DOCUMENT RESUME

ED 076 387 SE 015 852

AUTHOR	Higgins, Jon L., Ed.
TITLE	Research Reporting Sections, National Council of Teachers of Mathematics 51st Annual Meeting, Houston, Texas, April 25-28, 1973.
INSTITUTION	ERIC Information Analysis Center for Science, Mathematics, and Environmental Education, Columbus, Ohio.
PUB DATE	Apr 73
NOTE	109p.
AVAILABLE FROM	Center for Science and Mathematics Education, The Ohio State University, 244 Arps Hall, Columbus, Ohio 43210 ($1.75)
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	Attitudes; Evaluation; Instruction; Learning; Logic; *Mathematics Education; *Research; Teacher Characteristics; Teaching Methods; Test Construction
IDENTIFIERS	*Research Reports

ABSTRACT

This document is a compilation of abstracts of 20 research papers presented at the 51st Annual Meeting of the National Council of Teachers of Mathematics. Six reports concern methods of instruction, eight investigate patterns of learning, three deal with evaluation of attitudes, and three reports cover tests and test construction. (DT)

ED 076387

SMEAC/SCIENCE, MATHEMATICS, AND ENVIRONMENTAL EDUCATION INFORMATION ANALYSIS CENTER

. . . an information center to organize and disseminate information and materials on science, mathematics, and environmental education to teachers, administrators, supervisors, researchers, and the public. A joint project of The Ohio State University and the Educational Resources Information Center of USOE.

ED 076387

MATHEMATICS EDUCATION REPORTS

RESEARCH REPORTING SECTIONS

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS

51st ANNUAL MEETING

Houston, Texas
April 25-28, 1973

ERIC Information Analysis Center for
Science, Mathematics, and Environmental Education
400 Lincoln Tower
The Ohio State University
Columbus, Ohio 43210

PREFACE

The ERIC Information Analysis Center for Science, Mathematics, and Environmental Education has compiled abstracts of the research papers to be presented at the 51st annual meeting of the National Council of Teachers of Mathematics. Selection of papers was made by Professor Ray Carry of The University of Texas and a committee of the AERA Special Interest Group for Research in Mathematics Education. Minor editing has been done by the ERIC staff to provide a general format for the papers. Many of the papers that are abstracted here will be published in journals or be made available through the ERIC system. These will be announced in Research in Education and in the journal Investigations in Mathematics Education.

April, 1973

Jon L. Higgins
Associate Director for
Mathematics Education

Sponsored by the Educational Resources Information Center of the United States Office of Education and the Center for Science and Mathematics Education, The Ohio State University.

This publication was prepared pursuant to a contract with the Office of Education, United States Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official Office of Education position or policy.

CONTENTS

NCTM RESEARCH REPORTING SECTIONS
1973 Annual Meeting

April 26, 1973
Thursday
12:00-1:00 p.m. RESEARCH REPORTING SECTION I

Presider: L. RAY CARRY, The University of Texas at Austin
Austin, Texas

Reports: A STUDY OF THE INTERACTION BETWEEN SEX DIFFERENCES, STRUCTURE-OF-INTELLECT FACTORS AND TWO METHODS OF TEACHING A MATHEMATICAL RELATION

ROBERT R. HANCOCK, Eastern Illinois University
Charleston, Illinois

RELATING LOGICAL REASONING ABILITY TO STUDENT MATHEMATICAL ABILITY AND TEACHER VERBAL BEHAVIOR

JOHN W. GREGORY, University of Florida
Gainesville, Florida

ALAN R. OSBORNE, The Ohio State University
Columbus, Ohio

NEGATIVE INSTANCES AND THE ACQUISITION OF THE MATHEMATICAL CONCEPTS OF COMMUTATIVITY AND ASSOCIATIVITY

RICHARD J. SHUMWAY, The Ohio State University
Columbus, Ohio

FRANK K. LESTER, Indiana University
Bloomington, Indiana

VARIABLES AFFECTING PERFORMANCE ON ARITHMETIC WORD PROBLEMS IN A CAI SETTING

BARBARA W. SEARLE, Stanford University
Stanford, California

PAUL LORTON, JR., Stanford University
Stanford, California

PATRICK SUPPES, Stanford University
Stanford, California

A STUDY OF THE INTERACTION BETWEEN SEX DIFFERENCE, STRUCTURE-OF-INTELLECT FACTORS AND TWO METHODS OF TEACHING A MATHEMATICAL RELATION

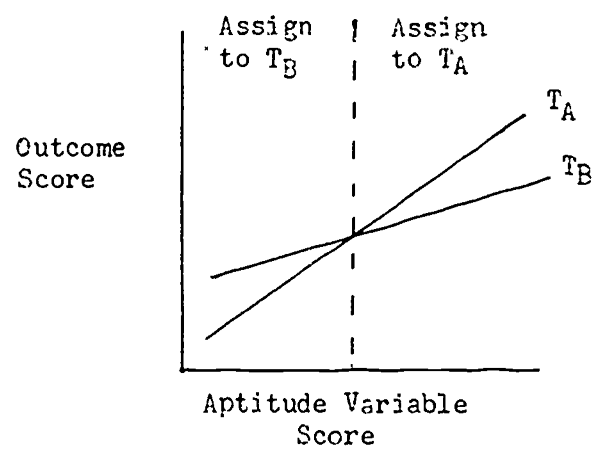
Robert R. Hancock
Eastern Illinois University

1. Purpose

This study was undertaken with a view to (1) attempting to identify personological variables which might be significantly related to differing cognitive preferences and (2) seeking evidence, in the form of disordinal interactions, that instructional materials which will interact with differing profiles of individual aptitudes can, in fact, be prepared. Specifically, this study investigated the interaction between personological variables of sex difference, certain mental factors chosen from Guilford's Structure-of-Intellect model with two methods of presenting concepts and principles associated with a contrived mathematical relation.

2. Rationale

Current thinking in education emphasizes the desirability of individualizing instruction in order to maximize 'payoff' in terms of achievement, retention, and transfer of training. However, educational research tends to deal largely with questions related to the significance of the difference between group mean scores; aptitude-treatment-interaction (ATI) research, on the other hand, seeks to provide a basis for employing differential treatments in order to exploit the cognitive preferences various individuals display for differing content and/or mode of presentation. The regression lines for T_A and T_B in the figure below illustrate the concept of maximizing outcome by making differential assignments on the basis of some aptitude variable score.



The work of Cronbach and Snow (1969) is a significant contribution to the area of ATI research.

3. Research Design and Procedure

Subjects, 176 undergraduates enrolled in a terminal, cultural level course in mathematics, were given a battery of nine tests developed by Guilford, et al and designed to measure cognitive ability in dealing with figural, semantic, and symbolic content. Reliabilities ranged from 0.59 to 0.78 and scores on these tests were treated as the independent variables on the study.

Ss were randomly assigned, by sex, to one of two treatment groups, these groups were shown to be homogeneous with respect to means and variances of independent variable scores and to satisfy the assumptions necessary for the subsequent regression analysis. For two fifty (50) minute class periods Ss studied one of two programmed units prepared for this investigation. These programs dealt with a contrived mathematical relation that had all of the characteristics of a linear order relation, however, this fact was disguised by the definitions employed and by the notation adopted. One program presented the material employing a verbal mode, the other a figural mode. Both programs had been examined by a panel of mathematics educators and judged to be of parallel content and to conform to the specified instructional mode. Results of a pilot study had shown both programs to be effective instructional devices.

Dependent variables were: (PT) time required to complete the program, (PE) the number of self-reported program errors, (CT-T) total criterion test score, and (CT-I, CT-II, CT-III) three criterion subtest scores at differing levels of cognitive activity, these cognitive levels were adapted from Bloom, et al. Reliabilities for the various criterion measure ranged from 0.63 to 0.86 and were obtained on the day following completion of the study of the instructional programs. No attempt was made to assess retention or the transfer of learning to new tasks.

Regression analysis with $\alpha = 0.05$ was carried out in the following manner: For each dependent variable, Y_i, and for each independent variable, X_j,

1. Simple, linear regression equations were determined for both of the treatment groups and an F-test made of the significance of regression due to linear regression. If both regression lines intersected within the range of observed scores on the independent variable, X_j, then

2. The difference between the regression coefficients was

tested for significance. If this difference was found to be significant, then it was concluded that, with respect to the dependent variable, Y_i, a disordinal interaction existed between the independent variable, X_j, and the two treatment modes.

A similar analysis was conducted in order to compare performance by the male Ss who studied the verbal material with those who studied the figural program. The same procedure was also carried out for the female Ss.

In addition to the regression analysis described above, t-tests were carried out to assess the impact of treatment main effects for the various experimental groups.

4. Findings

Results of this study failed to disclose any disordinal interactions, however, several significant results were obtained with regard to main effects. Significant differences which favored the verbal mode of presentation were found on mean scores for CT-I, CT-II, and CT-T; performance, although not significant at the 0.05 level, also favored the verbal treatment group on CT-III. Furthermore, it was found that female Ss performed significantly better on test CT-III, the test of highest cognitive level. Data relative to PT and PE were inconclusive.

5. Interpretations

It is felt that the clearcut superiority demonstrated by the verbal treatment group is related to the fact that all Ss were university students and had shown a priori if not preference for, at least the ability to learn successfully from, content material which is highly verbal. The study reported in this abstract is presently being replicated at the junior high school level in an effort to determine if younger students display this same pattern of learning preferences.

Female superiority on CT-III was as unexpected as it is difficult to explain. However, it does seem to indicate that further study of sex-difference as a variable of interest is desirable.

Investigator's Notes

During the course of this study it became apparent that requiring significant linear regression as a necessary condition for complete regression analysis would have a marked effect on the findings. One may easily draw unjustified conclusions relative to the occurrence of disordinal interactions when the linear model is assumed; thus, as a matter of methodology in regression analysis, it is suggested that, unless significant regression due to linear regression has been established, the investigator exercise due caution in claiming the existence of disordinal interactions.

One assumption that is implicit in ATI research is that optimum learning for students who score relatively higher on a test, for example, of figural aptitude, will occur when a figural mode of presentation is employed. But this does not provide an operational definition that allows the researcher to classify content material as 'figural' in the sense that it will be cognitively appealing to those students scoring high on a test of figural aptitude. That is, 'figural' in the sense of an aptitude measure might be distinct from 'figural' in the sense of instructional content. This focuses attention upon what appears to this investigator to be an area of major concern for anyone pursuing ATI research. In commenting on the lack of empirical evidence to solidly support claims of disordinal interaction, Glaser (1972) says, "This is an astounding conclusion; it implies that our generally used aptitude constructs are not productive dimensions for measuring those individual differences that interact with different ways of learning."

The study of aptitude-treatment-interaction appears to be a particularly challenging area of investigation when one contemplates the pedagogical implications of a decision making process based upon concepts and principles related at ATI. It is felt that research of the type described in this abstract is in the mainstream of current educational efforts to individualize instruction and should, therefore, be pursued with imagination and determination.

REFERENCES

Cronbach, L. J. and Snow, R. E. Individual Differences in Learning Ability as a Function of Instructional Variables (Final Report, OEG 4-6-061269-1216) Washington, D. C.: United States Government Printing Office, 1969.

Glaser, Robert. "Individuals and Learning: The New Aptitudes," Educational Researcher, Vol. 1, No. 6, June 1972, (5-13).

RELATING LOGICAL REASONING ABILITY TO STUDENT MATHEMATICAL ABILITY AND TEACHER VERBAL BEHAVIOR

John W. Gregory
University of Florida

Alan R. Osborne
Ohio State University

1. Purpose

This study investigated teacher verbal behavior and student mathematical ability as possible correlates of student conditional reasoning ability. In addition to testing the alternate hypotheses, this study sought to integrate analyses of teacher verbal behavior and psychological and linguistic analyses of the growth and development of children's logical abilities.

2. Rationale

Experimental studies have found that control groups not receiving instruction in the application of principles of logic tend to perform as well as experimental groups that have received formal instruction on tests of this ability (e.g., Ennis and Paulus, 1965; Roy, 1970; and Williams, 1971). Alternate hypotheses suggested by these investigations include possible influences of teacher effectiveness in terms of both content presentation and interaction with students and other subject matter coursework which requires abstract learning.

3. Research Design and Procedure

Twenty teachers in a large metropolitan school system were randomly selected from a total population of eighty-four teachers of seventh grade mathematics. During a three week period, five lessons of each teacher presented to one selected class were audiotaped. Subsequent analysis of the reacher' verbal behavior by three trained analysts identified teachers having significant differences ($p < .05$) in frequency of conditional moves. The teachers were then ranked on the basis of this analysis. The students in the selected class of each of the five teachers with highest ranks and the five teachers with lowest ranks served as subjects for determining the relationship of student conditional reasoning ability to teacher frequency of conditional moves, to student mathematical ability, and to the combination of student mathematical ability and teacher frequency of conditional moves.

Mathematical ability was determined as measured by the California Comprehensive Arithmetic Test. This instrument was administered to all students within the school system in October. Student conditional reasoning ability was determined by performance on the Cornell Conditional Reasoning Test having been administered in September (pretest) and January (posttest).

Multivariate analysis of covariance (pretest serving as covariate) was used to determine significant differences in posttest performance and adjust posttest scores for computing correlations within the two-by-three factorial design (high and low rank of teacher; below average, average, and above average student mathematical ability).

4. Findings

Mathematical ability was found to be a significant correlate of posttest performance on the three fallacy principles of denying the antecedent (r = .450), asserting the consequent (r = .646), and asserting the converse (r = .446) as well as the total test (r = .473). Students of the five teachers with high rank outperformed students of the five teachers with low rank on the total test (r = .429), on suggestive content items (r = .536), on items involving negation (r = .519), and the principle "p only if q, not q : not p" (r = .422). No interaction of mathematical ability and teacher frequency of conditional moves was found.

5. Interpretations

The results of this study indicate two major factors for mathematics educators. First, the positive relationship between mathematical ability and logical reasoning ability found in this study does not indicate that teaching mathematics implies that one is necessarily teaching logic. The direction of the arrow is still subject to question. However, the results seem to indicate that what is important is the manner in which the mathematics is presented, linguistically speaking.

The positive correlation found between the frequency of conditional moves by teachers and their students logical reasoning ability does not indicate the direction of the arrow, either. But it is difficult to conceive that student reasoning ability would alter teacher selection and utilization of particular linguistic patterns, especially with regard to the teacher's ordinary language.

REFERENCES

Ennis, R. H. and Paulus, D. H. Critical Thinking Readiness in Grades 1-12; Phase I; Deductive Logic in Adolescence. Ithaca: Cornell University, 1965.

Roy, Gregory Raymond. The effect of the study of mathematical logic on student performance in proving theorems by mathematical induction. (Unpublished doctoral thesis, State University of New York at Buffalo, 1970).

Williams, Billy Richard. Critical thinking ability as affected by a unit on symbolic logic. (Unpublished doctoral thesis, Arizona State University, 1971).

NEGATIVE INSTANCES AND THE ACQUISITION OF THE MATHEMATICAL CONCEPTS OF COMMUTATIVITY AND ASSOCIATIVITY

Richard J. Shumway
The Ohio State University

Frank K. Lester
Indiana University

1. Purpose

Two questions were examined: (1) What are the different effects of an instructional sequence of positive and negative instances and a sequence of all positive instances on the acquisition of commutativity and/or associativity; and (2) Assuming there are effects for negative instances, do the effects of negative instances for one concept transfer to another concept?

2. Rationale

Negative instances have been considered by mathematicians to be essential to the understanding of advanced mathematical concepts (Gelbaum and Olmsted, 1964; Steen and Seebach, 1970). Markle and Tiemann (1970) state explicitly that all instructional sequences designed for concept learning should include negative instances. Yet, a review of the research on concept learning generally confirms a debilitating effect for negative instances (Clark, 1971). This study was an attempt to reconcile this controversy regarding the role of negative instances in the acquisition of mathematical concepts.

3. Research Design and Procedure

Eighty-four undergraduate elementary education majors were assigned in equal numbers to four treatments. Concept A was defined to be commutativity of a binary operation and Concept B was defined to be associativity of a binary operation. The symbol A+ denoted a treatment of twenty positive instances of Concept A and the symbol A± denoted a treatment of ten positive instances and ten negative instances of Concept A. The symbols B+ and B± were defined similarly. The result of crossing levels of A (A+, A±) and levels of B (B+, B±) yielded a 2 x 2 factorial design with four treatments labeled A+B+, A±B+, A+B±, A±B±. Each treatment consisted of a sequence of forty instances. The feedback during treatment was a simple correct or incorrect.

Each subject took two ten item pretests designed to check the subject's ability to calculate with binary operations and parentheses necessary during treatments and posttests. The pretests were followed by one of the forty instance treatments. Each subject then took two twenty instance posttests designed to test the subjects' ability to classify new instances of Concepts A and B. All pretests, treatments, and posttests were administered using IBM 2741 computer terminals. Throughout the experiment, the following data were collected: number of correct answers, stimulus intervals, and postfeedback intervals (during treatments only).

The data were analyzed using a MANOVA program for a multivariate two-way analysis of covariance (Clyde, 1969). Because of the symmetry of the design, the results for Concept B were viewed as a potential replication for the results for Concept A. Hence, the analysis for Concept B was done separately from the analysis for Concept A.

Achievement variables were analyzed separately from time variables.

4. Findings

Multivariate and univariate analyses of the pretests indicated there was a significant interaction effect on stimulus intervals for both pretests. No other significant multivariate or univariate differences for correct answers or stimulus intervals on either pretest were found. There was a significant effect for levels of A on the correct answers for the posttest for Concept A favoring the treatment containing negative instances ($p < .001$). There was a significant effect for levels of B on the correct answers for the posttest for Concept B favoring the treatment containing negative instances ($p < .001$). Analysis of time variables indicated a significant difference in treatment post feedback intervals for levels of A on Concept A and a significant difference in treatment stimulus intervals for levels of B on Concept B ($p < .001$, $p < .01$).

5. Interpretations

The results supported the contention that negative instances enhance concept acquition but they also appear to require more time during treatments. No evidence for a transfer effect for negative instances from one concept to another was found.

VARIABLES AFFECTING PERFORMANCE ON ARITHMETIC WORD PROBLEMS IN A CAI SETTING

Barbara W. Searle, Paul Lorton, Jr., and Patrick Suppes
Stanford University

1. Purpose

(a) To identify structural variables that affect performance of students on arithmetic word problems presented at a computer terminal.

(b) To assess the usefulness of the identified variables as predictors of problem difficulty and other performance parameters.

(c) To use the identified variables in the restructuring of the CAI curriculum.

2. Rationale

A central theme of mathematics instruction is developing in students problem-solving skills that will generalize beyond the tasks and skills of the elementary-level mathematics curriculum. Instruction in the solution of arithmetic word problems is used as a method for teaching problem solving skills. Word problems are notoriously difficult for students, and despite intense interest and work, much remains to be learned about the sources of problem difficulty. Use of the computer makes possible the design of an instructional program that focuses on students' use of problem-solving skills, while minimizing the need for well-developed computational skills. At the same time the computer allows the collection of a large and detailed data base.

3. Research Design and Procedure

(a) The Problem-Solving (PS) Course

The emphasis of the problem-solving course is on methods of solution; the student constructs a well-formed algebraic expression and the computer carries out the actual computation. The student learns a set of simple commands that he uses to tell the computer which computations to carry out. The student is free to experiment with the computer calculator made available to him. His answer is evaluated by the computer only when he specifically instructs the computer to do so. Although the text of the problems is stored in the computer, the numbers used in each problem are generated by the program for each presentation. Thus, although more than one student sees the same problem statement, each has a different set of numbers to work with.

The student working at constructing a solution for a problem in the PS course is free to solve a problem using any combination of steps that produce the correct answer. The program confers this freedom by storing a solution string for the problem. The stored solution string allows the computer to calculate the correct answer from the values of the variables generated for the problem presentation.

The portion of the curriculum used for the current study contains (1) a set of instructional problems comprised of 14 nonnumerical problems (how to find characters on the keyboard, etc.) and 25 numerical problems illustrating different problem types, and (2) 100 numerical problems arranged in order of increasing difficulty, as described below.

The order of problems was determined in a systematic fashion, using data obtained from a pilot study. The steps followed in ordering problems were:

Step 1. A pilot study was conducted using a set of 65 problems. These were presented to students at a terminal, using response formats somewhat different from those of the PS course, but with the same constraint; that is, the computer carried out the calculations. The subjects were 16 sixth-grade students from a culturally disadvantaged area and 27 sixth-grade students from a middle-class area.

Step 2. The results of this test were examined and a set of variables thought to relate to problem difficulty were defined. These variables included the minimum number of different arithmetic operations required to reach a solution (OPERS), the minimum number of binary operations required to reach a solution (STEPS), the number of words in the problem statement (LENGT), whether the problem solution required conversion of units without the unit of conversion present in the problem statement (CONVR), whether the problem contained a verbal clue for an operation (VCLUE), whether the solution required an addition (ADD), a subtraction (SUB), a multiplication (MUL) or a division (DIV), and several other variables. A stepwise multiple regression analysis was used to determine the contribution of each of these variables to the multiple R. Five variables were found to account for 60 percent of the variability, and the contribution of each of the remaining variables was less than 1 percent. The five variables were, in the order in which they entered the regression, OPERS, CONVR, LENGT, DIV and VCLUE.

Step 3. A set of 700 problems was written and edited, and coded using the five variables identified in the pilot study.

Step 4. A multiple regression analysis was carried out on the pilot study data, using only these five variables. The regression coefficients obtained were used to predict the probability correct for the 700 problems written for the main curriculum, and the problems were ordered according to predicted probability correct. The predicted probability correct ranged from .95 to .07.

(b) The Students

The students were fourth, fifth, and sixth graders enrolled in the computer-based elementary mathematics drill-and-practice course given by the Institute for Mathematics Studies in the Social Sciences (IMSSS) at Stanford University. Approximately two-thirds of the students were from a primarily black California elementary school and the remainder from schools for the deaf in several parts of the country.

A student became eligible for the PS course when his average grade placement on the math drill-and-practice program reached 4.0. Thereafter, if his teacher chose to enroll him in PS, he received a problem-solving lesson every fifth day. Thus each student started the course at a different time throughout the year and worked at his own pace sequentially through the curriculum. The data reported here are averages for from 61 to 309 responses to individual problems.

Data collected included, for each problem, the number of correct responses on the first, second, and third try, the average number of steps, the average latency for each try, and the number of hints requested. Although results for hearing and deaf students have been examined separately only pooled data are presented in this report. The differences were small; the average probability correct was .692 for deaf students and .706 for hearing students.

4. Findings

The proportion correct (P) for each problem was obtained. Although predicted probability correct for the 125 problems used in the analysis ranged from .79 to .95, the observed P ranged from .03 to .94. The P for 64 percent of the problems fell in the range .60 to .94. For all but 5 problems, the observed P was lower than the predicted P. The mean difference between observed and predicted P was -.22. Thus, the pilot study yielded overestimates of performance.

The stepwise multiple linear regression analysis was repeated. The set of variables was augmented and some definitions of variables were changed. This analysis yielded a multiple R of .85 with six variables contributing significantly. These variables were OPERS, VCLUE, ORDER, ADD, SUB, and a newly defined variable, ALGER. ALGER characterizes a problem as a purely algebraic statement (find the sum of # and #), rather than as a word problem. Only two of the variables, OPERS and VCLUE, were the same as those contributing significantly to the multiple R of the pilot study.

5. Interpretation

The findings of greater variability in performance, compared with predictions and shifting of variables contributing significantly to the regression analysis, came as no surprise to the authors. First, it is clear that characteristics of the problem set, for example the frequency of occurrence of exemplars for the range of values for each variable and the way variables values are combined in problem types, affect the weighting for each variable in the regression analysis. Thus differences are expected, because of differences in the problem sets used for the pilot study and present study. Second, the population for this study was quite different from the population used in the pilot study.

Investigator's Notes

Our results are currently being used in a continuing effort to characterize problem difficulty. An attempt was made to broaden the problem set by examining the 256 theoretically possible problem types defined by the six variables OPERS, VCLUE, ORDER, ADD, SUB, and ALGER and writing new problems to exemplify combinations of the variables previously absent from the curriculum. Further, a more elaborate experimental design is being used. A computer program, constructed to periodically calculate individual regression coefficients for performance on a set of 25 problems, uses these coefficients to reorder the problems in the curriculum as described above. Each student therefore works on a problem set constructed for him. He is placed in this curriculum at an appropriately selected P level. Thus we are investigating whether and how the weightings for variables differ for different students, and if it is feasible to use this knowledge in specifically tailoring a curriculum to the individual student.

April 26, 1973
Thursday
1:30-2:30 p.m.

RESEARCH REPORTING SECTION II

Presider: THOMAS J. COONEY, University of Georgia
Athens Georgia

Reports: A TAXONOMIC APPROACH TO EVALUATING ATTITUDES OF PROSPECTIVE ELEMENTARY TEACHERS IN A MATHEMATICS EDUCATION COURSE

RALPH D. CONNELLY, Memorial University of Newfoundland
St. John's, Newfoundland, Canada

ATTITUDES AND PERCEPTIONS OF ELEMENTARY MATHEMATICS POSSESSED BY THIRD AND SIXTH GRADE TEACHERS AS RELATED TO STUDENT ATTITUDE AND ACHIEVEMENT IN MATHEMATICS

JOHN ARTHUR VAN DE WALLE, CEMREL, Inc. - CSMP
Carbondale, Illinois

PERCEPTIONS AND ATTITUDES OF STUDENT TEACHERS IN MATHEMATICS

GERALD KULM, Purdue University
West Lafayette, Indiana

THE EFFECTS OF A MATHEMATICS LABORATORY APPROACH ON STUDENTS' SELF CONCEPTS, ATTITUDES AND ACHIEVEMENT

DAVID L. PAGNI, California State University Fullerton
Fullerton, California

FRED SHARMAN, California State University Fullerton
Fullerton, California

JOHN RANDOLPH, California State University Fullerton
Fullerton, California

A TAXONOMIC APPROACH TO EVALUATING ATTITUDES OF PROSPECTIVE ELEMENTARY TEACHERS IN A MATHEMATICS EDUCATION COURSE

Ralph D. Connelly
Memorial University of Newfoundland

1. Purpose

(a) To determine whether or not Krathwohl, Bloom, and Masia's Affective Taxonomy rationale can be used effectively in developing an instrument for evaluating affective behavioral objectives of a mathematics education course and to determine whether or not the results from the taxonomized instrument will reflect the hierarchy of categories as set forth in the Affective Taxonomy.

(b) To ascertain whether or not significant affective change takes place during the Education 325 (Mathematics education) course, and what characteristics of individuals seem to influence change.

2. Rationale

It is generally held that teacher attitude and effectiveness in mathematics are important determinants of student attitude and performance in the subject. Research by Carner (1963), Banks (1964), and Phillips (1970) support this assumption. If attitude improvement is a desirable outcome of education, then mathematics educators ought to reexamine the objectives of mathematics education courses for prospective elementary teachers to insure consideration of the contribution these courses make to attitude change as well as to cognitive change. Krathwohl (1964) and Gray (1972) have made reference to a trend toward a lack of emphasis on affective objectives that apparently has continued.

It is assumed that there is value in expressing affective educational objectives in overt student behavioral terms, and that affect in any subject area can be measured with a taxonomy-type instrument.

3. Research Design and Procedure

Affective behavioral objectives for Education 325, Mathematics for Elementary Teachers, were formulated using the Affective Taxonomy rationale, and a Taxonomy-type instrument was developed for evaluating these objectives, using a five-point Likert scale.

The final Taxonomy-type instrument was administered to a sample of 146 students in Education 325 at the beginning of Summer Session, 1972, at the main campus of Kent State University. The instrument was administered again at the close of the session. The Dutton Attitude Scale (DAS) was administered along with the Taxonomy-type instrument both times.

A series of factor analyses were performed on the data. Alpha reliability coefficients were obtained. Scale analyses were executed for each of the general objectives for Education 325. A t test was used to compare pretest and posttest means of the DAS. An analysis of variance was performed on the pretest and posttest of the Taxonomy-type instrument across taxonomic levels. Correlations between various student descriptors and the Taxonomy-type instrument were obtained.

4. Findings

Alpha reliability coefficients of .91 and .94 were obtained for the pretest and posttest, respectively, of the Taxonomy-type instrument. Using the .05 level of significance, a significant positive correlation (r = .67) was obtained between the Taxonomy-type instrument and the DAS. Factor analysis of the items classified at each taxonomic level seemed to reflect the general affective objectives for Education 325 to a high degree.

Using Guttman scale analysis, coefficients of reproducibility for scales of items classified under the general objectives for Education 325 ranged from .88 to .91 on the pretest and .89 to .94 on the posttest.

A dependent measures _t_ test showed a significant increase on the DAS from pretest to posttest at the .05 level of significance.

An analysis of variance of the pretest and posttest of the Taxonomy-type instrument revealed a significant increase from the pretest to the posttest, and significant differences between means of all the taxonomic levels on both the pretest and posttest. Significant increases took place at the Valuing and Organization levels, and significant decreases took place at the Receiving and Responding levels.

Using the .05 level of significance, insignificant correlations were obtained between overall grade-point average (GPA) and scores on the Taxonomy-type instrument, achievement in Education 325 as measured by an instrument based on the _Cognitive Taxonomy_ and scores on the Taxonomy-type instrument, and between achievement in Education 325 and affective improvement in Education 325. An independent _t_ test revealed no significant differences in means of students with eleven hours of mathematics previous to Education 325 and students with five hours of mathematics.

5. Interpretations

(a) The Affective Taxonomy rationale can be used effectively in developing an attitude scale for evaluating affective objectives of a mathematics education course.

(b) Results on the Taxonomy-type instrument do reflect the hierarchy of categories as set forth in the Affective Taxonomy.

(c) Significant improvement in affect does take place as a result of Education 325.

(d) Attitudes toward mathematics as measured by the Taxonomy-type instrument are not significantly related to overall GPA nor to achievement in Education 325.

(e) Attitude improvement in Education 325 is not significantly related to achievement in the course.

(f) Mathematics background has no significant effect on attitudes toward mathematics measured by the Taxonomy-type instrument.

Investigator's Notes

Considerable difficulty was encountered by the judges aiding in classification of test items when these items were stated negatively. To fully utilize the Affective Taxonomy, perhaps only positively stated items should be used.

ATTITUDES AND PERCEPTIONS OF ELEMENTARY MATHEMATICS POSSESSED BY THIRD AND SIXTH GRADE TEACHERS AS RELATED TO STUDENT ATTITUDE AND ACHIEVEMENT IN MATHEMATICS

John A. Van de Walle
CEMREL-CSMP

1. Purpose

(a) To define and seek out the influence which different perceptions of mathematics (i.e., intuitive judgements about mathematics along a formal-informal continuum) held by teachers may have on student computational ability, comprehension of mathematical concepts and attitudes, at grades three and six.

(b) To determine the direction of causal relationships between the student and teacher variables.

2. Rationale

The most significant factor contributing to student attitude toward mathematics is the teacher and her attitude toward mathematics (Aiken, 1970; Kagen, 1965). Student attitudes also change with grade level (Dutton, 1968; Rays and Delon, 1968). On the other hand, the relationship between teacher attitude and student achievement in mathematics has not been consistently demonstrated. A possible explanation is a lack of attention to the teacher's perception of the mathematics taught. Young children can approach mathematics tentatively, experimentally, and creatively. This intuitive approach by students is hampered by formal, fixed forms of both texts and teachers (Hohn, 1961; Davis, 1965; Devault and Kriewell, 1969.)

It was hypothesized that the effect of a teacher with an informal view of mathematics is more likely to be reflected in student behavior at the higher cognitive levels. Students who not only learn to probe, but develop their own ideas, should perform better on tasks which require the ability to relate two or more concepts, or to integrate facts. Furthermore, it is possible that the influence will be more noticeable in the lower grades since, as the students are exposed to more teachers, the probability of exposure to one with a more rigid viewpoint increases.

By blocking on the teacher variables of attitude and perception, it was assumed that these variables are not highly correlated. It was also assumed that these factors influence the behavior of the teacher in the classroom.

3. Research Design and Procedure

A sixteen item Scale of Perceptions of Elementary School Mathematics (PESM) was assembled from three other instruments (Collier, NLSMA, Rettig) measuring a similar variable. Data from a small pilot study was used to aid in the selection of items. The PESM and Aiken's Revised Mathematics Attitude Scale (RMAS) were administered in September to sixty-six third grade and fifty-six sixth grade teachers from three different Ohio school systems. At the same time each teacher's class was given a computation test, a test of mathematical concepts and an attitude instrument. These student instruments were assembled by selecting items respectively from the Seeing Through Arithmetic Tests, The Wisconsin Contemporary Test of Elementary Mathematics and the NLSMA "non-test data". Mean scores were utilized in the analysis. An identical collection of data was made in March. Sixty-two third grade and fifty-one sixth grade classes (approximately 3100 students) returned completed data. Post-test reliability estimates for the student data ranged from .65 for the third grade attitude scale to .84 for the third grade achievement tests. Since class means were used as the unit of analysis the reliability of the student measures was deemed adequate. The reliability coefficients for the RMAS and PESM were .95 and .79 respectively.

The two teacher instruments were used as blocking variables, producing a two-by-two design at each grade. Multivariate analysis of covariance and correlational methods were employed.

Due to the nature of the study as a preliminary investigation the .10 confidence level was used to decrease the probability of Type II error.

4. Findings

A major limitation of the study was due to the high correlation between the pre-test teacher variables of perception and attitude (gr. 3: $r=.27$, gr. 6: $r=.51$) producing disproportionate cell sizes in the two-by-two design. Keeping this and other acknowledged limitations in mind, the following conclusions can be drawn:

(a) The multivariate analysis revealed a significant interaction of teacher attitude and perception at the third grade. The composite score which produced this interaction was essentially a difference of the computation score minus the comprehension score. The attitude of the third grade students did not seem to be affected. Teachers with informal perceptions coupled with positive attitudes were associated with student comprehension of mathematical concepts. On the other hand, a negative teacher attitude with an informal perception was associated with student computational ability. No significant interactions were found at the sixth grade level.

(b) Third grade students taught by teachers with an informal perception of mathematics performed significantly higher on all three student measures than students taught by teachers with a formal perception of mathematics. There were no significant differences due to teacher perceptions at the sixth grade level.

(c) There were no significant differences due to teacher attitude at the third or sixth grade level.

(d) No significant correlations between student and teacher variables were observed. The cross-lagged panel technique produced no significant indication of cause and effect relationships.

5. Interpretations

No clear-cut conclusions can be drawn from the results. At the sixth grade level no significant differences appeared. Perhaps this is due to the longer history of teacher and text influences diminishing the effects of a single teacher over six months. At the third grade level, the interaction effect is difficult to interpret. Coupled with the fact that teachers with informal perceptions of mathematics taught students who scored significantly higher on all three variables (attitude, computation, comprehension) it is safe to conclude that the dimension of teacher perceptions is worth further investigation. Observation and interview techniques should be employed to refine the PESM or to produce a more sensitive instrument.

No conclusions can be drawn from the cross-lagged paned correlations which were observed in an attempt to infer cause and effect relationships. An hypothesis often offered (Aiken, 1970) is that the causal relationships between students and teacher variables are two way. The results do not deny this possibility.

PERCEPTIONS AND ATTITUDES OF STUDENT TEACHERS IN MATHEMATICS

Gerald Kulm
Purdue University

1. Purpose

The purpose of this study was to investigate the effects of the skill of the supervising teacher, the type of mathematics program, and the type of community on the attitudes of student teachers in secondary school mathematics.

2. Rationale

In the school, the supervising teacher exerts an important influence upon the attitudes formed by a student teacher. Edgar (1972) reported that student teachers who had strong positive feelings toward their supervisors had greater increases in their feelings of autonomy than those having less positive feelings. These student teachers were also more satisfied with the evaluation of their teaching methods and their methods of disciplining students than those who had less positive feelings. Graening (1972) found significant positive correlations between student teachers and supervising teachers in teaching activities, teaching strategies, and perceptions of what should occur in teaching.

A more general effect in attitude formation is the teaching philosophy of the school or mathematics department where the student is teaching. Many schools have innovative programs that may have important consequences in forming attitudes toward discovery or individualized teaching approaches. The type of community in which the school is located can be important in forming attitudes, particularly if the student teacher has not had previous experience in a similar community.

3. Research Design and Procedure

A test was developed to assess perceptions and attitudes related to three aspects of mathematics teaching.

Subtest I (perceptions of pupil's attitudes toward mathematics). This subtest was developed by rewording some of the questions from the Ideas and Preferences Inventory of the National Longitudinal Study of Mathematics Abilities (Wilson, Cahen and Begle, 1968).

Subtest II (Attitudes toward discovery teaching). Items for this subtest were derived from the Collier (1972) Beliefs About Mathematics Instruction Scale.

Subtest III (Attitude toward the teaching profession). Items for this subtest were written by the author.

The subjects for the study were 34 secondary mathematics teaching majors who had just completed 6 weeks of student teaching in 25 Indiana junior and senior high schools. The subjects responded to each of the 27 items of a five-choice scale, from strongly agree to strongly disagree.

The director of mathematics student teaching rated each of the schools according to the Type of Mathematics Program (traditional or innovative) and the Type of Community (urban, suburban, or rural). He also rated each supervising teacher according to his effectiveness as a supervisor (average, good, excellent).

4. Findings

An internal consistency reliability estimate for each attitude test was computed by substituting the square of the average item-test correlation into the Spearman-Brown formula. The resulting coefficients were: Subtest I, .79; Subtest II, .64; Subtest III, .60; Total test, .73.

The scores on each of the three subtests and the total test were used as dependent variables in unequal cell analyses o. variance.

The difference between the means on both Subtest III and on the total test favored the traditional program; that is, student teachers who taught in a traditional mathematics program had a better attitude toward the teaching profession and a better total attitude toward mathematics teaching than those teaching in an innovative program ($p < .05$).

The second ANOVA showed that students teaching in a traditional program had more favorable perceptions of students' attitudes toward mathematics ($p < .05$). A Newman-Keuls test showed that there was a significant difference ($p < .05$) only between average and good supervisors. Students working with average supervisors had a better attitude toward the teaching profession than those working with good supervisors.

5. Interpretations

The test that was developed for this study appears to provide sufficiently reliable information about these perceptions and attitudes to assess beliefs held by groups of subjects. The number of subjects tested was not large and there may have been common beliefs and attitudes due to common preparation in the methods course. These factors must be kept in mind when interpreting the validity and application of the test.

There may be several explanations of the result that a traditional program produced better perceptions and attitudes on two of the Subtests and on the Total test. Most of the student teachers were probably taught mathematics in a traditional program and, hence, may have perceived students as working harder than students in an innovative program. Some of the innovative programs permit students to be more free in the classroom which may be seen as a discipline problem for an inexperienced teacher. This explanation supports Jenkins (1972) who found that student conversations and students wasting time predicted 70% of the variability in teachers attitudes toward students.

The result that student teachers working with average supervisors have a better attitude toward the teaching profession than those working with good supervisors may also be due to several factors. A student working with a good supervisor may have seen most other teachers as being less competent and the profession less respected. Another view is that a good supervisor in the eyes of the rater may be one who criticizes and actively engages the student teacher in an examination of his teaching. If the student is not able to accept criticism, he may form unfavorable attitudes toward the profession. The excellent supervisors may have been able to accomplish this criticism in such a way that negative attitudes were avoided and the average supervisors may have provided little feedback for the student teacher. The result that the type of community was not significant in the perceptions and attitudes of the student teachers tested may have been due largely to the fact that most students were placed near their homes so that they were familiar with the type of location in which they taught.

The innovative programs apparently did not induce enthusiasm for discovery teaching. The discovery approach was encouraged in the methods course but perhaps the student teachers did not feel sufficiently confident to use it enough to develop strong attitudes for or against discovery teaching.

Investigator's Notes

The results of this study should not be taken to mean that the best results will be produced by placing student teachers only in traditional schools with average rather than excellent supervising teachers. Instead, more specific attention needs to be given to teacher training methods that prepare students to teach in innovative instructional systems. Much of the content of methods courses and other training procedures are aimed at more traditional modes of instruction.

A master teacher is not necessarily a master supervisor. A common complaint of student teachers is that supervising teachers either criticize too much or too little. Since the supervisor has a significant effect on student teachers' perceptions, teachers must be trained in supervising skills.

THE EFFECTS OF A MATHEMATICS LABORATORY APPROACH ON STUDENTS' SELF CONCEPTS, ATTITUDES AND ACHIEVEMENT

David L. Pagni, Fred Sharman and John Randolph
California State University Fullerton

1. Purpose

To evaluate a mathematics laboratory curriculum project.

2. Rationale

The traditional way of teaching mathematics has been the lecture approach where the teacher is the only active participant in the learning process. This type of teaching permits limited interaction between the student and the teacher and even less interaction between the student and his learning environment. Books have been provided and students have been required to work through them. Though mathematically sound, these texts have failed to meet the needs of the student who lacks the ability to think abstractly, make interpretations, or who is unable to read at his grade level. Because of his lack of success with this type of curriculum and with school in general, the low achiever becomes frustrated and eventually gives up.

The need for a change in the traditional approach to mathematics was apparent when the achievement test results were examined for average and below average students at City Junior High School in 1969-70. Half of all 9th grade students did not show average growth in arithmetic fundamentals between 7th and 9th grades. Fully a third of the students could not meet the 8.0 grade equivalent requirement when tested in grade nine. The problem centered around the question, "Why is it that so many students who are exposed to arithmetic for many years fail to grasp the fundamentals of the subject?"

The designers of this curriculum project felt that the mathematics laboratory approach to studying mathematics would be a more viable way of meeting the needs of these youngsters. A curriculum based on the following learning precepts was developed and evaluated:

(a) Learning is based on experience.

(b) Sensory learning is the foundation of all experience and thus, the heart of learning.

(c) Learning is enhanced by motivation.

(d) Learning proceeds from the concrete to the abstract.

(e) Learning requires active participation by the learner (Hilgard, 1956).

The Math Laboratory at City was designed to offer students the opportunity to become involved in mathematics and to interact with their learning environment while the teacher assumed the role of coordinator of the learning experiences. Each student was pretested and assigned projects which remediated his particular areas of weakness. However, other aspects of the mathematics curriculum were included for all students. The students were able to work independently and were not necessarily all working on the same project at the same time. They were able to progress at their own rate. There was free mobility within the lab and students were allowed to work together.

Some of the components of the City Math Laboratory are:

(a) Electronic Calculators: These machines have a greater ability to arouse student interest in doing mathematics than any other device.

(b) Local Business Problems: The use of actual business forms and letterheads helps the students see the practical side of mathematics, thus resulting in increased student motivation.

(c) Flow Charting: A flow chart provides a pupil with a logical, pictorial diagram of a process - one to which he could refer at any time.

(d) Tapes: Both remediation and enrichment tapes were used.

(e) Experiments: Experiments allow the student to feel, see, and practice mathematics.

(f) Variety: The student is provided with many experiences and frequent success.

(g) Individualized Instruction: Individually prescribed skill building packets are available and are used successfully.

3. Research Design and Procedure

During the 1970-71 and 1971-72 school years, the mathematics laboratory approach was evaluated by comparing the progress of a selected groups of low achieving students with a similar group of students at another school on a pretest-posttest basis. The design corresponded to Campbell and Stanley's "Nonequivalent Control Group Design" (1963). In addition to the achievement test in mathematics, questionnaires on student self-concept and student attitude towards mathematics were administered. The one group pretest-posttest design was used for each questionnaire.

4. Findings

The experimental-control group results are summarized below.

	No. of Pupils	Pretest Mean	Posttest Mean	Difference	Std. of Dev. of Difference
Experimental Group	38	18.34	23.79	5.45	5.05
Control Group	34	12.77	14.32	1.55	5.19

The difference in achievement shown by the experimental group was significantly (α = .001) higher than the control group difference.

The following table summarizes the evaluation results of the City Mathematics Laboratory, one group pretest-posttest design:

Test or Questionnaire	AVERAGES Pretest	Posttest	Difference	SIGNIFICANT Positive Change?	Level of Significance
Mathematics Achievement Test	18.342	23.789	5.447	Yes	= .0001
Mathematics Attitude Questionnaire	11.839	13.607	1.768	Yes	= .005
Self-Concept of Ability Mathematics	15.792	18.563	2.771	Yes	= .0001
Self-Concept of Ability English	17.583	19.792	2.208	Yes	= .001
Self-Concept of Ability Social Studies	17.851	17.277	- .573	No	------
Self-Concept of Ability Science	19.846	17.385	-2.461	No	------
General Self-Concept of Ability	17.314	17.882	.567	No	------

5. Interpretations

The results indicated that the laboratory approach used at City had positive effects on students' self concepts, attitudes and achievement. It was especially heartening to find that students had developed higher self concepts toward mathematics - the reverse of a trend established by the "weeding out" process of mathematics instruction as students progress through the grades.

REFERENCES

Campbell, Donald T. and Julian C. Stanley. "Experimental and Quasi-Experimental Designs for Research on Teaching", Handbook on Research on Teaching, ed. by N. L. Gage, Rand McNally, 1963, pp. 218-19.

Hilgard, Ernest R. Theories of Learning, Appleton-Century-Crofts, 1956, pp. 485-487.

April 26, 1973
Thursday
4:30-5:30 p.m.

RESEARCH REPORTING SECTION III

Presider: NORMA G. HERNANDEZ, The University of Texas at El Paso
El Paso, Texas

Reports: OBJECTIVES AND FORMATIVE EVALUATION AS FACTORS IN MATHEMATICS ACHIEVEMENT AND ATTITUDES OF COLLEGE STUDENTS IN A NUMBER SYSTEMS COURSE

GRAYSON H. WHEATLEY, Purdue University
Lafayette, Indiana

ROBERT B. KANE, Purdue University
Lafayette, Indiana

GERALD KULM, Purdue University
Lafayette, Indiana

LEARNING THROUGH TEACHING

WILLIAM B. MOODY, University of Delaware
Newark, Delaware

R. BARKER BAUSELL, University of Delaware
Newark, Delaware

A COMPARISON OF VERBAL TEACHING BEHAVIORS IN SEVENTH GRADE MATHEMATICS CLASSES GROUPED BY ABILITY

H. BERNARD STRICKMEIER, California Polytechnic State University
San Luis Obispo, California

RALPH W. CAIN, The University of Texas at Austin
Austin, Texas

VALIDATING BEHAVIOR CATEGORIES IN STANDARDIZED TESTS

MILDRED E. KERSH, University of Washington
Seattle, Washington

OBJECTIVES AND FORMATIVE EVALUATION AS FACTORS IN MATHEMATICS ACHIEVEMENT AND ATTITUDES OF COLLEGE STUDENTS IN A NUMBER SYSTEMS COURSE

Grayson H. Wheatley, Robert B. Kane, Gerald Kulm
Purdue University

1. Purpose

The purpose of this study was to answer the following questions:

(a) Is there an effect on achievement and/or attitudes in the learning of mathematics when students are given written statements of the instructional objectives of the course?

(b) Is there an effect on achievement and/or attitudes in the learning of mathematics when nongraded formative evaluations with and without recommendations are utilized?

(c) Is there a combined effect of using both objectives and formative evaluations?

2. Rationale

Carroll (1963) and Bloom (1968) have developed a view of instruction described as mastery learning. Many studies (Block, 1971) have recently been conducted which investigate various aspects of the Bloom model. In the Bloom mode, a case is made for informing the learner of the task to be mastered; this is usually accomplished by providing the learner with a list of the behavioral objectives for the material to be learned. Also it is recognized that the learner needs feedback while he is learning. Provisions for feedback, called formative evaluations, also allow the instructor to determine the nature of the next step in the learning process for each individual. On the basis of the formative evaluation, recommendations for restudy and relearning are usually given to the student. In his initial conceptualization of mastery learning, Bloom (1968) gave much attention to individual rates of learning. Following Carroll, aptitude is defined as the time it takes the learner to reach mastery. This is in contrast to the usual practice of grading on the curve or grading with respect to class performance which gives the same time for all students and thus distributes students on achievement about the same as their aptitude. Bloom contends that with the proper instruction and provisions for learning rates, nearly all students can achieve mastery in a course.

Collins and Wheatley (1972) investigated the effect of using objectives, formative evaluation and recommendations with junior high mathematics classes. They found that each of these treatments, both singly and combined, contributed significantly to student achievement. The present study was an attempt to examine the role of the same variables in a mathematics course for elementary education majors.

3. Research Design and Procedure

The subjects for this study were 285 college freshman at Purdue University during the 1971-2 school year. These students were enrolled in a number systems course with approximately 30 in a class. At the beginning of the semester all students took a pretest covering the prerequisites for the course. Mathematics SAT scores were also recorded for the students.

In all, there were six treatment groups over the two semesters. There were three variables considered: Objectives, formative evaluations, and recommendations. One purpose of the study was to determine the effect of giving weekly (every third class) formative evaluations. These were short sets of questions corresponding to the objectives considered for that period. These formative evaluations were not graded, they were scored and returned promptly. The third objective was to determine the effect of giving recommendations for restudy accompanying the formative evaluations. These took the form of identifying the objective tested, the pages in the text describing this objective and, in some cases, additional problems. The same set of recommendations was given to all students taking the formative evaluations. It was left to the student to use the recommendations as he or she wished. Some gave graded quizzes frequently.

Four instructors taught six sections each semester of Mathematics for Elementary School Teachers I, a number systems course. During the fall semester two classes received formative evaluations, two received formative evaluations with recommendations and two classes received no treatment (control). In the spring two classes received objectives and formative evaluations with recommendations; two received objectives and formative evaluations and two received objectives only. This design is summarized in Figure 1.

Figure 1.
Description of Treatment Groups

Fall	Spring
FE	O, FE, R
FE, R	O, FE
Control	O

4. Findings

The Data were analyzed using an analysis of covariance model with pretest scores and Math SAT scores as covariates. The criterion measures were hour exams I and II, the final examination and an attitude scale constructed for use in this study. On the separate one-way ANCOVAs, significant differences were found only for the first hour exam in favor of the group which received all three treatment effects. In all other cases the F values were nonsignificant. A 2 x 2 (objective by formative evaluation) factorial ANCOVA was used to determine the explicit effects of objectives and formative evaluation. There were no significant differences for any of the F's computed. In summary, there was some evidence of a combined effect of the three treatment groups but generally the treatment groups did not perform differently from control classes.

5. Interpretations

There are several aspects of the design and execution of this study which should be considered when interpreting the results. The most significant of these is the lack of full provisions for individual rates of learning. Carroll and Bloom place this factor in a central position in describing a mastery learning model. If individual rates of learning are provided for, the correlation between pretest or other aptitude measures should be near zero since nearly all students should be able to reach mastery. In this study the correlation between pretest and final exam scores was 0.58 which indicates that mastery was not fully achieved. In this study little attention was given to individual rates of learning. When recommendations for remediation were provided it was then left up to the students to spend the time necessary to master the objectives. It is entirely plausible that the effect of the three treatment variables was dampened considerable by the lack of provisions for individual rates of learning.

The second factor relates to the nature of the groups. In order to insure no use of objectives by certain classes, treatment groups using objectives were all in the spring semester. The analysis of data indicated that these classes tended to be of somewhat lower aptitude. While analysis of covariance partially accounted for this, the results may have been different if all treatment groups had performed under the same conditions.

A third factor relates to the nature of the recommendations. In this study all students were given the same general set of recommendations. Ideally the recommendations could follow the form of a prescription tailored to each student and his particular needs.

Generally the results of this study showed no effect of giving objectives to students, or of using weekly formative evaluations on either achievement or attitudes. There was one notable exception; on the first hour exam the classes receiving objectives, formative evaluations and recommendations scored higher than the other groups. While this finding supports the use of these components of mastery learning, the advantage was not noted for the second hour exam or the final exam. If a completely random design had been possible together with provisions for individual rates of learning the results may have been different. In light of the many studies showing positive effects from using these treatments and the encouraging results on hour exam I it is recommended that additional research be conducted on the use of such treatments in university mathematics courses.

REFERENCES

Block, James H. Mastery Learning: Theroy and Practice. New York: Holt, Rinehart and Winston, Inc., 1971.

Bloom, B. Learning for Mastery, UCLA-CSEIP Evaluation Comment: No. 2, 1968.

Carroll, John. A Model of School Learning, Teachers College Record, 64, (1963), 723-33.

Collins, Kenneth and Wheatley, Grayson. An Investigation of Variables of Mastery Learning for Teaching Mathematics. Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1972.

LEARNING THROUGH TEACHING

William B. Moody and R. Barker Bausell
University of Delaware

1. Purpose

The purpose of this study was to determine the effect upon achievement of a strategy of teaching a version of the material to be learned. The material consisted of a specially developed number theory unit and an exponential notation unit presented as part of a college freshman level mathematics course. The subjects were prospective elementary school teachers. The subjects taught the units to fourth grade students.

2. Rationale

Prospective elementary school teachers are commonly provided with opportunities to both gain teaching experience and to learn the basic academic subject matter deemed necessary to teach the various disciplines. If the old addage is true that the best way to learn a topic is to teach it, then an excellent empirical rationale would be supplied for combining the two processes. In fact if mathematical content could be structured in such a way that college students could learn it through teaching it to small groups of elementary school children, then three pedagogical objectives could theoretically be achieved through one process: teachers could theoretically (1) learn necessary subject matter content and (2) gain teaching experience at the same time that (3) elementary school students were learning topics relevant to their needs.

Before so ambitious a teacher education program is instituted, however, several questions need to be answered. Among them are (1) can mathematical content be structured in such a way that it is relevant to both college and elementary school students, and (2) can students learn that mathematical content through teaching it? The present study was addressed to answering these two questions.

3. <u>Research Design and Procedure</u>

Volunteers were solicited from a freshmen level mathematics course designed for prospective elementary school teachers to teach two 30 minute units to fourth grade students. These 20 Ss served as the teachers in the present experiment; the remaining 50 Ss who did not volunteer served as the non-teaching control group.

Two mathematics units were constructed, one containing 10 instructional objectives for teaching fourth grade students elementary number theory, one containing 10 objectives involving exponential notation. Besides the 10 instructional objectives, each unit contained (1) an example of the type of item which the fourth grade students would be expected to answer as a result of the instruction, and (2) a mathematical development of each objective designed to supply the freshmen teachers with the necessary mathematical background to teach the topic.

A 51 item test was constructed based on the contents of the two units. Twenty-three items were designed to measure mastery of the number theory unit; twenty-eight were based on the exponential notation development. The test was constructed in such a way that the answers to all items could be found in one of the two units. The combined tests had a split-half reliability of 0.89, while the number theory and exponential subtests had reliabilities of 0.86 and 0.76 respectively.

All Ss in the mathematics course were given both experimental units and told that they would be tested on their contents in two weeks. Ss were told that this procedure was being employed in order to save class time. They were told that the results would count toward their final grade in order to insure adequate motivation. Neither topic was discussed in class.

The 20 teachers were scheduled to teach the two different trials over the course of the following two weeks. Teachers were not told until immediately before each instructional trial began which topic they were to teach, hence theoretically forcing them to prepare to teach both units on both trials. Teachers were randomly assigned to teach either number theory or exponents on the first trial, while all teachers taught the exponential unit on their second trial. The instructional trials lasted exactly 30 minutes each, during which each teacher taught two 4th grade students. At the end of the two week interval all Ss, both teachers and non-teachers, were administered the 51 item number theory-exponent test during their regular class period at the previously announced time.

4. Findings

A two group (teacher versus non-teacher) analysis of covariance was performed on the combined number theory-exponent test using Ss grade point averages and performances on the course's midterm exam as covariates. The two measures correlated 0.37 and 0.66 respectively with the dependent variable.

The $F = 7.3(1,66)$ was significant at the .01 level indicating that those Ss who instructed elementary school students mastered more of the mathematical content contained in the two units (adjusted $\overline{X} = 42.5$) than did Ss who simply prepared for the test on their own (adjusted $\overline{X} = 38.0$).

5. Interpretations

The present study was not designed to determine why the teachers learned more than the non-teachers, although the most parsimonious answer is that they simply studied the unit more since they were responsible to teach it to children. This interpretation was partly borne out by the fact that teachers who taught the number theory unit did not score higher on the number theory subtest than teachers who did not teach that unit. By the same token, no trend was observed for teachers who taught the exponent unit on both trials to score higher on that subtest.

Investigator's Notes

The present study was only heuristic in nature. There are far too many unanswered questions for it to constitute an empirical rationale for the implementation of the previously proposed teacher education program. It is not known, for example, whether an individual would learn more by teaching or by simply being a student in a traditional college classroom in which the faculty member is capable of serving as a more viable information source than was present for students in this study. It is also not known whether teaching would continue to serve as a study incentive once its novelty had eroded. It is hoped, however, that the results of the present study will suggest that further evaluation is worth undertaking.

The necessity on the part of the authors of using volunteers rather than randomly assigning _Ss_ constitutes a definite methodological weakness, one which the use of the covariance procedure only partially mitigates. Although it would be difficult to argue that learning differences as large as were observed could be a direct function of non-equivalence between the teacher and non-teacher groups after differences on two learning measures (GPA and midterm hourlies) had been partialled out, the superior learning on the part of the teachers in the study could have been a function of an _interaction_ with a variable contiguous with volunteerism and the independent variable (teaching). Even if the results are so attributable, however, a rationale for taking advantage of the interaction is provided by providing students with opportunities to learn certain topics through teaching them _if they so desire_, which, in the case of prospective teachers, is only a special instance of learning through doing.

A COMPARISON OF VERBAL TEACHING BEHAVIORS IN SEVENTH GRADE MATHEMATICS CLASSES GROUPED BY ABILITY

H. Bernard Strickmeier
California Polytechnic State University

Ralph W. Cain
The University of Texas at Austin

1. Purpose

The primary purpose of this study was to compare patterns of teachers' verbal behaviors in mathematics classes grouped by ability. A second purpose was to determine if teachers have different expectations of behaviors within classes of different ability levels and it teachers' perceptions of their verbal behaviors in a class are to some extent dependent on the ability level of the class.

2. Rationale

One argument advanced in favor of ability grouping has been that grouping makes possible with each group the use of instructional materials and methods particularly suited for the ability range within the groups. Although many studies concerning the effects of grouping have been conducted, little data has been collected concerning what, if any, differentiation of teaching method accompanies grouping. The present study was designed to determine if teachers varied a particular aspect of teaching method, verbal behavior, for classes of different ability level.

3. Research Design and Procedure

Ten classroom teachers, each teaching at least one high ability and one low ability section of seventh grade mathematics, were selected for this study. An initial meeting was held with each teacher to explain the procedures involved in the study and to collect information concerning the background of each teacher. Also at this meeting each teacher was asked to complete a fourteen item questionnaire to determine his expectations of behaviors within each of his two classes included in the study and to determine his perception of his verbal behavior in those classes. The teachers were not told the purpose of the study and they were not shown the instrument which was to be used to record verbal behavior.

Each teacher was observed by trained observers three times while conducting a high ability class and three times while conducting a low ability class. For each teacher the same high ability and the same low ability class was used for all observations. Observations were made in pairs, each subject being observed twice on the same day by the same observer - once in the high ability class and once in the low ability class.

The Observational Schedule and Record 5V was used for all observations to record and describe teacher verbal behaviors. The OScAR 5V is a twenty-four item instrument which records patterns of verbal interaction within a class. Five observers were trained in eight training sessions employing tape scripts, audio tapes and video tapes of actual mathematics classes. These observers then conducted the observations for this study.

The verbal behaviors of the ten teachers conducting low ability classes were compared with the verbal behaviors of the same ten teachers conducting high ability classes. The percentage of statements falling into each of the OScAR 5V categories served as measures of verbal behaviors. The percentages for the two groups were compared using a "t" formula for differences between correlated pairs of means. This same test was used to compare the amount of time devoted supervised study in high and low ability classes and to analyze the data collected concerning teachers' expectations and perceptions.

4. Findings

The analysis of the questionnaire data revealed that teachers did have different perceptions and expectations for the classes of different ability levels. Significant differences ($p < .05$) were found for the responses to eight of the fourteen questionnaire items. The analysis of the observational data indicated that the teachers' verbal behaviors were not different for the classes of different ability levels. Significant differences were found for only two of the twenty-four OScAR 5V categories. No significant difference was found the amount of time devoted to supervised study in high and low ability classes. A total of thirteen comparisons made between observed behaviors and differences in expectations and perceptions indicated by responses to the questionnaire revealed no significant differences in observational data which could be related to significant differences in questionnaire data.

5. Interpretations

The apparent lack of difference in verbal teaching behaviors between ability levels is the most significant finding of this study. Since much research indicates that if ability grouping is to have a positive effect on achievement, differentiation of teaching method must accompany this grouping, this study indicates that teachers practice a kind of self-deception concerning their verbal teaching behaviors. These findings may also indicate that teachers feel that they should differentiate their verbal teaching behaviors according to the ability level of the class, but are unable or unwilling to do so.

VALIDATING BEHAVIOR CATEGORIES IN STANDARDIZED TESTS

Mildred E. Kersh
University of Washington

(Editor's note: The following paper describes research which is currently in progress. The results of this research will be presented at the reporting section. A final paper giving findings and interpretations will be available in May, 1973 through the ERIC system. Readers may write to the SMEAC Clearinghouse for an individual copy of this final paper.)

1. Purpose

What behaviors are required by a student to answer successfully a mathematics item on a standardized achievement test? While most standardized tests do publish tables of specifications depicting a content/behavior matrix, the behavioral categories are usually validated by consulting content specialists. This study proposes to validate behavior categories by analyzing, through interviews, the behavior of students who are successful in completing a test item.

2. Rationale

The increasing frequency of use of formative tests in evaluating student progress through a mathematics curriculum emphasizes the importance of clearly specified and valid behavior categories. Ideally these formative tests are constructed to tell which specific content topics a students has mastered and to which behavioral level that mastery extends. Nitko (1971) maintains that construction of such criterion-referenced tests depends in large measure on a systematic plan which integrates specifying the content area and behaviors to be measured, constructing items that indeed test these behaviors, and distributing items in the behavior categories to achieve a representative sample of behaviors. He emphasizes that a test outline or table of specifications be maintained lest the validity of the test be affected by distorting the proper representation of behavior categories.

3. Research Design and Procedure

This study will examine a published table of specification of one standardized mathematics test, The Comprehensive Test of Basic Skills, to determine the validity of higher level cognitive categories. Higher level cognitive categories are defined, for the purposes of this study, to be those above Knowledge and Comprehension of the Bloom Taxonomy. The sub-test Concepts and the sub-test Application of the mathematics battery of the Comprehensive Test of Basic Skills will be administered to fifth grade students in a suburban elementary school and in an inner-city elementary school. Students who successfully complete items labeled by the publisher as requiring higher level cognitive behavior will be interviewed. In the taped interview the student, after his own correct answer selection and work sheet are returned, will be asked to show the investigator how he obtained the answer. The student's interview will then be analyzed and his responses categorized for behavioral level.

Since this investigation will take place in the Fall and Winter of 1972, no results or conclusions can be presented in this proposal. These data will, however, be presented at the NCTM Research Reporting Sections and will be available later through the

ERIC system.

4. Proposed Analysis

Analysis of the interviews may indicate that there is an acceptable correlation between behavior categories hypothesized by content specialists and behaviors actually exhibited by successful students. If so, test publishers would be justified in their use of expert opinions rather than costly validation procedures to categorize behaviors. Low correlations might have several results. First, the test might be revised to include items that do test higher cognitive levels, items that have been validated by some procedure other than opinion of content specialists. Or, second, the published table of specifications might be revised to indicate valid behaviors. The test itself would remain unchanged with its interpretation undergoing revision. As Cronbach (1969) has stated, "What one has to validate is a proposed interpretation of the test...."

Thirdly, investigation by interview will probably reveal that finding an answer involves more than one behavioral process. Moreover, the task of obtaining a correct response can perhaps be accomplished in a variety of ways. A description of these modes or patterns of processes could be used to understand the operational schema of individual students or classes of students. The central question of the study remains, however, focused on the behavior required by a student to produce a correct response.

Validation of a test requires the integration of many kinds of information. Data obtained by student interview could become an important source of such validity data.

REFERENCES

Cronbach, Lee J. "Validation of Educational Measures," Proceedings of the 1969 Invitational Conference on Testing Prcblems: Toward a Theory of Achievement Measurement. p. 36.

Nitko, A. J. "A Model for Criterion-referenced Tests Based on Use," ERIC: ED 049 318. February, 1971.

April 27, 1973
Friday
12:00-1:00 p.m. RESEARCH REPORTING SECTION IV

Presider: RALPH W. CAIN, The University of Texas at Austin
Austin, Texas

Reports: TRANSFORMATION VS. NON-TRANSFORMATION TENTH-GRADE GEOMETRY: EFFECTS ON RETENTION OF GEOMETRY AND ON TRANSFER IN ELEVENTH-GRADE MATHEMATICS

ANTHONE P. KORT, Niles East Township High School
Skokie, Illinois

DISCRIMINATORS FOR GEOMETRY ACHIEVEMENT IN THE NLSMA Y-POPULATION

LARRY SOWDER, Northern Illinois University
DeKalb, Illinois

COMPUTER-AUGMENTED CALCULUS: ITS EFFECT UPON PERFORMANCE

FREDERICK H. BELL, University of Pittsburgh
Pittsburgh, Pennsylvania

IMPROVING A CUPM-RECOMMENDED COURSE FOR ELEMENTARY EDUCATION MAJORS BY USING A PRETEST-INDIVIDUALIZED INSTRUCTION-POSTTEST ROUTINE

ROBERT KALIN, Florida State University
Tallahassee, Florida

CHARLES A. REEVES, Florida State University
Tallahassee, Florida

TRANSFORMATION VS. NON-TRANSFORMATION TENTH-GRADE GEOMETRY: EFFECTS ON RETENTION OF GEOMETRY AND ON TRANSFER IN ELEVENTH-GRADE MATHEMATICS

Anthone P. Kort
Niles East Township High School
Skokie, Illinois

1. Purpose

The major purpose of the research was to evaluate an innovative, transformation approach to tenth-grade geometry. The evaluation was accomplished through an investigation and comparison of the effects of the transformation approach and a non-transformation approach to tenth-grade geometry on retention of geometry and on transfer in eleventh-grade mathematics.

2. Rationale

An earlier study by Zalman Usiskin evaluated the same transformation approach on the basis of initial learning. Both the transformation and the non-transformation geometries were esentially Euclidean, incorporating SMSG-suggested modifications. Whereas the transformation geometry materials, Geometry - A Transformation Approach by Coxford and Usiskin, used transformations extensively as an intergrative concept, the non-transformation materials included no transformation content.

Two major questions were asked:

1. In the year following the study of tenth-grade geometry, can advantages be found in a transformation-based geometry with respect to retention of geometry as compared with a geometry not

based on transformations?

2. Do students who study geometry via transformations have advantages in learning eleventh-grade mathematics?

Neither type of geometry was expected to have superior advantages for overall retention of geometry or for overall learning in eleventh-grade mathematics. However, there was a rationale for expecting that a transformation geometry background would be advantageous for certain concepts.

3. <u>Research Design and Procedure</u>

The 184 subjects were in the eleventh-grade mathematics classes of Niles West Township High School in 1969-1970. In 1968-69, all subjects studied tenth-grade geometry at Niles West either by the transformation or the non-transformation approach. All possible subjects who had studied the transformation approach were selected as the experimental group. In each eleventh-grade mathematics class, a random selection of subjects who had studied the non-transformation approach was made so that each class and teacher contributed the same numbers of subjects to the control group as to the experimental groups.

The testing schedule was as follows:

(a) geometry posttest: June, 1969

(b) advanced algebra pretest: September, 1969

(c) geometry retention tests: January, 1970

(d) attitude inventory and eleventh-grade mathematics achievement tests: May and June, 1970.

Suitable standardized tests were used when possible. Otherwise, researcher-made tests were used.

Since data was collected for several independent and several dependent variables, a multivariate analysis of regression and co-variance was used. Two separate substudies were made for geometry retention and two for transfer in eleventh-grade mathematics. In two substudies, a 2 x 2 factorial design was used. The factors were the type of tenth-grade geometry studied (transformation vs. non-transformation) and the level of the eleventh-grade mathematics class (above average vs. average). In two substudies, only the type of geometry factor could be used. The rejection level for null hypotheses was .05.

4. Findings

The following results were obtained:

(a) Overall retention: The only significant difference was between the above-average and average control groups with the above-average group superior.

(b) Retention of congruence, similarity, and symmetry: The experimental group was significantly superior.

(c) Attitudes towards mathematics, achievement in eleventh-grade algebra, and achievement in relations and functions: There was no significant difference in the multivariate sense; in the univariate sense, the experimental group was significantly superior on the relations-functions test.

(d) Achievement in trigonometry, inverse relations and functions, and inverse circular relations and functions: The only significant difference was on the inverse relations-functions test with the experimental group superior.

5. Interpretations

The major conclusion was that tenth-grade geometry can be taught via transformations with possible advantages for retention of geometry and transfer in eleventh-grade mathematics. The primary implication was that tenth-grade geometry should be changed to extensively utilize transformations only if subsequent mathematics courses are altered to capitalize on a background in geometric transformations.

DISCRIMINATORS FOR GEOMETRY ACHIEVEMENT
IN THE NLSMA Y-POPULATION

Larry Sowder
Northern Illinois University

1. Purpose

The purpose of this study was to attempt to identify "best" discriminators between high and low geometry achievers, using data from the NLSMA Y-population. (A "best" discriminator in this study means a discriminator which was entered early into the analysis on the basis of giving the greatest decrease in the ratio of within to total generalized variance.)

2. Rationale

Several studies have been devoted to the prediction of success in high school geometry. Predictors usually chosen have been measures of previous achievement, IQ measures, geometry aptitude scores, teachers' estimates of success and, more recently, reading scores, structure-of-intellect measures, and student prediction of his grade. The availability of the NLSMA data-bank made it possible to examine data on many such variables from students across the country.

3. Research Design and Procedure

Data came from the NLSMA Y-population geometry students. 16446 data-cases were screened for missing data, for geometry achievement scores which permitted classification as a high (H) or low (L) geometry achiever, and for IQ measures which indicated that the subject was not extremely incapable or capable. The 1856 data-cases remaining were sorted by school, and equal numbers of cases were selected for each of the H and L groups within a school, depending in which group was more numerous. This sampling gave 310 cases for each group.

Data for these cases included 7 achievement measures, 2 IQ measures, 7 cognitive measures, 7 affective measures, and measures of geometry achievement (see Table 1). Subsets of these measures were analyzed for the H and L groups in several stepwise discriminant analyses (Biomed BMDO7M).

4. Findings

Each subset of measures chosen did discriminate between the two groups (.01) and correctly assigned 75-95% of the cases to the H or L groups.

As indicated by early entry into the analyses, 4 achievement measures, 3 cognitive measures, and a measure of debilitating anxiety were "best" discriminators. These measures are indicated by asterisks in Table 1.

5. Interpretations

The prominence of the achievement measures as discriminators is not unusual; noteworthy here is that one 5-item scale (Informal Geometry) was uniformly outstanding as a predictor. The 3 cognitive measures cited were designed to measure either spatial imagery manipulation or the ability to "handle a novel mathematical situation." That a debilitating anxiety measure discriminated serves as a reminder to us that evaluation-by-test may be unfair to some students.

TABLE I

Scales Used in the Study

Name	NLSMA Code	Maximum Score
Achievement measures		
Numbers 2	Y501	8
*Algebraic Expressions 3	Y502	22
*Algebraic Equations 4	Y503	18
Algebraic Inequalities	Y504	6
Graphs	Y505	6
*Informal Geometry 4	Y506	5
Analysis 3	Y507	12
Intelligence measures		
Lorge-Thorndike Verbal Raw Score	PY426	39
Lorge-Thorndike Non-Verval Raw Score	PY427	58
TOTAL IQ = PY426 + PY427		
Cognitive scales		
Card Rotations	PY608	112
*Paper Folding 2	PY609	10
Necessary Arith. Op.	PY610	15
Nonsense Syllogisms	PY611	15
Letter Se' :-Part 1	PY612	15
*Letter Puzzles 2	Y601	15
*Maps 2	Y602	17

(Table I continued)

Name	NLSMA Code	Maximum Score
Affective scales		
Pro-Math Composite	PY409	52
Ideal Math Self Concept	PY411	48
Facilitating Anxiety	PY415	45
*Debilitating Anxiety	PY416	50
Actual Math Self Concept	PY417	48
Orderliness	PY418	54
Messiness	PY419	48
Geometry achievement		
Sum of Y701 through Y707		45

COMPUTER-AUGMENTED CALCULUS: ITS EFFECT UPON PERFORMANCE

Frederick H. Bell
University of Pittsburgh

1. Purpose

The primary objective of this study was to assess the effectiveness of a computer-augmented approach to learning calculus fundamentals. In meeting this objective the following hypotheses were tested:

Students who study calculus using a computer-augmented approach do not perform differently than do students who do not augment their study with a computer, when performance is defined by

(1) achievement on a criterion test (Skill Posttest) administered immediately following instruction and designed to measure skill in applying techniques of calculus,

(2) achievement on a criterion test (Concept Posttest) administered immediately following instruction and designed to measure understanding of calculus concepts,

(3) achievement on an alternate form of the Skill Posttest administered one month following instruction (Skill Post Posttest),

(4) achievement of an alternate form of the Concept Posttest administered one month following instruction (Concept Post Posttest).

A secondary objective was to gather data regarding:

(5) the relative usefulness of various sources of assistance in learning computer programming,

(6) the relative usefulness of various sources of assistance in learning calculus,

(7) changes in interest in calculus and programming during the experiment,

(8) the students' evaluations of increased interest in the course as a consequence of using a computer.

2. Rationale

The widespread use of computers as instructional tools in mathematics classrooms coupled with increasing demands for accountability based upon performance criteria to justify public financial support of instruction necessitates studies to determine the effectiveness of various instructional components. It has been found that writing and executing computer programs assists secondary school students in mastering certain algebraic skills and concepts (Washburne, Robert M. Dis. Abst. 30A: 5179; June 1970.)

3. Research Design and Procedure

This study was carried out in 1969 using a freshman level mathematics course in the College of Agriculture at Cornell University; the duration of the study was 10 weeks. The subjects included both men and women of freshman, sophomore, junior, and senior status. Forty-nine students registered in the morning section (Control Group) studied calculus using group lectures and a calculus manual prepared by the researcher. Forty-six students registered in the afternoon section (Experimental Group) used a similar manual and were given the same lectures. The manuals differed in only one respect; the experimental materials contained six computer-augmented calculus projects to be completed by writing and executing computer programs. The control materials contained six comparable calculus projects to be completed without using a computer, together with six pre-calculus projects to be completed using a computer. Since neither group knew computer programming at the beginning of the experiment, it was necessary to teach the Experimental Group how to program, consequently, to avoid different treatments relative to this variable, the Control Group was also instructed in programming techniques.

Immediately after completion of the six projects, both groups were given the Skill Posttest followed by the Concept Posttest. Four weeks later the Skill Post Posttest and Concept Post Posttest were administered. The researcher "blind graded" the Concept tests and an assistant graded the Skill tests.

To obtain "alternate forms" test reliability indices, Fisher product-moment correlations were computed for the Skill Posttest and Skill Post Posttest scores ($r_{tt} = .669$), and for the paired Concept test scores ($r_{tt} = .661$). These four tests were constructed to correspond to the performance objectives of the course, the course lectures, homework assignments, and the manuals. In addition a panel of experts concurred that the Skill tests did measure skills and the Concept tests measured concepts. Consequently the four tests were judged to have face validity.

Analysis of covariance, with SAT Math scores taken as measures of the concomitant variable, was used to test hypotheses (1) through (4). To obtain data relevant to the secondary objective, the researcher prepared a questionnaire which was conpleted by the students in the final class meetings. Data obtained for (5) and (6) were used to rank sources of assistance according to the means of the coded responses. Data for (7) and (8) were summarized as mean responses to questionnaire items requesting expressions of interest in computer programming and calculus both before and after the course.

4. Findings

In testing hypotheses (1) through (4), analysis of covariance indicated no significant differences in performance between Control and Experimental Groups on either the Skill Posttest or the Skill Post Posttest, $F = 1.039$ and $F = .795$ respectively. Analysis of covariance on data from the Concept Tests yielded significant differences ($p < .01$) in favor of the Experimental Group on both tests; $F = 12.532$ for the Concept Posttest and $F = 7.769$ for the Concept Post Posttest.

A summary of findings for (5) and (6) are shown in Table 1.

Table 2 contains a summary of student responses relative to (7).

Relative to objective (8), students responded to the questions:

(a) How would you evaluate the contribution that the introduction of computer topics made to your general interest in this course?

(b) Which of the following best describes your attitude toward including computer topics with the other topics in this course? Possible responses to both of (a) and (b) were: 0-unfavorable, 1-mildly favorable, _ indifferent, 3-mildly favorable, 4-strongly favorable. Mean responses to question (a) for the groups were: 2.942 (Control Group) and 3.093 (Experimental Group). For question (b) the mean responses were: 2.769 (Control Group) and 2.963 (Experimental Group).

Table 1

Usefulness of Sources of Assistance in Computer Programming and Calculus*

Source of Assistance	Control Group (mean response)	Experimental Group (mean response)
	Computer Programming	
Programming Manual	1.62	1.35
Class lectures	2.37	2.50
Hardware use sessions	1.83	1.61
Evening review sessions	1.48	1.17
Office hours of professor	.96	1.13
Other students in course	1.31	1.17
Students not in course	1.00	1.13
	Calculus	
Calculus manual	2.51	2.70
Class lectures	2.56	2.54
Writing computer programs	.67	1.50
Evening review sessions	.83	.65
Office hours of professor	.79	.68
Other students in course	.89	.85
Students not in course	.71	.65

* 0 indicates no use, 1 little use, 2 fair use, 3 much use.

TABLE 2

Change in Interest in Computer Programming and Calculus During the Study*

Control Group		Experimental Group	
Mean interest before study	Mean interest after study	Mean interest before study	Mean interest after study
Programming 1.10	Programming 1.87	Programming 1.50	Programming 2.07
Calculus 1.08	Calculus 1.87	Calculus .94	Calculus 1.06

* 0 indicates no interest, 1 little interest, 2 moderate interest, 3 much interest.

5. Interpretations

The findings of this study support the hypothesis that writing and executing computer programs does aid students in understanding concepts of calculus and does not interfere with mastery of skills. Students participating in the study responded favorably to the inclusion of computer-oriented problems in the calculus course, indicating that it increased their interest in both computer programming and calculus.

Students found the textbook and class lectures to be the most useful sources of assistance in learning calculus, and computer programming was designated to be of little use, contrary to test results. Lectures and demonstrations were found to be much more helpful in learning programming skills than were the programming language textbook and other sources of assistance.

IMPROVING A CUPM-RECOMMENDED COURSE FOR ELEMENTARY EDUCATION MAJORS BY USING A PRETEST-INDIVIDUALIZED INSTRUCTION-POSTTEST ROUTINE

Robert Kalin and Charles A. Reeves
Florida State University

1. Purpose

The purpose of this study was to seek answers to these questions:

(a) How much competence do elementary education majors at a particular university have in two topics which occur in both CUPM and previously studied courses? Is it such that no further study is necessary?

(b) How accurately do these students perceive their competence in these two mathematical topics?

(c) Would a routine of diagnostic test followed by individualized assignment and independent study bring students to the required level of competence while recognizing their prior experience and competence?

2. Rationale

Ten years after adopting the course recommendations of the Committee on the Undergraduate Program in Mathematics, the Florida State University Department of Mathematics Education finds that there has recently developed some overlap between these courses and the mathematics our present elementary education majors studied while still in secondary school. Furthermore, students claim that this overlap is extensive, with some being quite vocal and positive in the claims.

On the other hand, very few of them do as well in their course examinations as such a situation would normally imply. Only a relatively small fraction of the students attempt an exemption examination, with very few passing it. And previous experience suggested that students believe themselves more competent than they really are; that indeed they are less competent than effective elementary school teachers should be.

These contradictions among content duplication, required competence, and student belief suggested that the three questions cited above be investigated.

3. <u>Research Design and Procedure</u>

The following procedure was developed for each of two two-week segments in the first mathematics course for elementary education majors:

(a) Administer a diagnostic pretest of twenty-two questions.

(b) Based upon a student's pretest errors, prescribe an individualized assignment to the text and to especially written supplementary units.

(c) Allow students to complete their individualized assignments over a ten-day period while the course proceeds in the regular manner with other content. These assignments are done independently without assistance except that a student may seek help voluntarily from graduate-assistant tutors.

(d) Administer a posttest of twenty-two questions which is equivalent in form and content to the pretest.

This routine was set up in accordance with the recommendations for competency-oriented modules currently in vogue. It was applied to two chapter-length topics: elementary set theory and modern numeration systems. It was tried with the sixty students in all three sections of the course offered in the Spring Quarter, 1972.

For comparison purposes, the diagnostic test in set theory was also administered to thirty elementary education majors who had satisfactorily completed all nine quarter hours of their mathematics requirement via conventional instruction, and were about to intern. For the same reasons the diagnostic test in numeration systems was administered to sixty-four elementary education majors who had successfully completed their study of this mathematics topic by means of conventional instruction in the preceding term.

4. Findings

The results may be summarized as follows:

(a) The diagnostic tests indicated that not quite one-fifth of these students did indeed know a great deal about each topic - more than their instructors had given them credit for. But about one-third of the group performed at an exceptionally low level. Furthermore, the entire group performed at a statistically significant lower level than the comparison groups.

(b) Students were not at all consistent in their perceptions of their knowldege of the two topics. Some overjudged their competence, some underestimated, few predicted very precisely. Their judgements were worse for the numeration systems than for the set theory.

(c) The individualized instruction approach brought the class as a whole up to satisfactory performance in both topics, indeed, a statistically significant better level than the comparison groups. But just less than one-fifth of the class was still performing at an unsatisfactory level.

5. Interpretations

These results indicated that students had every right to feel that they had covered this material before, and that conventional instruction was improper for, indeed insulting to them. Yet almost every one of the students needed to improve his mathematical competencies, although in widely varying ways and amounts.

The fact that most students misjudged their mathematical competencies tended to lower their morale. It is likely that those in the past who overjudged were not able under conventional instruction to recognize the need to make up some deficiencies. Those who underjudged likely became bored when the instructor explained things about which they already knew a great deal.

But the results also indicated that the individualized routine could help overcome these difficulties for at least 80% of the group. Being asked to study what you don't know while setting aside what you do know makes sense to the students. They like being treated as mature adults in a subject matter in which many such students normally have difficulties, indeed often have fears.

More needs to be done for the remaining 20% who experienced improvement, but not a sufficient amount. For this and other reasons, further developemnt is continuing along these lines:

(a) Instead of learning new content at the same time they do their individualized assignments, students will be given in-class help with their independent study. This will not be conventional instruction over two weeks, but rather tutoring and appropriately selected small-group work over one week. New individualized assignments are being written to challenge the best students and others to strengthen the weakest students.

(b) Experimentation is being done to determine if it would help to blend some pedagogy in with the mathematical theory. In particular, students will study the use of Z. P. Dienes' materials to teach place value.

(c) Pretests will be placed on computers for self-administration at the beginning of the unit. The test results and the individualized assignments will be given to each student before he or she leaves the computer terminal.

(d) This spring the revised materials will be tried out again with half of the students. The remaining half will be taught in the conventional manner, (no pretests, with classroom lectures assuming no prior knowledge on the part of the students). This will allow comparisons of achievement and morale not now available.

April 27, 1973
Friday
4:30-5:30 p.m. RESEARCH REPORTING SECTION V

Presider: KENNETH W. WUNDERLICH, Research and Development Center for Teacher Education, Austin, Texas

Reports: THE DEVELOPMENT OF THE TEST OF QUANTITATIVE JUDGMENT WITH IMPLICATIONS FOR CLASSROOM TEACHING

DONALD E. HALL, Castleton State College
Castleton, Vermont

CYNTHIA T. HALL, Castleton State College
Castleton, Vermont

SEX OF EXAMINER AS A VARIABLE IN PIAGETIAN CONSERVATION ASSESSMENT

CHARLES R. PARISH, Ball State University
Muncie, Indiana

GRAYSON H. WHEATLEY, Purdue University
Lafayette, Indiana

THE YOUNG CHILD'S KNOWLEDGE OF LINEAR PATTERNS

VINCENT J. GLENNON, University of Connecticut
Storrs, Connecticut

ROBERT G. CROMIE, St. Lawrence University
Canton, New York

THE EFFECTIVENESS OF POSITIVE AND NEGATIVE INSTANCES ON THE ATTAINMENT OF THE GEOMETRIC CONCEPT "SIMILARITY" BY SIXTH GRADE STUDENTS AT TWO INTELLIGENCE LEVELS

RICHARD A. HOUDE, Weston High School
Weston, Massachusetts

THE DEVELOPMENT OF THE TEST OF QUANTITATIVE JUDGMENT
WITH IMPLICATIONS FOR CLASSROOM TEACHING

Donald E. Hall
Castleton State College
On leave from the University of Massachusetts

Cynthia T. Hall
Castleton State College

1. Purpose

This study was concerned in measuring the ability of intermediate grade children on quantitative judgment and the determination of the relationships between this ability and sex, intelligence, mental age, chronological age, and grade level.

Specifically, the problem was to test these major null hypotheses:

(a) Children's ability to deal with the aspects of quantitative judgment tested will not differ significantly between sexes, ages, or grades.

(b) There is no relationship between children's ability to deal with aspects of quantitative judgment and their intelligence quotient or mental age.

2. Rationale

As far back as 1953, the Mid-Century Committee on Outcomes in Elementary Education cited in its "Elementary School Objectives" the importance of including as one of these newer mathematical worlds, the study of quantitative relationships (Kauffman, 1966). Likewise, current research evaluators stress the vital need for mathematical evaluation programs which include many concomitant aspects of mathematical reasoning. Among these are "growth in ability to make judgments in quantitative situations" (Glennon and Callahan, 1968).

If a mathematics program is to have meaning in the life of its students it must include quantitative thinking, for this type of mathematical understanding not only opens up a new mathematical world to the student but also suggests that the words of the psychologists who stress learning with understanding are being heeded. In these newer mathematical worlds large formal testing resources are not available. It is to this problem of devising a test in the area of quantitative judgment that these researchers have been involved. Quantitative judgment is defined as "the individual's ability to apply number and mathematical concepts and processes to quantitative situations encountered both within and outside of the classroom environment. Quantitative judgment includes thinking about amounts, estimating or guessing intuitively relative to <u>how much</u>, <u>how many</u>, <u>how far</u> and/or <u>how large</u>" (Hall, 1965).

This research has been primarily concerned with developing a test of quantitative judgment, refining and validating that test, and in the process learning more about quantitative judgment. This paper is a culmination of two doctoral dissertations. The original work on quantitative judgment was carried out under J. Fred Weaver when he was at Boston University. The test is now being researched further by the Office of Special Tests, Educational Testing Service, Princeton, New Jersey. It will be made available for classroom use.

3. Research Design and Procedure

(a) A pilot study was conducted for the Test of Quantitative Judgment (Form H). In this fifty-two items were administered to 161 pupils. Then on the basis of item analysis twelve items were eliminated leaving a total of forty items.

(b) The Test of Quantitative Judgment (Form H) was administered to 704 intermediate-grade pupils. The results of this testing were then subjected to analysis of data.

(c) Using specified criteria, thirty items from the original Test of Quantitative Judgment (Form H) were chosen to be included in the second Test of Quantitative Judgment (Form T).

(d) Sixty new or revised items were prepared and examined to constitute the pilot test for Form T.

(e) The pilot test for Form T was administered to 151 pupils and the results examined by item analysis.

(f) The Test of Quantitative Judgment (Form T) was prepared (thirty items from Form H and thirty items from the second pilot test) and administered to 637 intermediate-grade pupils. The raw test scores and related data were then analyzed.

4. Findings

(a) The raw score means were computed for both boys and girls at each grade level to test the null hypothesis that the means for boys and girls do not differ. A t-test was applied which rejected the hypothesis at the fifth and sixth grade level and accepted the alternate hypothesis at the fourth grade level.

(b) The raw score means were computed for grades to test the null hypothesis that the means do not differ by grades. Using the F-test the hypothesis of equal means was rejected at both the five and one per cent levels.

(c) The Kuder-Richardson formula 20 was selected to estimate the reliability of the Test of Quantitative Judgment. A reliability of .78 was computed.

(d) The refinement of the Test of Quantitative Judgment (Form T) is the most important aspect of its development to date. The selecting of new items or revising of old ones is a vital part of test construction, for this process increases both the reliability and validity of the test (Anastasi, 1954).

The two very important phases of item analysis are the determining of the level of difficulty for each item and its discriminating power (the ability of the item to differentiate between high and low scorers). Both of these indices were established for each item using Chung-Teh Fan's Item Analysis Table (1952).
The results were as follows.

ITEM ANALYSIS SUMMARY BY DIFFICULTY LEVEL
AND DISCRIMINATING POWER

Category	Number of Items
Very Easy	3
Very Difficult	0
Strong Discriminator	23
Medium Discriminator	20
Weak Discriminator	17

(e) The Spearman-Brown Formula was used to test the effect of lengthening the Test of Quantitative Judgment (Form T) on its reliability. If the test were lengthened from sixty to ninety items a reliability of .84 would be obtained; if the test were lengthened from sixty to 120 items, a reliability of .87 would be obtained.

(f) When test scores are correlated with an established test, the resultant correlation is considered a measure of congruent validity (Davis, 1964). The test of Quantitative Judgment was correlated with Sueltz's Test of Arithmetical Understandings. The correlations obtained were .57, .54, and .61 for grades four, five, and six. This indicated that the Test of Quantitative Judgment was not measuring the same thing as arithmetical understandings.

5. Interpretations

The Test of Quantitative Judgment (Form T) is a unique test. The questions require responses about things in the ordinary world of pupils. In making these responses the pupils will have to use quantitative judgments; the questions are designed to be solved by making estimations about quantities rather than by mathematical computation. This ability is developmental in nature as the means of pupils' scores increase with grade level. It is hoped that by bringing this test to the attention of teachers, they will give it priority in their curricula and thus increase the exposure of pupils to the creative, exciting, meaningful world of mathematics beyond computational arithmetic.

REFERENCES

Anastasi, Anne. Psychological Testing (New York: The Macmillan Co., 1954), p. 152.

Davis, Frederick B. Educational Measurements and Their Interpretation (Belmont, Cal.: Wadsworth Publishing Co., 1964), p. 81.

Fan, Chung-Teh. Item Analysis Table (Princeton, N. J.: Educational Testing Service, 1952).

Glennon, Vincent J. and Callahan, Leroy F. Elementary School Mathematics, (Washington, D. C.: Association for Supervision and Curriculum Development, NEA, 1968) pp. 83-84.

Hall, Donald E. "T_ Ability of Intermediate Grade Children to Deal with Aspects of Quantitative Judgment" (Unpublished doctoral dissertation, School of Education, Boston University, 1965), pp. 4-5.

Kauffman, Joseph F. Education (Washington, D. C.: Potomac Books, 1966), p. 23.

SEX OF EXAMINER AS A VARIABLE IN PIAGETIAN CONSERVATION ASSESSMENT

Charles R. Parish
Ball State University

Grayson H. Wheatley
Purdue University

1. Purpose

The purpose of this study was to examine the effects of sex of examiner on conservation of length responses under group testing procedures with second-grade subjects. A group test was developed which consisted of concrete materials. The reliability of the test was established using a sample of 406 second and third grade children. The KR-20 reliability coefficient was .95 and test-retest reliability (on a smaller sample) was .75.

2. Rationale

Available literature indicates that researchers have not studied to a sufficient degree the conditions under which conservation data have been collected. A number of researchers have begun to consider this problem and its implications for interpretation of normative data. In particular, Bittner and Schinedling, Zimiles, and others have identified the experimenter himself as a potentially important variable.

3. Research Design and Procedure

The sample for the study consisted of 51 Ss whose mean age at testing was 8.2 years. Twenty-six Ss were male and 25 were female.

Ss were randomly assigned to one of two test groups which were then randomly assigned to a male or female E.

For testing purposes, Ss were seated in a wedge configuration with E and the apparatus located at the apex of the wedge. This was done to minimize possible perceptual distortions. Ss responded by marking a box or star in response booklets provided by E. The test consisted of seven items: two practice items, two length conservation items, a response breaking item, and two length conservation items. Verbal instructions during the test were alternated systematically both within and between items to control for recency responses by the Ss.

4. Findings

Analysis of the data was performed by use of contingency table frequency analysis with restrictions suggested by McNemar. Results of the analysis revealed that at the .05 α-level there were significantly more operational conservers of length under female administration than under male administration. Under a given E there were no significant differences in the number of operational conservers at the .05 α-level between male and female Ss.

5. Interpretations

It is suggested that the relevance of variables such as sex of E may help explain to a large extent the level of conservation attainment and age of attainment of threshold and operational conservation. It seems clear that if normative data are to be interpretable, basic conditions of data collection must be carefully examined.

THE YOUNG CHILD"S KNOWLEDGE OF LINEAR PATTERNS

Vincent J. Glennon
University of Connecticut

Robert G. Cromie
St. Lawrence University

1. Purpose

(a) To examine the development of the ability of young, normal children to deal with patterns in three cognitive processes and in three cognitive modes.

(b) To use the result of this inquiry into the ontogeny of linear patterns to explore some of the educational implications in the use of patterning activities of the types involved.

2. Rationale

Sigel (1964) related the educational implications of theorizing such as that of Piaget by stating that curricula should be organized to take the stages of concept attainment into account.

Glennon and Alderman (1970) spoke of the concern among educators and cognitive psychologists for finding learning experiences which "ensure a better fit between the capacity of the learner and the difficulty of the cognitive task." Other authors (Ausubel, 1964; Goodlad, 1966) cited the lack of such a "fit" between school and pupil and were disturbed regarding the cummulative effects of a curriculum that is too demanding of the child.

Patterning activities can provide intrinsically interesting, non-reading oriented experiences on a manipulative, action-centered level. This study provides information on patterning tasks. Thus data from this study can be the basis for scientifically verified, ontogenetically-sequenced curricular materials which can be matched to the cognitive level of learners similar to those in the study.

It was assumed that the subjects must not be considered entirely naive with respect to patterning activities, but that such contamination is minimal and to be expected in designing curricula.

3. Research Design and Procedure

Inquiry into the ontogeny of linear patterns was accomplished through the administering of a fifty-four item instrument structured to reflect the three cognitive processes and the three cognitive modes. The processes were categorized as: reproduction (make a copy), identification (select a response from a set of possible responses), and extension (extrapolate in a specific manner from a model). The modes were classified as: enactive (cubical blocks), iconic (colored square regions), and symbolic (block-letter symbols). The three processes were crossed with the three modes to yield a three-by-three matrix. Six basic linear patterns with fixed repetends were applied within each of the nine cells of the matrix. The resulting fifty-four item instrument was administered to the subjects during a nine day period by four testers pretrained in administration of the tasks.

The subjects were fifty-six children (residents in an inner city urban development project) divided into three groups as follows:

Grade Level	N	Mean Age Months	Mean Age Years
Prekindergarten	15	57.3	4.8
Kindergarten	20	67.9	5.7
First Grade	21	79.3	6.6

Subjects at each grade level were chosen at random from those designated by teachers and administrators of their nursery or elementary school as falling into the top quartile of a ranked distribution (ranked as to their ability to interact with a standard curriculum).

The instrument was administered on an individual basis. Instructions were verbal. To acquaint each subject with the materials and the mechanics of the responses a training task occurred at the first possible presentation of the relevant materials.

Analysis of variance statistics were computed with respect to the differences in performance attributable to cognitive process, cognitive mode, and grade level. Planned orthogonal comparisons through t-tests provided evidence of the effects on performance attributable to the cells and patterns. Item analyses were conducted with respect to comparisons between item scores and cell scores, and between item scores and total scores. These analyses included difficulty levels, discrimination indices, and point biserial estimates of the Pearson product moment correlation coefficients.

4. Findings

The following statistically significant ($p < .01$) differences in performance resulted:

(a) Main effects. There were differences in performance attributable to cognitive process and to grade level.

(b) Simple main effects. There were differences in performance at each grade level attributable to cognitive process. There were differences in performance in the identification and extension processes attributable to grade level.

These findings and other observations led to the following conclusions:

(a) The ontogeny for cognitive process was established as reproduction, identification, extension.

(b) The types of reasoning involved to accomplish these tasks improve during the ages of 4-7 years.

(c) A grouping effect by cells reflected the process effect.

(d) Cells and individual tasks differentiated between those subjects who knew how to respond to the tasks and those who did not.

(e) Tasks in the extension process lend themselves to the design of classroom activities and inventory items.

(f) Prekindergarten and kindergarten groups found patterns with repetends of four elements the most difficult.

(g) All subjects knew how to respond in the reproduction process.

(h) For the overall group the order of presentation was established for patterns P1, P2, and P3, in that order.

(i) Subjects responded easily to the tasks.

(j) Subjects responded equally well in all modes.

5. Interpretations

The results of this study can be used to ensure a closer fit between the capacity of the learner and the nature of the cognitive task involved. The findings indicate that activities in reproducing patterns would be worthwhile as introductory experiences. However, there should be only a limited number of these activities. The identification and extension processes offer more challenge without appearing too discouraging.

Statistically, the process effect is significant at each mode. But, the only statistically significant effect for mode appeared in the simple main effect for mode at the extension process. Thus, it seems that if activities are planned in extending patterns using materials and subjects similar to those employed in this study, they should be ordered with respect to mode: symbolic, iconic, and enactive. This order is the reverse of that which most mathematics educators and cognitive psychologists would expect. More research should be conducted to investigate this area.

THE EFFECTIVENESS OF POSITIVE AND NEGATIVE INSTANCES ON THE ATTAINMENT OF THE GEOMETRIC CONCEPT "SIMILARITY" BY SIXTH GRADE STUDENTS AT TWO INTELLIGENCE LEVELS

Richard A. Houde
The University of Tennessee

1. Purpose

Mathematics teachers present concepts to their students using both positive and negative instances of the concepts. The question of what type of instances (positive or negative) produce best learning in students of different abilities is an important one and a difficult one to answer. The major aim of this study was to examine the interaction between intelligence and the use of positive and negative instances on the attainment of the geometric concept "similarity" at the sixth grade level.

2. Rationale

The experimental concept attainment literature has revealed few studies dealing with the effects of positive and negative instances on the attainment of relational concepts. Furthermore, of all the studies employing non-relational concepts, most have dealt with concepts defined over finite universal classes. Since many mathematical concepts are relational and are defined over infinite universal classes ("less than," "equal to," "congruence," etc.), the reported results may not be applicable to the learning of mathematical concepts. This study is an attempt to supply experimental evidence which applies directly to the learning of mathematical concepts by examining the effects of positive and negative instance presentation on the attainment of a relational mathematical concept over an infinite universal class.

A review of the literature pertaining to the geometric concept "similarity" has revealed that experimenters used primarily only two geometric figure types, triangles and rectangles, as a basis for judging concept acquisition or concept attainment. Since the instances of the concept "similarity" involve all geometric figure types, it seems reasonable to extend the set of figure types used, hence this study subjected experimental subjects to a much wider range of geometric figure types.

Dodwell (1971) noted that the Piagetian revolution of the past decade had produced little research concerning children's visual perception and their understanding of geometrical concepts. He further remarked that this was all the more surprising when one considered that children's understanding of geometry and other spatial concepts is thought to be related to both their perceptual and intellectual development. Hence, this study may be considered valuable due to the fact that it employs children's visual perception of geometric figures to examine the interaction between intelligence and the use of positive and negative instances in learning a geometrical concept.

Since this experimental study was limited to the investigation of a single concept involving visually perceived shape, a revised version of the definition of "concept" proposed by Shumway (1971) was employed. The revised definition is: a concept is a partition of all stimuli from a class X into two subclasses X_1 and X_2 where the partition is determined by certain predetermined characteristic properties of the stimuli. All stimuli of class X which possess the predetermined characteristic properties are assigned to class X_1 and all other stimuli from class X are assigned to class X_2.

Stimuli of the class X_1 are called positive instances of the concept and stimuli of the class X_2 are called negative instances of the concept. Class X is called the universal class over which the concept is defined. Also, following the format established by Shumway (1971), a subject is said to have attained a concept over a class X if given any stimulus from X, he is able to classify the stimulus on either a positive or negative instance of the concept.

Most geometry textbook authors such as Moise and Downs (1967), Coxford and Usiskin (1971), and Levy (1970) do not define the geometric <u>concept</u> "similarity." Rather, they propose definitions of functions or relationships that describe the necessary conditions for two geometric figures in the plane to be similar or not similar to each other. Since this experiment was concerned with the geometric <u>concept</u> "similarity," those textbook definitions were modified to satisfy the purposes of this experiment.

The following definition of the geometric concept "similarity" was prepared for this study:

Let S be the class of all ordered pairs of geometric figures in the plane. The geometric concept "similarity" is a partition of S containing two subclasses S_1 and S_2 such that if (a,b) is in S_1, then there is a "similitude" (Levi, 1968) between a and b and if (a,b) is in S_2, then there is no similitude between a and b.

In particular, if (a,b) is in S_1, then (a,b) will be called a positive instance of the concept "similarity" and we will say that "a is similar to b." If (a,b) is in S_2, then (a,b) will be called a negative instance of the concept "similarity" and we will say that "a is not similar to b."

3. Research Design and Procedure

Three types of instructional booklets, each containing a series of 32 instances of similarity, were developed. One type contained a series of all positive instances, another type contained a series of all negative instances, and third type contained an alternating series of positive and negative instances. Each instance consisted of two geometric figures drawn in ink on 8 1/2 x 11 inch white paper using straightedge and compass. The figures were drawn having approximately the same orientation. One sentence accompanied each instance of the concept and was placed above the drawings of the geometric figures. If the instance was positive, the sentence was, "These two figures are similar." If the instance was negative, the sentence was, "These two figures are not similar."

A 40 item multiple-choice test, designated as the Similarity Criterion Test (SCT), was constructed (see example). Each test item was drawn on 8 1/2 x 11 inch plain white paper. A line divided each piece of paper into two parts. Above the line, one figure (called a reference figure) was drawn and beside it was stated the question, "Which figure below the line is SIMILAR to this figure?" Below the line, five figures were drawn among which only one served as a correct answer to the question. In all but three questions, each of the figures below the line was the same type of geometric figure as the reference figure. Furthermore, all the figures in the test items were drawn having nearly the same orientation. With few exceptions, the types of geometric figures used in construction of the test items had not been used to form any instance in the instructional series booklets. The figures were drawn in ink using compass and straightedge.

A variety of geometric point sets were selected for use in construction of the instances of similarity and the test items. A partition of geometric figures into eight categories was devised to aid in the selection procedure.

The sixth grade children at Halls Middle School in Knox County, Tennessee were divided into high, average, and low IQ groups using the scores from an Otis-Lennon Mental Ability Test. Sixteen students were randomly assigned to each of four instructional treatment conditions (all-positive series, all-negative series, alternating series, and control) by selecting eight students from the high and eight students from the low intelligence groups. In this manner, eight students were assigned to each of eight treatment combinations constituting a 2 x 4 factorial design.

4. Findings

Of the 64 subjects originally assigned to the eight experimental treatment groups, three subjects failed to participate in the experiment. This resulted in unequal numbers of SCT scores for the eight cells of the 2 x 4 factorial design. Thus, the analysis of variance techniques used to analyze the factorial design are based upon statistical methods advanced by Rao (1952) and Anderson (1958) and described by Winer (1962). If a null hypothesis was rejected ($p < .05$), then Duncan's New Multiple Range Test (Winer, 1962) for unequal cell frequencies was used to examine the differences between the means.

Analysis of data revealed that over all levels of instructional series, the high IQ mean score was significantly greater than the low IQ mean score ($p < .001$) and that over both IQ levels, the four instructional series mean scores differed significantly ($p < .01$). In particular, multiple comparisons of the four instructional series mean scores revealed that the alternating series and the all-positive series mean scores were each significantly greater than both the control and the all-negative series mean scores ($p < .05$).

Also, it was found that the IQ-Instructional Series interaction effect was significant ($p < .05$). Comparisons of the high IQ and low IQ mean scores within each instructional series group revealed that the high IQ mean score was greater than the low IQ mean score within each of the all-negative series, alternating series, and control groups. In particular, within both the all-negative and alternating series groups, the high IQ mean score was significantly greater than the low IQ mean score ($p < .05$). However, the low IQ mean score was greater than the high IQ mean score within the all-positive series group. This difference was not significant ($p > .05$).

5. <u>Interpretations</u>

Hence, with respect to the children and the materials used in this study, it was concluded that:

(a) The geometric concept "similarity" is attained more efficiently with all positive instances or with half positive and half negative instances than with all negative instances or with no instances.

(b) Grouped over all levels of instructional series conditions, children of high intelligence perform significantly better than children of low intelligence.

(c) Using all positive instances, perhaps children of low intelligence can perform as well as children of high intelligence.

(d) All negative instances tend to confuse children of low intelligence.

(e) Children of high intelligence use either positive or negative instances efficiently. However, the high IQ group using both positive and negative instances exhibited the best performance of the eight treatment groups.

REFERENCES

Anderson, T. W. An Introduction to Multivariate Statistical Analysis. New York: John Wiley & Sons, Inc., 1958.

Coxford, Arthur F. and Zalman P. Usiskin. Geometry A Transformation Approach. River Forest, Illinois: Laidlaw Brothers, Publishers, 1971.

Dodwell, Peter C. "Children's Perception and Their Understanding of Geometrical Ideas," Piagetian Cognitive-Development Research and Mathematical Education, eds. Myron F. Rosskopf, Leslie P. Steffe, and Stanley Taback. Washington: National Council of Teachers of Mathematics, Inc., 1971.

Levi, Howard. Topics in Geometry, ed. Howard W. Eves. The Prindle, Weber & Schmidt Complementary Series in Mathematics, Vol. XI. Boston: Prindle, Weber & Schmidt, Inc., 1968.

Levy, Lawrence S. Geometry: Modern Mathematics via the Euclidean Plane. Boston: Prindle, Weber & Schmidt, Inc., 1970.

Moise, Edwin E. and Floyd L. Downs Jr. Geometry. Reading, Massachusetts: Addison-Wesley Publishing Company, Inc., 1967.

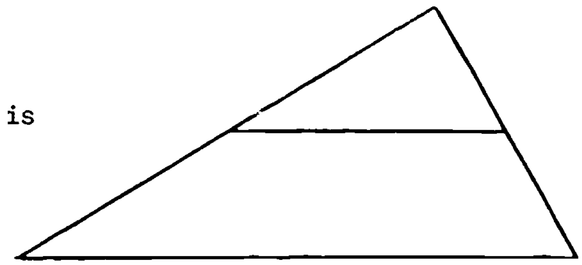
Rao, C. R. Advanced Statistical Methods in Biometric Research. New York: John Wiley & Sons, Inc., 1952.

Shumway, Richard J., "Negative Instances and Mathematical Concept Formation: A Preliminary Study," Journal for Research in Mathematics Education, 2:218-227, 1971.

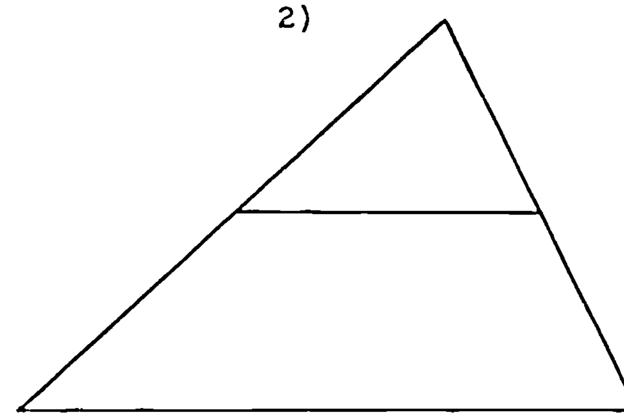
Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Company, Inc., 1962.

12. Which figure below the line is

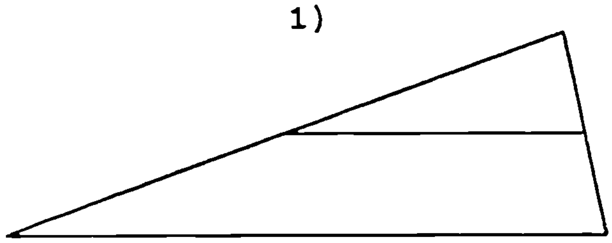
SIMILAR to this figure?



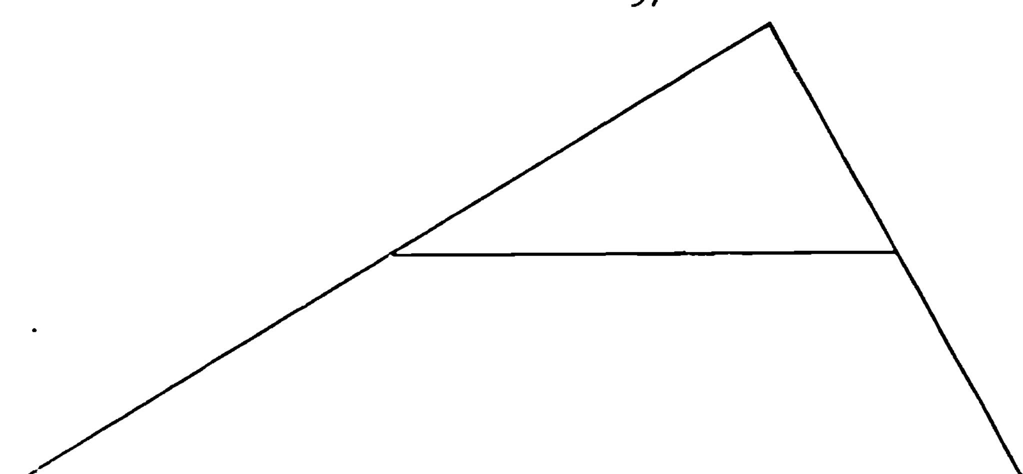
2)



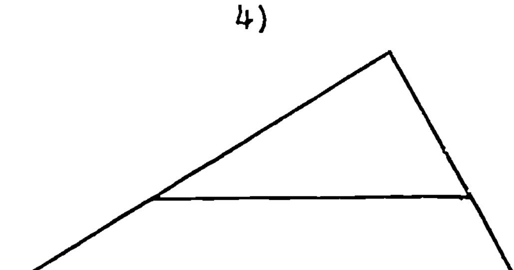
1)



3)



4)



5)