences, this abstractor does not consider any such analyses which use achievement level to be appropriate. The use of separate regression equations in the determination of predicted achievement for classification of achievement levels is fully capable either of creating artifactual results or of suppressing actual results. While the use of such separate procedures and even sex-based norms is common in studies involving intelligence, it does eliminate meaningful discussion of sex differences.

One sex difference which occurs before this process is used and which the investigators do not mention is the TASC-PAT correlation (for males, -0.18; for females, +0.07). This switch, despite the very similar MASC-TASC and MASC-PAT relationships, is one which the abstractor finds interesting.

If the MASC were available, its comparability to the more common GASC and TASC would warrant further psychometric evaluation. In addition, despite the abstractor's conceptual problems with contamination of the predictor variable by anxiety, this approach does appear to be the one which would be productive if investigated further—especially with other and more complete achievement measures.

Smith, Lyle R. and Edmonds, Ed M. TEACHER VAGUENESS AND PUPIL PARTICIPATION IN MATHEMATICS LEARNING. Journal for Research in Mathematics Education 9: 228-232; May 1978.

Abstract and comments prepared for I.M.E. by KENNETH A. RETZER, Illinois State University.

1. Purpose

Investigating the joint effects of three degrees of teacher vagueness and two levels of student participation upon the learning of mathematics, Smith and Edmonds contributed this study to the literature of the individual effects of each teaching behavior.

2. Rationale

Building upon (a) studies which found negative correlations between frequency of teacher vagueness terms and student achievement and (b) studies which reported active participation of mathematics students more effective than passive intake of information, the investigators have explored their joint effects.

3. Research Design and Procedures

Using a 2x3 factorial design in which 34 college psychology and sociology students were randomly assigned to each cell, the investigators used 20-minute videotaped mathematics lessons on sums of consecutive positive integers (SCPIs). Each subject saw definitions, examples, and applications which had been videotaped from overhead projector transparencies while listening to the same person read the script, which contained an identical sequence of concepts, generalizations, and examples. The pupil participation dimension was dichotomized as (a) "passive intake" lessons in which the subjects were shown examples of how to identify SCPIs and write numbers as SCPIs, followed by a detailed analysis of each problem and its solution, and (b) "active participation" lessons in which the teacher procedure was identical except that students were given opportunities to work the problems posed individually before their solutions were presented. The degree of vagueness dimension of the treatments was supplied in each of three pairs of lessons by using 0, 72, or 144 vagueness terms. A frequency distribution of vagueness terms is supplied.

Student comprehension was determined from a 20-item test given immediately after each lesson. The test measured comprehension of (a) the concepts of powers, divisors, and primes; (b) generalizations characterizing a positive integer as an SCPI if and only if it is not a power of 2, and prescribing a procedure for finding the number of different ways a positive integer can be written as an SCPI; and (c) the processes of how to write positive integers which are not powers of 2 as SCPIs.

4. Findings

An analysis of variance revealed a significant difference ($p < .05$) between high vagueness groups and the no-vagueness groups. There were no significant differences in the pupil participation main effect or in the interaction.

5. Interpretations

The authors suggest that the failure of this study to support prior research findings that student participation significantly influences achievement may result from a qualitative difference in participation between the college-age students in this study and the elementary and high school students of other studies. This study does support the conclusions of earlier studies that the frequency of vagueness terms negatively influences achievement. The authors suggest that this study reveals that a high degree of vagueness per se affects achievement because they were able to control the variables correlated with vagueness which were left uncontrolled in prior studies.

Abstractor's Comments

From the perspective of a researcher, it seems that this study was well-conceived, adequately designed and conducted, and clearly reported. It builds upon and extends prior research and provides a basis for further research. An interested researcher might want to write for scripts of the lessons or copies of the test if more details are desired or if other bases for results or conclusions were sought.

A teacher or teacher educator who is a consumer of such research should find the conclusions interesting. Probably no teacher is intentionally vague but this study provides motivation to eliminate vagueness.

Attempts to eliminate vagueness would prove challenging, however. This study operationally defines vagueness as the use of vagueness terms, which are adequately operationally defined for research purposes by the list provided. However, it would not be clear to the practitioner whether one can eliminate vagueness by avoiding use of the terms on that list, making his/her sentences more concise and his/her quantifications more precise, being more conscious of his/her lesson objectives, preparing his/her lessons better, or some other means.

The absence of significant difference in the participation main effect will probably do little to change a practitioner's view with regard to the desirability of pupil participation. Some would agree with the authors that college-age students can actively participate mentally without that participation being overtly observable. Others might criticize the opportunity to attempt individually to solve problems before solutions are demonstrated as hardly what they would call active participation when contrasted with alternatives such as teacher-student dialogue or small-group interaction.

Finally, content specialists might differ with the authors in their characterization of the mathematical content involved in the treatments based on the report. This reviewer would characterize that content as basically teaching a procedure—how to write SCPIs—together with the concept of an SCPI, along with the prerequisite concepts needed to understand the definition of an SCPI and the prescription for writing its name in summation notation. There seem to be no generalizations taught. Using the notion of contrapositive, the first two generalizations listed seem to be better described as equivalent to a definition for an SCPI and the third purported generalization seems to be a prescription for carrying out a procedure which, in itself, might be of dubious value in mastering the criterion procedure of this experiment.

These comments are not intended negatively, however, but serve to demonstrate that in addition to the positive characteristics of the research mentioned above, its content and constructs are sufficiently interesting to warrant discussion among both practitioners and researchers.

Zammarelli, J. and Bolton, N. THE EFFECTS OF PLAY ON MATHEMATICAL CONCEPT FORMATION. British Journal of Educational Psychology 47: 155-161; June 1977.

Abstract and comments prepared for I.M.E. by THOMAS C. O'BRIEN, Southern Illinois University at Edwardsville.

1. Purpose

The aim of this study was to assess the effect of play experience on concept formation involving modulo 4 combinations.

2. Rationale

The work of Kohler and Wertheimer gives implicit support to the idea that play experience can be helpful in problem solving, and the work of Piaget, Bruner, and Dienes stresses the importance of personal activity in intellectual development. Yet only a few studies have attempted to substantiate the importance of play in cognitive advance. Further, Bennett's review points up contradictory evidence on the relation between divergent thinking and mathematical ability. Thus the present research was undertaken to assess the effects of active play on the structural thinking involved in mathematical concept formation.

3. Research Design and Procedures

The Sample

Twenty-four children, aged 10 years 2 months to 12 years 2 months, were selected from the second and third years of a middle school.

The Tasks

The Concept Task. Wooden blocks in four different colors (yellow, green, red, blue) were used. The experimenter showed the child a colored block. The child chose a colored block of his or her own, and the experimenter responded with a third block according to the table below, as well as with a statement such as "Red plus red makes yellow."

	Y	G	R	B
Y	Y	G	R	B
G	G	R	B	Y
R	R	B	Y	G
B	B	Y	G	R

The response block (yellow, in the example above) was then used as the new first block, the child chose a second block, and the experimenter chose the "resultant." The sequence was continued until the child felt that he or she could predict the resultant block for any of the 16 color combinations. At this point a standard verbal test was administered. If an 80 percent success rate was achieved, the child was interviewed and asked how he or she knew which color combination produced the resultant colors. If an 80 percent success rate was not achieved, the child was reassured that there was no penalty and the sequence with the blocks continued until the child was ready for another test.

The Toy. The experimental condition in this research involved a toy version of the table shown above in which children pressed a button in the left column (yielding a colored light) and a button in the top row (yielding a colored light); the combination thus produced a third colored light on the table toy.

The Procedure

Four days after children had completed the initial concept task and were interviewed, each child was given a sheet of paper which listed the 16 color combinations and was asked to list as many of the products as possible, guessing when unsure.

Three groups of 8 children were formed based on a 1-12 ranking scale supplied by the child's mathematics teacher; the ranking compared children with their classmates in terms of mathematical ability.

One group, the Play group, was then given 7½ minutes to play with and explore the toy.

A second group, the Yoke group, was given toys "yoked" to the toys in use and watched the combinations of lights and their products.

A third group had no further experience and returned for testing after 7½ minutes.

Then each child was presented with a 4x4 matrix of colors and asked to complete the matrix. Upon completion, each child was asked how he or she had accomplished the matrix task, the explanations were tape-recorded, and the results independently scored.

<u>The Data</u>

For each child, the following data were available:

1. The teacher's rating of mathematical ability.
2. The number of blocks needed to achieve an 80 percent success rate on the Concept Task.
3. The conceptual level exhibited under interview upon completion of the Concept Task (on a scale of 1 to 4). A score of 1 was given for memorization without awareness of patterns or relationships, 2 for heavy reliance upon memory with some basic relationships mentioned, 3 for several rules and strategies which when used together would predict the result of any combination, and 4 for the production of one rule which could be demonstrated to work in all cases.
4. The number of combinations remembered prior to the treatment.
5. The number of combinations remembered following the treatment.
6. The conceptual level of explanations following the treatment.

The data in 1-4, then, are initial (pre-treatment) scores and those in 5 and 6 are final (post-treatment) scores.

4. Findings

Mann-Whitney U Tests showed no statistically significant differences among the three groups on the four initial measures. The differences between the Play group and other groups on the final measures were statistically significant at the .02 level or better. The means were as follows:

		1	2	3	4	5	6
Group 1	Play	7.87	242	1.37	9.50	15.75	3.43
Group 2	Yoked	5.37	206	1.5	10.62	12.25	1.87
Group 3	Control	6.25	262	1.5	10.12	10.62	1.81

5. Interpretations

The results "provide a striking confirmation of the hypothesis that play with a specially designed toy can lead to greater understanding of the rules embodied in a mathematical concept and a better memory of such rules than can be provided by observation of the same stimuli but without manipulation."

It remains the task of future research using both experimental and observational techniques to suggest how manipulative activity in play situations might fruitfully be integrated with formal teaching. Such research should include an assessment of long-term memory, especially since the present study was confined to short-term results.

Abstractor's Comments

This is a neat, straightforward, well-reported study. However, a number of trifling questions might be raised.

a. The authors seem to use the expressions "base four" and "modulo 4" as though they were interchangeable.

b. The terms "productive thinking," "problem solving," "divergent thinking," "activity methods of teaching," "play," "concept formation," "autonomous activity," and "creativity" all slip a bit loosely around in this paper, as though the research had to do with all of them.

   In the same sense, "play"--a central issue in the present study--could have been defined more precisely than merely to describe children's activity (whatever it was) with the table toy.

c. It is not clear that the 1-12 rankings given participants were comparable, since within-teacher bias is possible and since the rankings were within-classroom comparisons and several different classrooms appear to have been used. Whether the rankings were provided by one teacher or several teachers is unclear.

d. Several small sections of the text are unclear: (1) "...for our purpose, the mathematical abstractions had to be structured such that they could be represented on a piece of machinery to be used as our pre-task toy." What pre-task toy? (2) Concerning the scoring of tape-recorded responses by an independent, experienced rater: "There was substantial agreement between the two rankings: 32/48 rankings were identical and the other 16 differed by an average of 0.7." What two rankings? If the two rankings are those from the pretest interview and the posttest interview, the claim does not square with the data which, fortunately, is reported in full. (3) The "Yoke" condition is not fully described. Did "Yoke" children see the input bulbs light sequentially followed by the product bulb, or did they see the inputs and outputs lighted simultaneously? (4) What sort of investment did children in Groups 2 and 3 have in the outcome? Did they know that they would be responsible for their performance? Did any members of the groups take notes? (6) It may be risky to associate gains immediately assessed as the result of 7½ minutes of treatment with anything called "cognitive advance." There are differences between wavelets and tides.

But this is nitpicking. The study is relatively simple and clear. There are several issues which could fruitfully be explored as extensions of the present research:

a. What is the sticking power of that which was learned?
b. The interviews following the Concept Task must have been fascinating. What a rich source of information! What relationships are there between the data obtained here and the findings of Bell (1976), Erlwanger (1973), and Wason (1968)?
c. The suggestion that the effects of activity-based instruction be studied in both the short run and the long run is rich with promise. Certainly, given the curriculum work of Britons such as Edith Biggs, David Fielker, Arthur Morley, Dick Tahta, Bob Jeffery, Jan Stanfield, Leo Rogers, and dozens if not hundreds of others, researchers would not have far to go for issues or samples.

But there is one serious objection. The claim that "play" (however defined) is the source of the gains reported is invalid. One cannot give Group 1 hamburgers, tomato sauce, French fries, and Coca Cola, and Group 2 Coca Cola only, and, having demonstrated that Group 1 does better in some respect, claim that it was the tomato sauce that made the difference. In the present study, Group 1 children not only "played"; they were in control of their generating of evidence. (What if in some Group 1' children worked the table toy to someone else's tune?) Further, they were engaged, whereas Group 2 may not have been and Group 3 was not. (What if some Group 1' members were asked to view the "yoked" toy and to guess the outputs at each given input?) Other possible explanations of the phenomena reported in the present study--Group 1's gain and superiority--may well be hypothesized and many may withstand experimental scrutiny.

In brief, this is a good preliminary study demanding further refinement.

REFERENCES

Bell, A. W. A Study of Pupils' Proof-Explanations in Mathematical Situations. Educational Studies in Mathematics 7: 23-40; July 1976.

Erlwanger, S. H. Benny's Conception of Rules and Answers in IPI Mathematics. Journal of Children's Mathematical Behavior 1: 7-26; Autumn 1973.

Wason, P. C. "On the Failure to Eliminate Hypothesis ..., -- a Second Look." In Wason, P.C. and Johnson-Laird, P.N. (eds.) Thinking and Reasoning. Harmondsworth, Middlesex: Penguin Books Ltd., 1968. Pp. 165-174.

MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN JOURNALS AS INDEXED BY CURRENT INDEX TO JOURNALS IN EDUCATION March 1979*

EJ 190 564 Richardson, Lloyd I; And Others. Comparison of the Performance of First-Grade and Mentally Retarded Students on the Peabody Mathematics Readiness Test. American Journal of Mental Deficiency, v83 n1, pp83-85, July 1978.

EJ 191 055 Vance, J. H. Attitudes toward Mathematics and Mathematics Instruction of Prospective Elementary Teachers. Alberta Journal of Education Research, v24 n3, pp164-72, September 1978.

EJ 191 184 Giese, Madeline. What's Up with Those Calculators in the Classroom? Part One. Calculators/Computer Magazine, v2 n1, pp67-70, January 1978.

EJ 191 185 Roberts, Dennis M; Glynn, Shawn M. A Comparison of Numerical Problem Solving under Three Types of Calculation Conditions. Calculators/Computers, v2 n1, pp71-74, January 1978.

EJ 191 235 Davis, Edward J. A Model for Understanding Understanding in Mathematics. Arithmetic Teacher, v26 n1, pp13-17, September 1978.

EJ 191 282 Power, Colin; McBryde, Bruce. An Evaluation of Two Computer Mathematics Programmes. Australian Mathematics Teacher, v34 n1-2, pp1-11, February-March 1978.

EJ 191 292 Friesen, Charles D. A Look at General Education Mathematics Programs. Two-Year College Mathematics Journal, v9 n4, pp218-21, September 1978.

EJ 191 358 Lowerre, George F.; And Others. Using Electronic Calculators with Third and Fourth Graders: A Feasibility Study. School Science and Mathematics, v78 n6, pp461-64, October 1978.

EJ 191 360 Gussett, James C. What Does Industry Really Want in a Mathematics Major? School Science and Mathematics, v78 n6, pp472-74, October 1978.

EJ 191 583 Troutman, James G. Cognitive Predictors of Final Grades in Finite Mathematics. Educational and Psychological Measurement, v38 n2, pp401-04, Summer 1978.

---

*Due to a change in publishers, issues of CIJE have been late in arriving.

ED 191 599 Brush, Lorelei R. A Validation Study of the Mathematics Anxiety Rating Scale (MARS). Educational and Psychological Measurement, v38 n2, pp485,90, Summer 1978.

MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN RESOURCES IN EDUCATION April - May 1979

ED 162 483 Reid, D. Kim. The Effects of Cognitive Tempo and the Presence of a Memory Aid on Conjunctive Concept Attainment in Educable Mentally Retarded Boys. 31p. MF and HC available from EDRS.

ED 162 487 Cronk, George A., Jr. Title VI-B Evaluation Report 1976-77. New York State Education for the Handicapped Act. 42p. MF and HC available from EDRS.

ED 162 552 Shapiro, Leonard. Final Report: ISCI. Consulting Group on Instructional Design. 57p. MF and HC available from EDRS.

ED 162 709 Raines, Roy H. A Comparative Analysis of Learning Styles and Teaching Styles of Mathematics Students and Instructors. 290p. MF and HC available from EDRS.

ED 162 855 Hassett, Matthew J. Effects of Student Attitude Data on the Development of a PSI Program in Mathematics. 21p. MF available from EDRS. HC not available from EDRS.

ED 162 861 Lachat, Mary Ann; And Others. An Evaluation of the Oregon System in Mathematics Education. Final Report. 185. MF and HC available from EDRS.

ED 162 862 Chang, Ping-Tung. Persistence in Small Group Instruction in Developmental Mathematics Courses - on Relationships Among Academic Performance, Sex Difference, Attitude and Persistence. 4p. MF and HC available from EDRS.

ED 162 863 Osborne, Alan R., Ed. Investigations in Mathematics Education, Vol. 11, No. 2, Spring 1978. MF and HC available from EDRS.

ED 162 876 Sherrill, James M. NCTM Commission on the Education of Teachers of Mathematics Survey of Pre-Service Elementary Mathematics Teacher Education Programmes. 224p. MF available from EDRS. HC not available from EDRS.

ED 162 877 Sherrill, James M. NCTM Commission on the Education of Teachers of Mathematics Survey of Pre-Service Secondary Mathematics Teacher Education Programmes. 146. MF available from EDRS. HC not availabe from EDRS.

ED 162 878 Sherrill, James M. NCTM Commission on the Education of Teachers of Mathematics Survey of Pre-Service Secondary Mathematics Teacher Education Programmes in Canada. 142. MF available from EDRS. HC not available from EDRS.

ED 162 879 Davidson, Neil Andrew. The Small Group-Discovery Method of Mathematics Instruction as Applied in Calculus. Technical Report No. 168. 384p. MF available from EDRS. HC not available from EDRS.

ED 162 896 Downes, John P.; And Others. 76 Questions: A Synthesis of the Research on Teaching and Learning Mathematics. 112p. MF and HC available from EDRS.

ED 162 899 Doyle, Mercedes; Graesser, Arthur, II. Mathematical Problem Spaces of the Mathematically Anxious and the Mathematically Comfortable. Paper Presented at a Meeting of the Western Psychological Association, San Francisco, California, April 1978. 23p. MF available from EDRS. HC not available from EDRS.

ED 162 901 Burstein, Leigh. Implications from the Beginning Teacher Evaluation Study for the IEA Second Mathematics Study. 25p. MF and HC available from EDRS.

ED 163 026 Wong, Shirley M. A Measurement of the Basic Consumer-Related Mathematical Competency of Urban Grade 12 Students, Report No. 78:20. 24p. MF and HC available from EDRS.

ED 163 067 Gibson, Janice T. A Current View of Soviet Research in Cognitive Psychology. 8p. MF and HC available from EDRS.

ED 160 088 Evans, Franklin R. The GRE-Q Coaching/Instruction Study. GRE Board Professional Report GREB No. 71-5aP. 34. MF available from EDRS. HC not available from EDRS.

ED 163 091 Marco, Gary L. The Performance of First-Year Graduate Students on the Graduate Record Examinations. Graduate Record Examinations Special Report Number 68-2. MF available from EDRS. HC not available from EDRS.

ED 163 749 Krakow, Joanne; Curcio, Frank. Concrete Operational Development and Arithmetic Progress in Learning Disabled Boys: A Cross-Lagged Panel Analysis. 11p. MF and HC available from EDRS.

ED 163 973 Blaine, Lee; And Others. Technical Problems in Implementing University-Level Computer-Assisted Instruction in Mathematics And Science. First Annual Report. Technical Report No. 293. 29p. MF and HC available from EDRS.

ED 163 979 Bond, Nicholas A.; And Others. Studies of Verbal Problem: II. Prediction of Performance from Sentence-Processing Scores. Technical Report No. 87. 57p. MF and HC available from EDRS.

ED 163 992 Dugdale, Sharon; Kibbey David. Elementary Mathematics with PLATO. 33p. MF and HC available from EDRS.

ED 164 247 Becher, Rhoda McShane. Teacher Behaviors Related to the Mathematical Achievement of Young Children. 10p. MF and HC available from EDRS.

ED 164 259 Hewitt, Gloria C. The Status of Women in Mathematics. 15p. Document not available from EDRS.

ED 164 262 Heller, Kirby A.; And Others. Expectancies, Utility Values, and Attributions for Performance in Mathematics. 13p. MF and HC available from EDRS.

ED 164 271 Brendzel, Sharon. Sex Differences in Proportional Reasoning. 10p. MF available from EDRS. HC not available from EDRS.

ED 164 274 Proceedings of the Research-on-Teaching Mathematics Conference, May 1-4, 1977. Conference Series No. 3. 246p. MF and HC available from EDRS.

ED 164 275 Schmidt, William H.; And Others. Training Manual for the Classification of the Content of Fourth Grade Mathematics. Technical Manual. Research Series No. 4. 47p. MF and HC available from EDRS.

ED 164 307 NLSMA Technical Reports, Nos. 1-44. 473p. Document not available from EDRS.

ED 164 355 Stalker, George H.; And Others. USMES School Study: Effects of School Strategies on USMES Utilization. 121p. MF and HC available from EDRS.

ED 164 356 Arbetter, Carolyn Clinton; And Others. USMES Student Study: Effects of USMES on Student Problem-Solving Abilities and Attitudes. 99p. MF and HC available from EDRS.

ED 164 357 Arbetter, Carolyn Clinton; And Others. USMES Team Study: Using Resource Teams to Disseminate and Implement USMES. 247p. MF and HC available from EDRS.

ED 164 495 Phillips, Beverly. Adolescent Thought: Curriculum Implications. 12p. MF and HC available from EDRS.

ED 164 582 Antill, Ellen; Speidel, Gisela E. Standardized Test Results: KEEP and Control Students. 1975-1976, Technical Report #69. 20p. MF and HC available from EDRS.

ED 164 591 Mathematics: Texas Assessment of Essential Objectives. 61p. MF and HC available from EDRS.

ED 164 622 Holmes, Mark. Report on the Intermediate Evaluation Project. 23p. MF and HC available from EDRS.