This issue of the journal contains abstracts and critical comments for 10 published reports on research in mathematics education. The reports concern children's school mobility and achievement, preservice teachers' sources of decisions, spatial visualization in boys and girls, rote versus conceptual emphases in teaching probability, using drawings in conditional reasoning programs, international study achievement results for grade 12, shifts in reasoning, sentence-solving strategies, effects of whole class, ability grouped, and individualized instruction, and teaching basic fact strategies for addition and subtraction. Research references from the "Current Index to Journals in Education" (CIJE) and "Resources in Education" (RIE) for October through December 1985 are also listed. (MNS)
INVESTIGATIONS IN MATHEMATICS EDUCATION

Volume 19, Number 2 - Spring 1986

THE ERIC SCIENCE, MATHEMATICS AND ENVIRONMENTAL EDUCATION CLEARINGHOUSE in cooperation with Center for Science and Mathematics Education The Ohio State University


Hansen, Randall S.; McCann, Joan; and Myers, Jerome L. ROTE VERSUS CONCEPTUAL EMPHASES IN TEACHING ELEMENTARY PROBABILITY. Journal for Research in Mathematics Education 16: 364-374; November 1985. Abstracted by GLENDA LAPPAN 18

Markovits, Henry. THE CURIOUS EFFECT OF USING DRAWINGS IN CONDITIONAL REASONING PROBLEMS. Educational Studies in Mathematics 17: 81-87; February 1986. Abstracted by ROBERTA L. DEES 23

McKnight, Curtis C.; Travers, Kenneth J.; and Dossey, John A. TWELFTH-GRADE MATHEMATICS IN U.S. HIGH SCHOOLS: A REPORT FROM THE SECOND INTERNATIONAL MATHEMATICS STUDY. Mathematics Teacher 78: 292-300, 270; April 1985. Abstracted by THOMAS O'SHEA 30


Steinberg, Ruth M. INSTRUCTION ON DERIVED FACTS STRATEGIES IN ADDITION AND SUBTRACTION. Journal for Research in Mathematics Education 16: 337-355; November 1985. Abstracted by ROBERT B. ASHLOCK .................................................. 53

Mathematics Education Research Studies Reported in Journals as Indexed by Current Index to Journals in Education October - December 1985 .................................................. 59

Mathematics Education Research Studies Reported in Resources in Education October - December 1985 .................................................. 63
1. Purpose

The purpose of the paper was to summarize previous findings on the stated research topic and to consider their implications for schools.

2. Rationale

Several previous research studies are cited (ranging from 1933 to 1980). Analysis of the previous and relevant research gives a rather confusing pattern of inconsistent results. Thus, it is difficult to make definitive statements in relation to schooling and academic achievement. A suggested reason for the inconclusiveness of the findings is that previous studies have not taken into account the variable of prior attainment and socioeconomic status. Previous literature had suggested that research on mobility concentrate on three phases of mobility: 1) the situation before the move; 2) the actual moving situation; and 3) the situation after the move.

These factors led Blane and others to recommend a new and major longitudinal study of school mobility and attainment using data contained in the National Child Development Study (NCDS).

3. Research Design and Procedures

The NCDS was a long-term study carried out by the National Children's Bureau in London. It included all children born during the week 3-9 March 1958 and living in England, Scotland, and Wales. The
present study used data from the NCDS study. Data were collected when the children were aged 7, 11, and 16. Information included data from parents, school records, medical exams, exam results, and student questionnaires (at ages 11 and 16). The data used in the study are from the sample of children for whom a full set of variables was available.

The study consisted of two phases. First, school behavior and social and home circumstances of mobile and non-mobile children were compared during the elementary and secondary school years. This was a descriptive stage of the study and was used to justify the use of multivariate analysis in the rest of the study. The purpose of these analyses was to adjust for initial differences. Three sets of data analysis were carried out (elementary years, secondary years, in total schooling). The findings in this report dealt with the content area of mathematics.

4. Findings

These tests of mathematics achievement were administered (at ages 7, 11, and 16). Scores were standardized to give a sum of zero and a standard deviation of one for the multivariate analyses. As mentioned above, the descriptive analyses were inconclusive, although for the most part the attainment scores of the more mobile children were lower than those of the non-mobile children.

At each stage of the study, the father's occupation was recorded. These data were used to categorize according to the Registrar General's 1966 classification of occupations. Study of various groups at the differing age levels was conducted. The results again showed a general trend in that mobile children showed general decline in ability.
Three sets of multivariate analyses were carried out:
1st set - School progress between 7 and 11/School behavior at 11.
2nd set - School progress 11 and 16+/School behavior at 16.
3rd set - School progress 7 and 16+/School behavior at 16.

Changing elementary schools does affect progress in mathematics between ages 7 and 11. One change does not appear to have much effect; however, more than one change in schools does have an adverse effect.

When looking at mathematics attainment at age 16, there is no strong adverse affect of school mobility when the variables of sex, attainment at age 11 and social circumstances, and type of school at 16. There was a significant interaction between secondary school mobility and social class at age 16. Also, elementary school mobility was not found to be related to mathematics attainment at age 16.

When scores from public exams were considered for age 16, school mobility did have a negative effect on achievement after mathematics attainment and social circumstances at age 11, sex, and school types at age 16 were controlled for. There appeared to be some depression of test scores for those who had changed schools during their high school careers.

Case study evidence yielded the interesting result that for children who attended three or more schools, lower socioeconomic families cited family problems as the primary moving reason, while higher socioeconomic families cited the father's occupation in the primary reason for moving.

5. Interpretations

The main conclusion of the study was that school mobility was not a major concern for school people. It was suggested that mathematics
syllabi and examination syllabi be more consistent. It was also suggested that teachers should be aware of the mobility patterns of students so as to deal with problems experienced by children who are mobile. Teachers are encouraged to consider associated factors rather than relying too heavily on school mobility as a cause for differences in school attainment.

Abstractor's Comments

1. The study was an ambitious one as it was longitudinal, done with a large sample, and dealt with an important topic.

2. The writing style left a lot to be desired. The report rambled on and on and jumped from idea to idea, often repeating many points. As a reader, I found it very hard to follow. Important statements were hard to pick out.

3. The point of the study was to look for relationships between school mobility and attainment in mathematics. However, the results and discussion sections made points about mathematics course and exam syllabi. This seemed off the point and not well-discussed in terms of background and rationale.

4. As well as reading this report, interested parties would also be well-advised to obtain the original sources of data in order to more fully clarify the situation.

5. At the start, one had hopes for a definitive statement and conclusion. There was an attempt at one, but it was lessened by the accompanying discussion.

Abstract and comments prepared for I.M.E. by OTTO BASSLER, George Peabody College for Teachers of Vanderbilt University.

1. Purpose

The purpose of this study was to identify and classify the sources that influence teaching decisions made during initial teaching experiences. Three components were used to classify sources of teaching decisions:

a. teacher enculturation--occurs by observing and reflecting on one's own actions.

b. teacher education--occurs in school classrooms and through direct teaching experience.

c. teacher schooling--occurs outside the school classroom and is presented by specially prepared persons.

2. Rationale

Teachers are required to make many decisions. As rational, thinking individuals they make decisions in planning for instruction as well as during instruction. The types of decisions, reasons for decisions, and alternative decisions have been the basis for previous research; however, little research on sources of teaching decisions has been done. Knowledge of the sources of teaching decisions may help to plan more effective teacher education programs and may lead to a better understanding of these programs.
3. Research Design and Procedures

Five preservice teachers served as subjects. All were female, were enrolled in a mathematics methods course in a fall quarter, and were enrolled in student teaching the following winter quarter.

The methods course had two types of field experience: micro-teaching lessons, 20-minute lessons taught to two or three eighth-grade students, occurred at a middle school once a week for the first four weeks of the methods course; and small group lessons, one-hour lessons taught daily to a small group of ninth-grade algebra students for three consecutive weeks near the end of the fall quarter. In the student teaching quarter each subject was assigned to one or two cooperating teachers for a period of ten weeks.

Interview data were gathered before and after all four microteaching lessons, two selected small group lessons, and four selected student teaching lessons. Before each lesson the subjects were asked to identify the sources of teaching decisions in their lesson plans; after each lesson they were asked to identify sources of spontaneous decisions made in the lesson.

Data were tallied and categorized by sources within components for each type of teaching experience. For Teacher Enculturation, sources were past teachers' performance, methods instructors' performance, classmates' performance, cooperating teachers' performance, reflecting on own learning, and reflecting on own teaching. Sources in Teacher Education were microteaching suggestions, small group suggestions, cooperating teachers' suggestions, supervisors' suggestions, and mathematics textbooks. Teacher Schooling sources were methods course content and other college courses. A fourth component, Miscellaneous, was added and included the sources: seemed logical, personal style, and indeterminant. When an independent rater was compared to the investigator on tallied sources from taped interviews, a correlation of .96 (n = 48) was obtained.
4. Findings

The most often cited source for making teaching decisions was the content of the methods course. It ranked second for microteaching lessons, first for small group lessons, second for student teaching lessons and first overall.

The overall second ranked source of decisions was mathematics textbooks. They were ranked first in student teaching lessons and third in both microteaching and small group lessons.

Ranked third overall was the source 'microteaching suggestions,' which was probably inflated due to its first place ranking in microteaching experience. This source was not ranked high for making decisions in small group or student teaching situations.

Past teachers' performance was ranked fifth overall and was ranked second in the small group instruction at the close of student teaching.

Cooperating teachers' performance and suggestions were ranked third and fourth respectively when teaching decisions were made during the student teaching experience. These were slightly higher than the ranking of supervisor's suggestions.

Miscellaneous sources were ranked as rather important in making teaching decisions. Overall ranking of 'seemed logical' was seventh, 'personal style' was eighth, and 'indeterminant' was fourth.

5. Interpretations

The investigator acknowledged several limitations to the study. They were: all data were based on self report, subjects sometimes could not identify the source of an idea or decision, and the investigator served as both methods instructor and investigator. With these limitations clearly stated, some interpretations and implications for further research were offered:
a. "The preservice teachers relied heavily on textbooks as a source of their teaching decisions." Since textbooks have such a strong influence on teaching decisions, more research is needed to help teachers make better decisions about the role, capabilities, and limitations of textbooks.

b. "The preservice teachers in this study received little help and advice from their cooperating teachers." Never-the-less, the cooperating teachers' performance and suggestions were ranked third and fourth as the basis for making teaching decisions during student teaching. Research is needed for better utilization of cooperating teachers as they help to educate preservice teachers.

c. "...the content of a methods course had a significant impact on the teaching decisions of five preservice teachers." There needs to be additional research to determine if these findings generalize to other students, teachers, methods courses, and different conditions.

Abstractor's Comments

The author is to be commended for viewing the findings of this research as suggestions for further research rather than as the basis for drawing conclusions. He did provide additional support to the ranking of sources for decisions made by student teachers—namely, textbook, methods course (taught by the supervising teacher and investigator), and cooperating teacher. It would have been surprising if any other source had been among the top three, for the textbook provides the content and organization as well as teaching suggestions, the methods course provides the teaching suggestions of the supervisor and investigator, and the cooperating teacher is the designated authority in the student teaching experience. Needless to say, the supervising and cooperating teachers are evaluators of the student teaching experience.
Little indication of the items on the interview guide were included in the paper; hence, the reader is left to wonder what aspects of the lesson were singled out to be questioned. Were these content decisions? methodological decisions? classroom management decisions? Did they deal with decisions that were pedagogically sound? emphasized different strategies? There are many questions which consider the soundness of the decision that also would be most important to investigate.

The author did indicate that 14% of all responses were viewed as incorrect or indeterminate as one of the limitations of the study. This category of decisions is intriguing and certainly needs more investigation. Do student teachers make these decisions using a source they have experienced but cannot identify or because they have no source upon which to make the decision? If these decisions are critical to the instructional program and teacher educators are not providing a basis for making these decisions, then it is imperative that the student education program be modified.

This study raised interesting questions that need to be studied. As such, it accomplished one of its goals which "was to provide an impetus to investigate more closely sources of teachers' decisions and their applications to teacher education."

Abstract and comments prepared for I.M.E. by FRANCES R. CURCIO, Queens College of the City University of New York, Flushing.

1. Purpose

The purpose of this longitudinal study was to examine the effect of spatial visualization and verbal skills on girls' and boys' mathematical problem-solving ability.

2. Rationale

The related literature presented suggests a relationship among spatial visualization, verbal skills, the learning of mathematics, and sex-related differences in mathematics performance. Some of the studies reviewed included mathematics tasks that were overtly spatial, influencing a direct relationship between spatial visualization and mathematical problem-solving ability. As a result, the relationship between spatial visualization and the "broader spectrum of mathematics tasks remains unclear" (p. 184).

Studies that examined the relationship between sex-related differences and spatial visualization skills on students' ability to solve problems were cited. Conflicting results in the ability to predict problem-solving performance based upon sex and spatial visualization were pointed out.

Since verbal skills have been documented to be related to mathematics performance, and sex-related differences with respect to verbal skills tend to occur favoring girls, it was questioned why girls don't also excel in mathematics (p. 185). Furthermore, it was pointed out that either a spatial or verbal approach can be used to solve many mathematical problems.
As a result of a review of the literature, the inclusion of spatial visualization, verbal skills, sex, and mathematical problem-solving ability as variables for this study was supported.

3. **Research Design and Procedures**

After administering the Space Relations subtest of the Differential Aptitude Test (to measure spatial visualization), the vocabulary test of the Cognitive Abilities Test (to measure verbal skills), and the SRA Mathematics Concepts test (to measure mathematics achievement), to 669 sixth graders in four out of the five middle schools in Madison, Wisconsin, 36 girls and 33 boys discrepant in spatial visualization (SV) and verbal skills (V) (i.e., students scoring in the upper third on SV and in the lower third on V—high/low, or, in the lower third on SV and the upper third on V—low/high) were assigned to one of four groups. The groups consisted of 18 high/low girls, 18 low/high girls, 17 high/low boys, and 16 low/high boys.

The students were presented with word problems and fraction problems on the **Instrument to Measure Mathematical Thinking (IMMT)** during individual interview sessions conducted by trained interviewers in the spring of the sixth, seventh, and eighth grades. Over the course of the three years, some problems were modified and others were deleted or inserted, based upon students' maturing knowledge and abilities. In the eighth grade, the subjects of this study and their peers (n = 103), selected randomly, were given the same SV and V tests as well as the Mathematics Basic Concepts subtest of the Sequential Test of Educational Progress.

The subjects experienced three phases for each problem in the interview: the **verbalization** phase (consisting of asking the student to read silently and state the problem in his/her own words); the **solution**; and the **explanation** phase (consisting of asking the student to describe how the picture was used to solve the problem).
The interviews were audiorecorded. The protocols transcribed from the tapes, students' pictures, and interview records were coded each year by two trained advanced doctoral or postdoctoral students in mathematics education.

The interview data were analyzed according to six dimensions: correct solutions, verbal information, translation picture information, mental movement, solution picture information, and use of picture.

Descriptive statistics were reported for each of the four groups by grade. The results of the standardized tests, and the IMMT interview data coded and analyzed by dimensions, were presented. Spatial/verbal group and sex provided the sources of variance for analyses of variance.

4. Findings

Five research questions were presented and discussed separately.

"Question 1. Do girls and boys with discrepant spatial visualization and verbal skills differ on the ability to solve mathematical problems?" (p. 191)

The differences between the spatial/verbal groups with respect to the number of correct solutions were not statistically significant. That is, "high/low students did not differ from low/high students in their abilities to accurately solve mathematical problems" (p. 191).

In general, the boys' groups outperformed the girls' groups in their ability to solve mathematical problems. The differences between these groups were found to be statistically significant (p < .05).
"Question 2. Do girls and boys with discrepant spatial visualization and verbal skills differ on the ability to verbalize relevant data in mathematical problems?" (p. 193)

Results from the verbal information dimension indicated that girls provided more complete verbal information than boys, and the differences between the groups were significant (p < .05) in Grade 6 (all solutions to fraction problems, and, correct solutions to fraction problems), and Grade 7 (correct solutions to all problems). "Six significant spatial/verbal group differences were found; in every case, the groups low in spatial visualization ability and high verbal ability provided more verbal information" (p. 193).

"Question 3. Do girls and boys with discrepant spatial visualization and verbal skills differ on the ability to translate symbols into pictorial representation?" (p. 193)

Results from the translation picture dimension yielded five significant differences between spatial/verbal groups in favor of the high/low groups in Combined Grades (for total correct solutions, correct fraction problems solutions, and incorrect word problem solutions), and Grade 8 (for total incorrect solutions and incorrect word problems solutions). The one significant sex difference was due to the higher mean for the sixth grade girls' incorrect solutions to fraction problems.

"The high/low students tended to translate symbols into pictures more completely than low/high students" and "the high/low groups usually did a more complete job in translation when they were able to find the correct solutions" (p. 193).

Based upon the graphs of the means of the translation picture scores for boys and girls during the three years, the researchers observed that
The low/high girls consistently put less information in their translation pictures (for both word problems and fraction problems) than any other group. The high/low girls put more information in their translation pictures for word problems than any other group in both Grades 6 and 8. When one looks at only the correct solutions to the fraction problems...a different trend is evident. The high/low girls put the most information in their translation pictures when they correctly solved fraction problems in Grade 6, but by Grade 8, the high/low boys put the most information in their pictures. Once again, the low/high girls put less information in their pictures than any other group. (p. 197)

"Question 4. Do girls and boys with discrepant spatial visualization and verbal skills differ on the ability to use spatial visualization skills overtly during mathematical problem solving?" (p. 199).

Results from the mental movement dimension were used to answer this question. One significant sex difference was found in favor of the boys in Grade 8 (all solutions of word problems), and two significant spatial/verbal group differences were found in favor of the high/low group also in Grade 8 (all solutions for total problems and fraction problems) (p < .05).

Students were not questioned specifically about their use of mental movement because a direct question "might have influenced their response." Also, it is not evident from the data whether any students "used much overt mental movement, with the exception of the fraction problem in Grade 8" (p. 199). The researchers indicate that these results should be interpreted with caution.

"Question 5. Do girls and boys with discrepant spatial visualization and verbal skills differ on the ability to use pictorial representations during mathematical problem solving?" (p. 200)
The results of two dimensions, solution picture and use of picture, were used to answer this question. "Significant sex-related differences were found in the solution pictures used in the fraction problems for all solutions and correct solutions, with boys having pictures with more information" (p. 200). Significant differences favoring the high/low groups occurred for correct total solutions, and a significant difference favoring the low/high group occurred for incorrect total solutions.

Although the high/low groups tended to score higher on use of pictures, there were "few spatial/verbal differences found" (p. 201). Also, it was reported that girls used pictures more frequently than boys.

5. Interpretations

Although many mathematics educators may believe that spatial visualization skills are highly important when learning mathematics and that perhaps developing these skills should become a major goal of mathematics education, the results of this study suggest that further research and data "are needed before one can safely conclude that an emphasis on spatial visualization skills will improve mathematics learning" (p. 203). Students who demonstrate a strength in either spatial visualization or verbal skills might process problems differently, but neither exclusively influences their ability to obtain correct answers.

The data indicated that low/high boys solved more problems than any other group. These boys also had the highest mathematics achievement scores in Grades 6 and 8. It seems as though their mathematics achievement scores and their success on the IMMT were not inhibited by low spatial visualization skills. On the other hand, the mathematics achievement scores of the low/high girls were the lowest among the sample. "Although one cannot conclude that low
spatial visualization skill caused the low/high girls' lower mathematics achievement, the hypothesis that females are more debilitated than males by low spatial visualization should be investigated" (p. 204).

Abstractor's Comments

This research is a contribution to the literature on the relationship among spatial visualization skills, verbal skills, and sex-related differences in mathematics. As a result of this study, the authors have provided us with a rich base of hypotheses recommended for further testing.

Longitudinal studies such as this one are needed to chart the performance of boys and girls during a period when sex-related differences tend to manifest themselves (i.e., during pre- and early adolescence) (Fennema, 1974; Hilton & Berglund, 1974). What was interesting in this study is that although boys tended to outperform girls by getting more correct solutions, drawing more accurate fraction pictures, and using more pictorial information when solving problems, and girls tended to verbalize more information about the problems, the intrasex differences were larger than the intersex differences. It seems as though there are factors other than sex alone that might cause differences that appear to be sex-related. As suggested by the authors, further study might focus on sex-by-spatial/verbal group interaction as well as examining the characteristics of high/low and low/high girls and boys.

Some questions come to mind as one reads a description of the design and procedures: How many problems of each type (i.e., word and fraction) were given to the students during the interview sessions using the IMMT? Were the problems presented in the same order, a mixed order, or a random order? If the problems were presented in a fixed order, was a fatigue factor taken into
consideration (i.e., would children do better on problems presented in the beginning of the interview because they were more alert, and worse on problems at the end of the interview because they were tired)? Also, how long did the interviews last—was there a time limit put on the children? Were all questions asked during one session or were there several interview sessions per student? All or some of these factors might have an effect on the results.

This study has opened a new door for examining the relationship between spatial visualization and mathematical problem solving by using problems that were not overtly spatial. Although word problems and fraction problems were used on the IMT, perhaps other types of problems could be used to further test this relationship (e.g., probability, statistics, logic) to include broader mathematical tasks. The IMMT can be thought of as an experimental prototype. It can be used as a foundation for developing more sophisticated procedures to examine the use of verbal and spatial skills in mathematical problem solving.

References


Hansen, Randall S.; McCann, Joan; and Myers, Jerome L. ROTE VERSUS CONCEPTUAL EMPHASIS IN TEACHING ELEMENTARY PROBABILITY. Journal for Research in Mathematics Education 16: 364-374; November 1985.

Abstract and comments prepared for I.M.E. by GLENDAL LAPPAN, Michigan State University.

1. Purpose

The purpose of the study was to compare the performance of students who read one of the three texts that varied the emphasis placed on rote versus conceptual understanding of six basic concepts of elementary probability.

2. Rationale

The authors hypothesize that problem solving involves three stages: categorization of the problem, retrieval of the appropriate formula, and translation—correct substitution of values from the problem into the retrieved formula. In an earlier study the authors found that rote instruction resulted in better performance on formula problems while conceptual instruction yielded better performance on story problems. These results fit the hypothesized multistage model in that the students who read a conceptual text were more likely to solve problems when they retrieved the appropriate formula.

In this study the authors return to the earlier study to perform a more molecular analysis on the story problem data collected in order to better understand the errors made by students from the different treatment groups.
3. Research Design and Procedures

The subjects were 48 volunteers from an introductory psychology course who had no previous exposure to instruction in probability or statistics. The subjects were randomly assigned to one of three text conditions—explanatory, standard, and low explanatory.

The standard and low-explanatory texts were 4 pages long. The subjects in these groups were given 15 minutes to read the material. The explanatory text was 14 pages in length; these subjects were given 25 minutes to read the material. The same six formulas were developed in each text. However, the explanatory text used a relative frequency definition of probability, used pictorial aids, and developed the formulas in terms of their relation to other formulas. In the standard text probability was treated as a measure. No pictorial aids or explanations of relations between formulas were used. The low-explanatory text was similar, but even more abstract.

To prevent rehearsal of formulas, all subjects were given a short task not involving probability immediately after the study period. Then the subjects were given a form of the test involving 12 story problems. Forty-eight hours later they returned to take a second form with an additional 12 story problems.

These 24 story problems included four problems for each of the six formulas studied.

4. Findings

The standard and low-explanatory groups performed similarly and are presented together as non-explanatory.
The students in the explanatory condition scored higher than the non-explanatory group on all but six problems; four of these differences were statistically significant.

The error analysis focused on difficulties in classification and translation of the problems. Classification errors were of two types: classification error due to absence of key words and classification error due to presence of irrelevant information.

In the data on key words the explanatory group made fewer errors in problems of the types \( P(A \text{ or } B \text{ or } C), A, B, C \text{ mutually exclusive,} \) and \( P(A \text{ and } B \text{ and } C), A, B, C \text{ independent,} \) when the words "or" or "and" did not appear explicitly in the problems.

On problems which contained irrelevant information, the non-explanatory group tended to use all the information provided. In the explanatory group this error was virtually absent.

Of the errors involving translation, some seemed to be based on a lack of understanding of the relation between frequency and probability. For example, in problem 1A,

"A marble is drawn from a jar containing 10 red, 30 white, 20 blue, and 15 orange marbles. What is the probability of drawing a red or white marble?"

the errors in the non-explanatory group were often because of a failure to divide the sum of the red and white marbles by the total number of marbles or because they divided by 100. This was the most difficult formula 1 problem for the non-explanatory group (.41 proportion correct) and the easiest for the explanatory group (.94 proportion correct).

Other translation errors involved problems where a probability had to be translated into its complement (i.e., subtracted from 1). In these problems the explanatory group made fewer errors.
5. Interpretations

The authors conclude that variations in test performance on problems requiring the same formula for solution were due to various elements in problem presentation that affect the ease of classification and translation. The students in the non-explanatory group tended to have more difficulty than the explanatory group with classification when key words were absent or irrelevant information was present and with translation errors when complements of given probabilities were needed or when verbal statements had to be translated into probabilities.

Abstractor's Comments

This abstractor would classify a study of this sort as a laboratory experiment which raises some interesting questions, but which gives little real information on appropriate instructional models. Studying three types of probability situations, additive, multiplicative, and conditional, in 15 to 25 minutes is unlikely to ever occur in a real teaching situation. The difference in instructional time also raises questions about the results. The explanatory group had two thirds again as much instructional time (25 minutes compared to 15 minutes) as the non-explanatory group.

In spite of these concerns, the abstractor agrees with the authors that the error analysis suggests that an explanation of basic concepts seems to enhance the development of translation skills. This points the way to the development of appropriate teaching models for elementary probability.
The set of 24 test questions devised by the authors is also of interest. One suggestion is that for future research the questions be revised so that over each of the six formulas the questions are varied in specific ways to reflect the classification and translation errors uncovered in this study.

Abstract and comments prepared for I.M.E. by ROBERTA L. DEES, Department of Mathematical Sciences, Purdue University Calumet.

1. Purpose

The author has tested the hypothesis that the use of line drawings, as concrete referents, would facilitate the conditional reasoning performance of college students.

2. Rationale

The author states that conditional reasoning is an important manifestation of formal thought; he also cites research results that indicate its difficulty even for adults (Karplus and Karplus, 1970; O'Brien, 1973; also several references from developmental psychology literature). The hypothesis that referential drawings would be helpful seems intuitively reasonable, and in fact, the author's informal survey of a small sample of educators revealed unanimous agreement that such concrete support should facilitate performance on conditional reasoning questions. However, the author says, "some preliminary results appeared to indicate the contrary" (p. 81). Thus he decided to investigate the question in a systematic study.

3. Research Design and Procedure

Subjects were 80 first-year college students, average age 18 years, 5 months. The instruments were two paper-and-pencil questionnaires, each containing three conditional reasoning problems. Each problem consisted of a statement about a conditional relation, followed by four multiple choice questions. On one questionnaire, called the verbal, all three problems were presented in words only; on the
other, called the drawings questionnaire, the first two problems were accompanied by simple drawings, and the questions referred to these drawings. The third problem was couched in verbal form on both questionnaires, and so served as a control problem.

Students' answers were called Formal if they were correct. The non-Formal responses were classified as follows: if students responded as if the conditional were instead a biconditional, their responses were called Transductive after Knifong (1974); all other responses were called Intermediate. Roughly half the subjects, group A (37), were given the verbal questionnaire, and the rest, group B (43), were given the drawings questionnaire. A week later, all students were given the opposite form of the questionnaire.

4. Findings and Interpretations

The author first summarized the results from the initial trial. (A line of type has apparently been omitted: I believe the first sentence in the Results section should read: "Table I shows the results obtained on the verbal questionnaire by group A and on the drawings questionnaire by group B.") For each of the three problems, the table contains the number and percentage of responses in each of the three categories, which the author ranks from best to poorest as Formal, Intermediate, and Transductive. Thus, his definition of "doing better" is more (percentagewise) responses in Formal as opposed to the other two categories, and/or a higher percentage of responses in Intermediate than in Transductive.

Using the Mann-Whitney U test, he found no significant differences on the control problem, concluding that the two groups were equivalent. On the other two problems, he found significantly poorer performance from the drawings condition than from the verbal condition. The second analysis compares each group's initial performance with that on the other form a week later. Using Wilcoxon matched-pairs
signed-ranks tests to analyze same-group performance, the author found no significant difference between questionnaires on the control problem, concluding that simple repetition did not produce large effects on performance. However, group A, which did the verbal form first, did significantly better than group B on one problem, better (but not significantly) on the other. He concludes, "Thus, for the CAR problem, doing the drawings questionnaire first resulted in a significant decrement in performance on the verbal questionnaire even at a one-week interval" (p. 84). Comparison on the drawings form showed that group A, which did the verbal form first, did better than group B for both problems. "Thus, doing the verbal form first resulted in relative improvement on the problems using drawings" (p. 85). He remarks that this significant effect was not the result of more formal responses, but of an increase in intermediate responses. He concludes: "These results indicate that using line drawings as concrete references in conditional reasoning problems markedly reduces performance compared to purely verbal presentations" (p. 85).

The author examines one possible explanation: Not only were drawings added, but also the instructions were altered to make references to them. However, he discards the changed instructions and concludes that the presence of the drawings was the major factor in the poorer performance on the drawings form.

Markovits does not attempt to generalize the result; however, he says, it indicates that specific examples for illustrating complex reasoning problems should be chosen with care. In this case, he says, the drawings performed two possible functions. They presented concrete instances of both P and not-P. On the other hand, they may have concentrated a subject's attention on a specific instance of P or not-P, while not adding any more information than was present in the verbal form. Therefore, they could thus provoke a "concrete" mode of reasoning in which no attempt is made to go beyond the "givens" of the problem to the possibilities...
relatively poor performance on the drawings questionnaire indicates that this second effect is predominant. Thus making a conditional reasoning problem more 'concrete' by tying it down to specific examples while conserving the logical structure of the basic problem may make it more difficult for subjects to reason correctly (p. 86).

**Abstractor's Comments**

In reading this article, I had two difficulties in following the author. One, the omission mentioned above, would have been less troublesome if Table I had been labelled more clearly. The two groups were referred to in the text and in Tables II and III as groups A and B; why not in Table I also?

The other problem I had was in visualizing the actual test items; I believe a sample item would have been clearer and would not have taken much more room than the description.

There were a few statements about which I would have liked more information. How were the subjects chosen? Why 37 and 43? My guess is that intact classes were used; why not say so? Was conditional reasoning a part of the course content, before or after the experiment, or was it totally extraneous?

The author mentions the "preliminary results" which led him to go against intuition in designing this study. I would have liked to know what those preliminary results were.

Now to more substantive matters. I am not satisfied with his classifying Intermediate responses as better than Transductive responses. While a Transductive response is erroneous (thinking that "P implies Q" is equivalent to "Q implies P"), it at least represents a systematic attempt to reason. I judge it better than
the "Christmas trees" or random marks that college students often make when they have no idea of the correct answers. Such guesses, if they were made, would have turned up in Markovits's Intermediate category, since it was to contain "all other response patterns" (p. 83). His judgment, that Intermediate responses are better than Transductive, is important, since an increase in Intermediate responses is the basis for one of his conclusions.

If the observed effect does warrant the author's conclusions, then it is curious indeed, and worthy of further consideration. The obvious question is, why did this happen?

Markovits presents one possibility: that the drawings provoked a "concrete" mode of reasoning. This conjecture begs for further exploration. The drawings may have provoked the response, "Oh, this is something I should be able to figure out from the drawing, or possibly from my common sense." If so, why was there then no attempt to go "beyond the 'givens'", as he says? Because when they had given their concrete answer, they were finished. They did not connect the drawings with a regular verbal statement of a conditional relation.

I am reminded of Schoenfeld's (1986) talk before a group of mathematics teachers, about his investigations of students' constructions and proofs. Asked to construct a circle internally tangent to both sides of an angle, he said, students could eventually do it. But having just written a proof which supported their method did not make the construction any easier, even when that proof was still visible on the chalkboard. The students still approached it as a totally new problem. Sitting in the audience were, I suspect, many teachers who had told their geometry students, "A picture is not a proof." Surely we never intended for them to separate drawings and logic so permanently.
But a drawing may tend to dominate a more concrete student's thinking if it is not yet fully developed. Consider what it means to learn a definition. I suggest that there are at least two stages. (Assume that the student has command of the individual vocabulary words used.)

Stage 1: Learning (as in memorizing) the definition—that is, becoming able to repeat it, write it, even paraphrase it, and give examples (which could have been memorized also). Stage 2: Being able to apply the definition to decide whether something fits it or not—that is, presented with an instance, being able to decide "Is" and given a non-instance, being able to decide "Is not." Even when we think a student has reached Stage 2, some curious things happen.

Present a rectangle, and ask a student whether it is a parallelogram or not. That the student can quote the definition of a parallelogram does not guarantee success; a student may answer, "No, it is a rectangle." The properties of the rectangle somehow dominate and the student does not notice that the properties of a parallelogram do appear to be present. Unfortunately, the student can not then apply the various properties with respect to a rectangle, as in "Alternate interior angles of a rectangle are congruent." The incomplete concept is that of class inclusion, but here I believe the drawing does interfere, or dominate to the point that the student makes an incorrect answer. Optical illusions are other examples; even when we know that the two segments are the same length, our eyes stubbornly tell us that one is longer.

Markovits's case is not as clear for me, because what he is asking students to do is more complex than these examples. If the presence of the drawing does interfere, with which part of the process is it interfering? To answer, we have to examine the process. To correctly answer his conditional reasoning problems, a student must somehow go through these steps.
Step 1. Understand the form of a conditional statement, and the underlying truth table (whether in form of a table or not).

Step 2. Identify whether the given statement is of that form (an implication as opposed to a conjunction, for example).

Step 3. If yes, identify which clause is P and which is Q.

Step 4. From the given information, draw the conclusion called for in the question.

If the author's conjecture is correct, then the student might not, upon seeing the drawings, identify this as a logic problem at all and so might not attempt to execute Steps 2, 3, and 4. Consequently, we will never know whether they could have done so correctly.

There are other intriguing questions. For those who did the drawings first: when, a week later, they saw the verbal form of the CAR problem, did they, using good problem-solving strategies, think, "Wait; I saw something about heavy cars last week. Is it the same or not? What did I get for an answer then?" In other words, if the previous experience with drawings interfered, what was the nature of the interference?

References


Abstract and comments prepared for I.M.E. by THOMAS O'SHEA, Simon Fraser University.

1. Purpose

The study was designed to gather details of twelfth-grade mathematics programs from a number of countries. Information included the content of the mathematics curriculum, achievement and attitudes of students, and the ways in which mathematics is taught.

2. Rationale

Mathematics educators, concerned citizens, and national officials need to have information in order to analyze their school programs, to identify areas of strength and weakness, and to plan future directions for mathematics education in their own countries.

3. Research Design and Procedures

The research was conducted as part of the Second International Mathematics Study (SIMS). A representative sample of 237 twelfth-grade mathematics classes and their teachers was selected to respond to a number of internationally-developed instruments. Included in the sample were 46 calculus classes in which students followed the AP calculus syllabus and 191 "precalculus" classes in which students had studied various combinations of trigonometry, college algebra, analytic geometry, and introductory topics in elementary functions and calculus. Data were gathered at the beginning and end of the 1981-82 school year.
4. Findings

The teachers were asked to indicate whether the mathematical content needed to respond to each achievement item that had been taught during the year, was assumed to have been taught in previous years, or was assumed never to have been taught. This information, along with achievement data, is shown in Table 1. The number of achievement test items for each content area is shown in parentheses. All items were in multiple-choice format with five choices.

<table>
<thead>
<tr>
<th>Content Area</th>
<th>Taught before</th>
<th>Taught this yr.</th>
<th>Pre</th>
<th>Post</th>
<th>Taught before</th>
<th>Taught this yr.</th>
<th>Pre</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>ts/relations (7)</td>
<td>31</td>
<td>50</td>
<td>48</td>
<td>54</td>
<td>50</td>
<td>40</td>
<td>66</td>
<td>64</td>
</tr>
<tr>
<td>number systems (17)</td>
<td>39</td>
<td>42</td>
<td>33</td>
<td>38</td>
<td>75</td>
<td>14</td>
<td>43</td>
<td>48</td>
</tr>
<tr>
<td>algebra (26)</td>
<td>34</td>
<td>52</td>
<td>35</td>
<td>40</td>
<td>53</td>
<td>41</td>
<td>53</td>
<td>57</td>
</tr>
<tr>
<td>geometry (26)</td>
<td>21</td>
<td>40</td>
<td>24</td>
<td>30</td>
<td>41</td>
<td>26</td>
<td>35</td>
<td>38</td>
</tr>
<tr>
<td>fns. &amp; calc. (46)</td>
<td>8</td>
<td>37</td>
<td>18</td>
<td>25</td>
<td>9</td>
<td>83</td>
<td>26</td>
<td>49</td>
</tr>
<tr>
<td>mb. &amp; Stats. (7)</td>
<td>29</td>
<td>14</td>
<td>36</td>
<td>39</td>
<td>50</td>
<td>6</td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>site math. (4)</td>
<td>29</td>
<td>21</td>
<td>25</td>
<td>29</td>
<td>62</td>
<td>8</td>
<td>36</td>
<td>38</td>
</tr>
</tbody>
</table>

An "opportunity to learn" (OTL) index was defined as the sum of the "taught before" and "taught this year" percentages. Thus the OTL index might be thought of as an upper limit on how well students might be expected to perform. Because the tests were developed and administered internationally, no one country found that the tests fit their curriculum exactly. For example, in Table 1, the OTL for elementary functions and calculus was only 45% for the precalculus classes, yet these items constituted about 35% of the test.
Precalculus students showed modest achievement gains in all content areas, whereas the achievement of calculus students ranged from a very large improvement in elementary functions and calculus to a small loss in sets and relations. In comparison with the overall sample of countries, the median of the U.S. sample was at a level markedly below the overall median level of performance. The U.S. calculus classes, however, performed at or near the international median in almost all content areas.

Twenty of the achievement items were also included in the First International Mathematics Study in 1964. Modest gains in performance on these items were indicated, but most occurred as a result of superior performance by the calculus students.

The teachers were about evenly divided between men and women (52% and 48%). Median responses indicated a teacher who typically was 41 years old, with 18 years of teaching experience, 10 at the Grade 12 level. The median preparation for teaching consisted of 16 semester courses in mathematics, 2 in mathematics methods, and 4 in general methods and pedagogy. Thus teachers in the sample classrooms were well-trained and experienced.

From the teacher responses to a set of items related to effective teaching "the picture that emerged was one of an emphasis on an orderly, subject-matter-oriented classroom, with clear concern for reinforcement and substantive feedback." Of eleven possible goals for teaching mathematics, teachers rated the most important to be developing a systematic approach to problem solving. The lowest rated goal was understanding the nature of proof.

Teachers spent approximately 40% of instructional time on developing new material, 20% on review, 10% on administration, and 30% on supervising students' work (including testing). The textbook was
the most commonly and consistently used resource among a narrow range of instructional resources. About 20% of teachers reported that students did not, or were not allowed to, use calculators.

The students reported that they spent most of their time listening to teacher presentations (130 min/wk), doing seatwork (60 min/wk), and taking tests (45 min/wk). In responding to several attitudinal scales, students gave high ratings to the usefulness and importance of mathematics to society but gave the lowest rating to mathematics as a process rather than as a static body of rules and content to be mastered.

5. Interpretations

The subsample of calculus classes performed at or above the international median in all content areas, whereas the precalculus classes performed below the median. This reflects, in part, the higher ability of students in calculus classes. However, the proportion of U.S. students in calculus classes is less than that for most of the developed countries in the study. Thus, the students studying calculus are highly selective and it is disappointing that they did not do better.

Curriculum for precalculus classes showed considerable diversity in content and time allocation. This may indicate curriculum drift. By contrast, the calculus classes had a more focused curriculum and spent a greater amount of time on a selected set of goals. They also spent larger blocks of time on specific content areas, rather than fragmenting time into one- or two-day coverage of concepts and skills, as was the case for precalculus classes. This pattern suggests a lack of focus in the curriculum for precalculus students, with the absence of a clear set of shared goals and a strategy of devoting large blocks of time to intensive study of the content related to these goals.
**Abstractor's Comments**

The authors of the article have taken on a formidable task. The SIMS study is an extremely ambitious undertaking which involves many countries and numerous instruments for collecting data. A 12-page article must, of necessity, be limited in its coverage. In this case, the authors chose to focus on the findings themselves rather than on broader issues of interpretation and inference.

For the most part, the findings are presented clearly, and only the most salient have been selected. There are, however, some discrepancies between the achievement information presented in Table 3 and that shown in Figures 4 and 5. It would also have been helpful to have the opportunity to learn indices included in Figures 4 and 5 so one could assess the degree to which students achieved relative to the extent to which the material was taught.

In their conclusion, the authors claim that the calculus classes performed as well as the international sample in all content areas. From Table 3 this was clearly not the case for geometry where the U.S. sample median was 38 and the international median was 42. This poorer performance may be because calculus classes covered only about two-thirds of the geometry content tested.

One of the fears of the personnel responsible for carrying out the SIMS study was that the results would be used as an international "horse race" as was the case with the First International Mathematics Study. In a sense, the authors have done this by focusing on international comparisons of achievement. The problem is exacerbated by presenting only the mean scores for each content area. Why, for example, were no international comparisons given of the selectivity of school systems, teacher qualifications and experience, classroom organization, homework expectations, and student attitudes? All these
variables may relate to achievement, and they are also of interest in their own right.

The U.S. report was the third of a series of national reports to come out of the SIMS. Each has dealt with issues of concern to the mathematics education community in the individual country. I look forward to seeing reports from the international SIMS committee which will integrate findings across countries and draw inferences for mathematics education globally.
1. **Purpose**

Two studies were conducted to investigate "systems of rules" (systematic errors) used by children in grades 6 through 9 in selecting the larger of two given decimals. The investigators sought to "demonstrate the evolution of an expert's knowledge through an elaborated learning path" for this limited domain.

2. **Rationale**

The investigation stemmed from the bases of a) differences in reasoning used by experts and novices have been demonstrated; b) other investigations have dealt with sets of rules which explain developmental trends in the acquisition of whole number place value and other mathematical content; and c) investigations have found two major systematic errors associated with decimal comparison. The researchers felt that, although not a longitudinal study, finding grade level differences in the use of these errors would "establish an order between the systems of rules...from the more primitive to the more advanced as measured relative to the expert's knowledge."

3. **Research Design and Procedures**

Two specific error types made by students in comparing decimals served as the central focus of both studies. Choosing one of two decimals as larger on the basis of the number of decimal places leads to both error types: Rule 1 - choosing the longer as bigger; Rule 2 - choosing the longer as smaller.
A third rule, Rule 3, is choosing the decimal with zero tenths as smaller. Although all items reported indicate the decimal with zero tenths was also longer (Rule 1 would be incorrect; Rule 3 is correct), the researchers saw Rule 3 as "a version of Rule 1."

The first study involved 21 sixth graders who had just completed a chapter on decimals. No other curricular or population/sample description is presented. All students responded to about 60 questions during an individual 45-minute interview. Questions centered on 1) decimal-pair comparison; 2) ordering fractions, wholes, and decimals; 3) the density of numbers; 4) the relative size of the product of a decimal and whole number; 5) writing numbers from oral presentation; and 6) stating the value of a specified decimal digit.

The second study focused upon the results from administering a 30-item test of decimal comparison. The subjects were 74 seventh-, 106 eighth-, and 60 ninth-grade children (no additional description is provided). The test was designed to discern Rule 1 and Rule 2 use with and without the presence of a zero in the tenths or the right-hand-most column of a decimal.

4. Findings

The pattern of the sixth graders' answers combined with their explanation of reasoning led to the identification of 19% who used Rule 1; 33%, Rule 2; 14%, Rule 3; 14% whose errors could not be classified; and 19% who answered all questions correctly. Of particular note is the rationale provided by those using each of the various rules:
Rule 1  \[4.63 > 4.8\] "because 63 is bigger than 8"

Rule 2  \[4.4502 < 4.45\] "because the 5 is hundredths and the 2 is ten-thousandthc which is smaller"

Rule 3  \[4.08 < 4.7\] "because there's a zero here"

Giving two equivalent decimals of different length (2.35 vs. 2.350) led to inconsistencies in use of the rules (most responded correctly).

In writing decimals and stating columnar values, generally Rule 1 users made errors and Rule 2 users were correct.

The instrument used was reliable (0.92) and a factor analysis supported the consistency of items relative to the two error types. A third factor did appear, though unexpected, involving the items with equivalent decimals.

Percentages of Rule 1 and 3 (combined) or Rule 2 users are reported in the undefined categories of high ability and low ability within grade levels. In all grade X ability groupings, Rule 2 was used by a percentage greater than, or equal to, that of the Rules 1 and 3 users. There is a steady decline in the percentage of students of high ability who use Rule 1, from seventh to ninth grade. Percentages of Rule 2 users do not steadily increase or decrease across grades for either ability group. High ability students did progressively increase in percentage of those answering correctly; 56% of the seventh graders, 62% of the eighth graders and 82% of the ninth graders got every item correct.

It is particularly interesting that the percentages of low ability students who could not be classified ranged from 27% to 44%.
5. **Interpretations**

The researchers felt that the results support the assumption that basically only two types of errors are made in comparing decimals. No discussion was offered relative to the rather high percentages of students who made "unclear" errors.

Explanations are offered for the bases of the particular systematic errors. Sixth graders who used Rule 1 seemed to treat the decimal portion as though it were a whole number. That is, the same rule used with wholes was related to decimals. Rule 2 users seemed to think of decimals as fractions, "but they failed to make the coordination between the size of the part, the fraction (e.g., tenths, hundredths, etc.) and the number of such parts."

The interview responses of Rule 1 and Rule 2 users concerning decimal place value led the researchers to conclude that Rule 1 is more primitive than Rule 2. They conclude that they have discovered the "intermediate, transitional systems of rules, each relying on a previous partly learned knowledge...rules (that) can be so ordered as to form a learning path for this domain.”

**Abstractor's Comments**

This investigation reaffirms the importance of diagnosing errors rather than simply declaring that a particular frequency of errors was made. The attempt of the researchers to so strongly state that the data supports a consistent, ordered development-by-stages theory for learning comparison of decimals does raise several questions:

1. There were large percentages of students apparently not using Rule 1 or 2. Are there other systematic errors existent, but not discerned?
Unfortunately the report fails to include sample sizes for the grade X ability level cells. Using the maximum sample sizes (which lead to very conservative percentages), roughly 20% of all students in both studies were in fact making errors not classified. One possible error (which is closely related to Rule 2 from the description of reasoning provided) is comparing decimals simply upon the basis of the last digit, regardless of decimal length. The theory purported offers no explanation for errors other than Rule 1 and Rule 2.

2. Are the results consistent with those of other studies?

Using data provided in the report from related studies, the table below suggests less than consistency of results.

<table>
<thead>
<tr>
<th></th>
<th>Rules 1 and 3</th>
<th>Rule 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>6th</td>
<td>Others</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Current</td>
<td>33</td>
</tr>
<tr>
<td>7th</td>
<td>Others</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Current Hi</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Current Lo</td>
<td>27</td>
</tr>
<tr>
<td>8th</td>
<td>Others</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Current Hi</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Current Lo</td>
<td>22</td>
</tr>
<tr>
<td>9th</td>
<td>Others</td>
<td>-</td>
</tr>
<tr>
<td>Up</td>
<td>Current Hi</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Current Lo</td>
<td>7</td>
</tr>
</tbody>
</table>
At grades 6 and 7, others have found much greater percentages using Rules 1 and 3 than the current study. Different, too, is the percentage of Rules 1 and 3 users versus Rule 2 users. Others find a much greater proportion of Rules 1 and 3 users. Such is not the case in the current study for which data are available. In fact, the percentages of Rule 2 users have been reported consistently higher or equal to that of Rules 1 and 3 users. These inconsistencies suggest that something other than age might be the significant variable.

3. If age is not a significant variable, how does one explain the "clear picture of the trend of change in the distribution of rules" over grade levels?

Data reported as percentages are misleading if sample sizes are ignored. More dangerous is the comparison of percentages of different sized populations. An additional subject in one sample is not equivalent in percentage to a single subject in the other study if sample sizes are not the same. An analysis of the report at hand is even further hampered by the lack of data reported concerning the n's per cell.

It appears as though there is a decreasing trend in use of Rules 1 and 3 on the part of high ability students, from 7% to 4% to 3% (see Table 1). But with only one more ninth grader and one less seventh grader (again, using the most conservative n's possible), absolutely no trend remains.

The "clear picture" the investigators refer to is less than clear from the data they report. The trend of which they speak is tenuous at best, as is any suggestion that Rule 1 users become Rule 2 users, who in turn become experts.
4. What do the data suggest if not "the evolution of an expert's knowledge through an elaborated learning path?"

Totally ignored throughout the report is the significant number of students who are classified as acting like "experts." An expert, as defined by the researchers, is "one who has integrated all these number sets (whole numbers, fractions, and decimals) into a one coherent number system." It appears that the operational definition for expert is "getting all items correct." Now it might be that the percentages are again misleading, but 19% in sixth grade perfectly correct progressing to 82% in ninth grade seems rather significant. Since we know that instruction is taking place, it seems only logical to attribute the increase in experts (and thus reduction in any error classification) to learning. Unfortunately for the theory purported, it appears that the effect is due to instruction over time, not some inherent, physiological, age-related evolution.

The most significant contribution offered by this study is the identification of bases for errors made. The interviews with the sixth-graders did in fact reflect cognitive (albeit, defective) schema. Of particular interest for future investigations might be:

a) Do particular models (e.g., decimals as money: .43 is 43 cents) instill greater use of Rule 1 versus Rule 2?

b) What models do texts use? When a new model is presented, what is the effect relative to Rule 1 versus Rule 2 use?

c) Are systematic errors in decimal comparison merely the result of an abbreviated rule statement or of more deep-seated problems?
d) What are absolute prerequisites for decimal comparison mastery? Is it necessary and sufficient to understand fractions (whole number place value) prior to mastery of this skill?
1. Purpose

The three-part study investigated error patterns in solving open sentences, strategies leading to these errors, and the effects of instruction on two of the strategies.

2. Research Design and Procedures

The open sentences represented the twelve possible derivations from the form $A \square B = C$ obtained by having the placeholder ($\square$) occupying each possible position ($A$, $B$, or $C$) for each of the four basic operations. Two sentences were prepared for each form; one of the sentences involved basic facts, the other a pair of two-digit numbers. The students were fourth- and fifth-graders for the first part of the study, fifth- and sixth-graders for the second part, and sixth-graders for the third part.

3. Findings

One error pattern identified in the first part of the study was failure on open sentences of the form $\square - B = C$ and $\square \div B = C$. Another was failure on sentences of the form $A - \square = C$ and $A \div \square = C$. The two strategies hypothesized for the production of these errors were called "finding the solution" (FS) and "inverse operation" (IO). FS is characterized by the student actually attempting to find the numerical solution, but failing in the attempt. IO is characterized by applying a rule—if the placeholder is in position $A$ or $B$, perform the inverse of the operation indicated. Since this rule leads to a
correct solution in some open sentences (for example, \( \Box - A = B \) and \( \Box + A = B \)), it is applied to all.

In the second part of the study, students making the two types of error were interviewed and asked to explain their answers. Fifth-graders interviewed confirmed that they used the FS strategy; none of the fifth-graders mentioned an IO strategy. Some of the errors by sixth-graders showed the IO pattern, and the interviews confirmed that they had used this strategy.

The instruction that comprised the third part of the study focused on the meaning of the operations. Posttest results immediately after instruction showed no student using either of the strategies. A retention test after six weeks, however, showed only two students using FS, but a large number using IO. The group using IO included some of those who had used this strategy on the pretest together with some who had used FS and some whose pretest errors were unclassified.
1. Purpose

"The purpose of the present research was to investigate the mathematics achievement effects of three commonly proposed methods of dealing with student heterogeneity: individualized instruction, within-class ability grouping, and whole-class instruction."

2. Rationale

One of the most consistent problems faced by the school mathematics teacher is differences in student preparation and learning rates. For every lesson there are students for whom the pace is too fast and the material too difficult and there are students for whom the pace is too slow and the material too easy. The most common way of dealing with the problem of heterogeneity is to use some form of ability grouping. The three most commonly used and most commonly researched methods are between-class ability grouping, within-class ability grouping, and individualized instruction. The present study compares the three methods of ability grouping.

3. Research Design and Procedures

The article presents the results from two experiments which were designed in the same way. First, the differences between the two experiments.
Experiment 1 used 345 students in 15 grade 4-6 classes in an urban area in which heterogeneous class assignments were mandated as part of a desegregation plan. Approximately 71% of the students were white, 26% black, and 3% Asian-American. Experiment 2 used 480 students in 22 grade 3-5 classes in a relatively homogeneous rural area which used between-class ability grouping to reduce the heterogeneity of mathematics classes. Approximately 91% of the students were white, 7% black, and 2% Asian-American. In Experiment 2, in addition to the three treatments described below, there was an "untreated control group, in which teachers used traditional whole-class instructional methods."

Both experiments used the following experimental treatments:

**Missouri Mathematics Program (MMP):** A regular sequence of teaching, controlled practice, independent seatwork, and homework, with emphasis on a high ratio of active teaching to seatwork, teaching mathematics in the context of meaning, frequent questions and feedback, rapid pace of instruction, and management strategies intended to increase student time on-task.

**Ability Grouped Active Teaching (AGAT):** Same as MMP, but used within-class ability grouping. On the basis of an initial test, students in each AGAT class were divided into a high group (about 60% of the students) and a low group. Teachers were instructed to "push the pace for the high group."

**Team Assisted Individualization (TAI):** Students worked in heterogeneous four- or five-member learning teams on individualized mathematics materials at their own levels and rates. Students in the teams helped one another with problems and took responsibility for almost all checking, routing, and other management tasks which freed the teacher to work with three regularly constituted teaching groups.
composed of students (drawn from many teams) performing at the same level in the materials.

In both experiments mathematics achievement was measured using the appropriate forms of the Comprehensive Test of Basic Skills (CTBS) with the district-administered California Achievement Test (CAT) as a covariate. Since different grades used different tests, all achievement scores were transformed to T scores. Two eight-item scales (Liking of Math Class and Self-Concept in Math) were used as pre- and posttests to measure attitude.

All teachers were observed to determine whether or not they were implementing the critical features of their treatments.

The adjusted CTBS scores were analyzed using random-effects analysis of variance; the factors were treatment and class/teacher within treatment. If the overall nested analysis of variance was statistically significant ($p < 0.10$), individual-level planned comparisons between treatment means were compared using a modified Bonferroni procedure. Individual-level analyses of covariance were conducted to look for interactions between treatment and students of different levels of past performance, race or sex.

4. **Findings**

In both experiments all MMP, AGAT, and TAI teachers were found to be implementing the major components of their methods.

In both experiments there were no pretest differences with respect to computation. The computation achievement results for TAI and AGAT were almost identical and significantly higher than MMP. In addition, in Experiment 2, the computation results for all three experimental treatments were significantly higher than the control group.
While in Experiment 1 there were no pretest differences with respect to Concepts and Applications, there were pretest differences found in Experiment 2. Even though random assignment was used, there were high pretest scores in AGAT classes and low scores in the Control classes. There were no main effect differences with respect to Concepts and Applications in either experiment.

In Experiment 1 the results on the Liking of Math Class attitude scale reflected the achievement results, with TAI and AGAT yielding similar results and both being higher than MMP. In Experiment 2 on the same attitude scale TAI exceeded all other groups. AGAT and MMP did not differ, but MMP exceeded the Control group while AGAT did not. On the Self-Concept in Math attitude scale in Experiment 1, TAI students scored much higher than AGAT and MMP students who did not differ from one another. The three experimental groups and the control group did not differ on the same scale in Experiment 2.

No significant interactions were found in either experiment.

5. Interpretations

Given the differences between the urban, integrated, untracked schools in Experiment 1 and the rural, mostly white, tracked schools in Experiment 2, the results are amazingly similar. In both experiments TAI and AGAT increased computational skills markedly more than MMP and, in Experiment 2, traditional whole-class instruction. No differences were found between TAI and AGAT with respect to achievement. No differences were found in Concepts and Applications. All achievement differences were main effects with no interaction effects found to be significant. Such similarities were not expected; in particular, it was assumed that TAI and AGAT would be more effective for those students performing farthest from their class means and in settings with the greatest degree of student heterogeneity.
The results for TAI lend support to the argument that if the inherent problems of management, motivation, and lack of direct instruction could be solved, individualized instruction could be made instructionally effective.

The results for AULT lend support to the argument that if the difficulty of managing multiple ability groups could be solved, within-class ability grouping may be a particularly effective procedure. Most impressive was the result that low ability students in AGAT gained significantly more in Computation than did all students in MMP or Control classes.

In addition to the achievement and attitude results of the study, the TAI treatment was supported by the teachers. At the end of the experiment the teachers were allowed to select any of the methods other than the one they used in the experiment in which to receive training and materials. Every eligible teacher selected TAI; every TAI teacher continued to use the program during the next school year.

**Abstractor's Comments**

It is extremely difficult to report the results of the present type of study in a journal such as *AERJ*. One cannot provide an adequate description of the treatments to give the reader a "feel" for what happened in the classroom. The present article is no exception. After reading the article several times, I still am not clear what the teachers were doing in anything more than just general terms.

Some questions simply do not get addressed in the article. Given the care with which the study was obviously designed and implemented, it is likely the answers to the following questions are available, but no room existed in the article.
a. Why was the TAI treatment group given a break from the treatment every fourth week?

b. Why was there no Control group in Experiment 1, given there was one in Experiment 2?

c. What did the Control teachers do? Given all the ecological studies completed in the last decade we surely know that descriptions such as "traditional whole-class instructional methods" have different meanings to each reader.

d. What does a single Grade Equivalent score mean when three different grades were involved?

e. The authors make the point of mentioning, by name, the teachers trainers for TAI and AGAT in Experiment 1, but do not mention who trained the other teachers in Experiment 1 or any of the teachers in Experiment 2. Did having the primary developer of TAI train the teachers in Experiment 1, but not Experiment 2, have any effect?

f. Why was the alpha level set at 0.1? Maybe I'm just a traditionalist, but when I see an alpha level other than 0.05 or 0.01 I wonder why.

The lack of significant interaction effects is extremely surprising! TAI and AGAT were designed to accommodate diverse achievement levels. Were the levels not diverse enough? It may be as the authors suggest, the treatments are better because "they provide more effective instruction in general." Both TAI and AGAT are "highly structured instructional models." Many more hours went into preparing the lessons than a classroom teacher could ever hope to spend. What the authors may have to offer the reader is a lead to some very good instructional materials and techniques.
Given the last paragraph one shouldn't necessarily agree with all the things the authors read into the results. They preface their conclusions about TAI and AGAT with some rather large "ifs". The problems which must be overcome before getting on the bandwagon for TAI or AGAT are larger than the results of the study. One should take the study for what it is—a well-designed and implemented study to compare some special treatments. The results certainly provide food for thought to mathematics education researchers, school district office personnel, and classroom teachers.

Abstract and comments prepared for I.M.E. by ROBERT B. ASHLOCK, Belhaven College and the RTS Graduate School of Education, Jackson, Mississippi.

1. Purpose

To document how children spontaneously use strategies for deriving basic addition and subtraction facts from known facts, and to study how training in the use of strategies for deriving facts from known facts influences children's use of strategies to solve addition and subtraction problems.

2. Rationale

Although there is considerable documentation that children use counting strategies to solve addition and subtraction problems, less is known about the use of non-counting strategies. Strategies for deriving facts from known facts have been described in several studies and there is evidence that many children use such strategies spontaneously, usually before they have learned the basic facts at a recall level. However, many other children continue to use counting strategies. There is little research to show how actual instruction in strategies for deriving facts from known facts affects the thinking processes children use when solving addition and subtraction problems.

3. Research Design and Procedures

A teaching experiment was conducted with one second-grade class \(N = 23\) in a middle-class neighborhood. Beginning in early September an instructional unit was taught by the regular classroom teacher for eight weeks. Four main interviews were conducted with each child:
a pretest, an interview in the middle of the instructional unit, a posttest, and a long-range-effects test. Also, short daily interview were conducted. Lessons were observed and anecdotal data collected. A timed group test was given along with the first three main interviews.

During the instructional unit number facts were presented in relation to derived facts strategies; facts with a similar structure were grouped together. For addition, doubles were presented first, then addend pairs differing by one, addend pairs differing by two, and facts easily related to 10. Different strategies were presented for solving each category; e.g., three strategies were presented for addend pairs differing by two. The main strategies for subtraction were the think addition strategy and strategies derived from addition strategies.

The first half of the unit focused on strategies for addition problems; the second half was devoted to subtraction. Children modeled strategies with manipulatives, discussed the strategies, and completed worksheets. The worksheets were corrected daily and systematic errors were noted.

Interviews included word problems with different semantic structure; join, separate, compare, join missing addend, and missing minuend. Also included in the interviews were addition and subtraction number combinations. The word problems and the number combinations represented derived fact strategies that were taught. The interviews focused on which word problems were solved, what strategies were used and whether children were able to apply their knowledge of strategies for deriving facts from known facts.
4. Findings

On the pretest about one-fifth of children's responses to addition combinations and about one-fourth of their responses to subtraction combinations incorporated varied strategies for deriving facts from known facts. Certain derived fact strategies were used consistently by a few children. It was observed that the strategies were used both with and without finger counting. Derived fact strategies were also used with word problems, especially with the join missing addend problem. However, they were not used with the missing minuend problem.

For addition combinations, the use of derived fact strategies more than doubled during instruction, and the increase was still observed two months after the end of instruction. Also, more children were observed using the strategies. However, the use of derived fact strategies did not increase as much for subtraction. What did increase was the proportion of subtraction derived fact strategies based on addition strategies.

Although instruction focused on number combinations, about the same percent of children who used derived fact strategies with combinations also used them with word problems.

An examination of children's individual patterns revealed three profiles with reference to use of derived fact strategies: children who increased in their use of the strategies along with recall, those who moved from mainly counting to mainly using derived fact strategies, and others who used few if any derived fact strategies throughout the study.

Classroom observations and daily interviews produced specific data which are summarized in the report. For example, derived fact strategies were sometimes learned as rote procedural rules without an understanding of relationships involved. Although data from the timed
tests did show significant increases in the proportion of correct answers for both addition and subtraction, correlations with individual interviews did not suggest that using derived facts strategies was a quicker method of solution than counting.

5. Interpretations

During the study children did change solution strategies considerably, from mainly using counting strategies to using derived facts. The lesser increase for subtraction may be accounted for by the fact that less time was spent teaching the subtraction derived fact strategies, difficulty of the subtraction tasks (as other research suggests), difficulty in identifying doubles in subtraction, and the fact that subtraction strategies have greater memory requirements.

However, the study did not resolve the question of prerequisites needed for learning derived fact strategies. It was clear that children do not have to wait until they use counting-on strategies regularly before they can learn to use derived fact strategies.

Furthermore, the study did not establish whether extensive use of derived fact strategies leads to recall of number facts. It is possible that instruction in derived fact strategies influences how the facts are represented in, and retrieved from, long-term memory. In fact, with practice a derived fact strategy may become an automatic retrieval process which is used unconsciously and is not distinguishable from recall.

By the end of the study most children could use a variety of derived fact strategies as well as a variety of counting strategies, but they did not always use their most sophisticated strategies.
Abstractor's Comments

Amidst calls to teach children how to derive basic facts from facts they can recall and evidence that some children appear to do this on their own, this study of "spontaneous" derived fact strategies and the effects of training children to use derived fact strategies is most welcome. Certainly, a teaching experiment was the appropriate method of investigation.

As is typical with a teaching experiment, subjects do not represent a random sample; they constitute a small intact group. The "spontaneous" problem-solving strategies of children in this group were influenced by particular previous experiences at home and during grade one. Appropriately, the investigator does not imply that these children's "spontaneous" derived fact strategies are typical of other groups. We wonder if replications of the study with different groups of children would produce different effects.

Twenty years ago Gray demonstrated the efficacy of teaching a derived fact strategy in grade three with multiplication facts using the distributive property. He made sure his students had no previous instruction in multiplication. Steinberg has demonstrated the effects of teaching strategies at the beginning of grade two with addition and subtraction facts, but her students apparently had previous instruction in basic addition and subtraction facts. (She even noted that half of the children had difficulty recalling combinations that sum to 10.) What is needed is a similar study with children in grade one who have not previously been taught the basic addition and subtraction facts. In other words, how effective are derived fact strategies if they are introduced as the initial approach to figuring out the specific facts where they apply? There might be less dependence on counting, for children do cling to any procedure they find successful.

The report is not without limitations. For example, the investigator states that the first half of the instructional unit was devoted to teaching strategies for addition problems and the second half to strategies for subtraction problems; yet, after noting that children increased their use of derived fact strategies less with subtraction than with addition, she states that less time was spent teaching the subtraction strategies. A clarification would be helpful.

Notwithstanding, this is a generally well-conceived piece of research which addressed an important area of inquiry with implications for instructional practice. The report should be read if only to be stimulated by the excellent discussion on the relationships of derived fact strategies to the recall of basic facts.
MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN JOURNALS AS INDEXED
BY CURRENT INDEX TO JOURNALS IN EDUCATION
October - December 1985


ED 257 304 Abbott, Randy V. The LD Student in Math: Teacher Attitudes—A Summary. 8p. MF01/PC01 Plus Postage. Available from EDRS.


ED 257 628 Evaluation Report on Pilot Programs for Summer and After-School Instruction in Science and Mathematics as Reported by the State Board of Education. Submitted to the Governor, Lt. Governor and The Sixty-Ninth Legislature. 28p. MF01/PC02 Plus Postage. Available from EDRS.

ED 257 630 Chi, Micheline T. H.; Glaser, Robert. Problem-Solving Ability. 27p. MF01/PC02 Plus Postage. Available from EDRS.

ED 257 632 Pedersen, Katherine; And Others. Causal Models of Continuation in the Mathematics Study Sequence. 31p. MF01/PC02 Plus Postage. Available from EDRS.

ED 257 633 West, Jerry; And Others. An Analysis of Course-Taking Patterns in Secondary Schools as Related to Student Characteristics. High School and Beyond: A National Longitudinal Study for the 1980's. 135p. MF01/PC06 Plus Postage. Available from EDRS.


ED 257 642 Vanfossen, Beth E.; And Others. Curriculum Tracking: Correlates and Consequences. 30p. MF01/PC02 Plus Postage. Available from EDRS.

ED 257 643 Reyes, Laurie Hart; Stanic, George M. A. A Review of the Literature on Blacks and Mathematics. 38p. MF01/PC02 Plus Postage. Available from EDRS.

ED 257 653 Suydam, Marilyn N., Ed.; Kasten, Margaret L., Ed. Investigations in Mathematics Education. Volume 18, Number 1. 74p. MF01/PC03 Plus Postage. Available from EDRS.
ED 257 654 Falls, Timothy H.; Voss, Burton. The Ability of High School Chemistry Students to Solve Computational Problems Requiring Proportional Reasoning as Affected by Item-In-Task Variables. 29p. MF01/PC02 Plus Postage. Available from EDRS.

ED 257 655 Spade, JoaZ.; And Others. Effective Schools: Characteristics of Schools Which Predict Mathematics and Science Performance. 39p. MF01/PC02 Plus Postage. Available from EDRS.

ED 257 658 Brown, Rose Marie. A Determination of Attitudes toward Mathematics and an Analysis of Factors Which are Associated with Negative Attitudes toward Mathematics of Students at an Urban Community College: A Descriptive Study. 144p. MF01/PC06 Plus Postage. Available from EDRS.


ED 257 670 Song, Myung-Ja; Ginsburg, Herbert P. The Development of Informal and Formal Mathematical Thinking in Korean and American Children. 40p. MF01/PC02 Plus Postage. Available from EDRS.


ED 257 684 Berebitsky, Roger D. An Annotated Bibliography of the Literature Dealing with Mathematical Anxiety. 40p. MF01/PC02 Plus Postage. Available from EDRS.

ED 257 801 Byers, Joe L. The Relation between Academic Aptitude and Commitment to Teaching among MSU Students. Research and Evaluation in Teacher Education. Program Evaluation Series No. 5. 13p. MF01/PC01 Plus Postage. Available from EDRS.


ED 257 884 Malcolm, Shirley M.; And Others. Equity and Excellence: Compatible Goals. An Assessment of Programs That Facilitate Increased Access and Achievement of Females and Minorities in K-12 Mathematics and Science Education. 73p. MF01/PC03 Plus Postage. Available from EDRS.

ED 257 925 Evaluation of Summer Basic Skills Centers, 1984. Education Consolidation and Improvement Act, Chapter I and Chicago Effective Schools Project. 83p. MF01/PC04 Plus Postage. Available from EDRS.

ED 258 089 Bellando, Jayne; Winer, Jane L. Computer Anxiety: Relationship to Math Anxiety and Holland Types. 33p. MF01/PC02 Plus Postage. Available from EDRS.

ED 258 554 Johanson, Roger P. School Computing: Some Factors Affecting Student Performance. 14p. MF01/PC01 Plus Postage. Available from EDRS.

ED 258 698 Ashcraft, Mark H. Children's Mental Arithmetic: Toward a Model of Retrieval and Problem Solving. 23p. MF01/PC01 Plus Postage. Available from EDRS.

ED 258 726 Halford, Graeme S. How Do We Define the Limits to Children's Understanding? 21p. MF01/PC01 Plus Postage. Available from EDRS.

ED 258 791 Geary, David C.; And Others. Chronometric Analysis of Complex Addition: Verification of a General Model. 15p. MF01/PC01 Plus Postage. Available from EDRS.


ED 258 807 Reading/Mathematics, Kansas Minimum Competency Assessment Report. School Year 1982/83. 70p. MF01/PC03 Plus Postage. Available from EDRS.


ED 258 810 McNicol, Shirley; And Others. A Study of Arithmetical Problem Solving Abilities of Young Children through the Use of Calculators. 207p. MF01/PC09 Plus Postage. Available from EDRS.


ED 258 829 Kloosterman, Peter. Sex-Related Differences in Students' Reactions to Failure on Algebra Word Problems. 24p. MF01/PC01 Plus Postage. Available from EDRS.

ED 258 830 Johnson, William L. The Mathematics Laboratory and Computer-Assisted Instruction. 32p. MF01/PC02 Plus Postage. Available from EDRS.

ED 258 846 Sleeman, D. Basic Algebra Revisited: A Study with 14 Year Olds. Revision. 37p. MF01/PC02 Plus Postage. Available from EDRS.

ED 258 847 Sleeman, D. An Attempt to Understand Students' Understanding of Basic Algebra. 36p. MF01/PC02 Plus Postage. Available from EDRS.

ED 259 029 Mittelholtz, David J.; And Others. Learning and Intelligence: Ability Differences in the Organization and Recall of Elementary Statistical Concepts. 16p. MF01/PC01 Plus Postage. Available from EDRS.


ED 259 894 Connecticut Basic Skills Proficiency Test 1982-83. Mathematics, Basic Writing Skills in the Language Arts, Reading, Summary and Interpretations. 49p. MF01/PC02 Plus Postage. Available from EDRS.


ED 259 900 Mathematics Program Analysis of Grade 10 Stanford Achievement Test, Test of Academic Skills. Fall 1984 Administration. 24p. MF01/PC01 Plus Postage. Available from EDRS.

ED 259 904 Mathematics Program Analysis of Grade 6, Stanford Achievement Test, Intermediate Level II. Fall 1984 Administration. 37p. MF01/PC02 Plus Postage. Available from EDRS.

ED 259 905 Mathematics Program Analysis of Grade 8 Stanford Achievement Test, Advanced. Fall 1984 Administration. 33p. MF01/PC02 Plus Postage. Available from EDRS.


ED 259 947 LeFevre, Patricia. Rational Number Learning and Instruction from a Cognitive Perspective. 26p. MF01/PC02 Plus Postage. Available from EDRS.

ED 259 949 Shalin, Valerie L.; Bee, Nancy V. *Structural Differences Between Two-Step Word Problems.* 26p. MF01/PC02 Plus Postage. Available from EDRS.

ED 259 950 Schoenfeld, Alan H. *Students' Beliefs About Mathematics and Their Effects on Mathematical Performance: A Questionnaire Analysis.* 40p. MF01/PC02 Plus Postage. Available from EDRS.


ED 259 958 Shire, Beatrice; Durkin, Kevin. *Developments in the Use of Proximity and Ratio Cues in Velocity Judgments.* 31p. MF01/PC02 Plus Postage. Available from EDRS.

ED 260 098 Sofer, Sheldon; And Others. *Peer Teachers as Mirrors and Monitors: Second Year Evaluation Report.* 26p. MF01/PC02 Plus Postage. Available from EDRS.

ED 260 101 Elmore, Patricia B.; And Others. *A Longitudinal Study of Career Interests and Mathematics Attitudes for Students at the Eighth and Twelfth Grade Levels.* 28p. MF01/PC02 Plus Postage. Available from EDRS.