This publication is an abstracted compilation of 15 investigations selected from other journals. The information includes purpose, rationale, research design and procedures, findings, interpretations, abstractor's comments, and references. This journal includes the following reports: (1) "A Constructivist Approach to Numeration in Primary School: Results of a Three Year Intervention with the Same Group of Children"; (2) "Identifying Fractions on Number Lines"; (3) "Teachers' Pedagogical Content Knowledge of Students' Problem Solving in Elementary Arithmetic"; (4) "Representation of Addition and Subtraction Word Problems"; (5) "Effect of Cognitive Entry Behavior, Mastery Level, and Information about Criterion on Third Graders' Mastery of Number Concepts"; (6) "The Effects of Computer-Assisted and Traditional Mastery Methods on Computation Accuracy and Attitudes"; (7) "An Investigation into the Use of Microcomputers to Teach Mathematical Problem-Solving Skills to 13 Year Olds"; (8) "Differences among Women Intending to Major in Quantitative Fields of Study"; (9) "Characteristics of Unskilled and Skilled Mental Calculators"; (10) "Intrinsic Orientation Profiles and Learning Mathematics in CAI Settings"; (11) "Putting the Student into the Word Problem: Microcomputer-Based Strategies that Personalize Math Instruction"; (12) "The Mathematics of Child Street Vendors"; (13) "A Cognitive Approach to Meaningful Mathematics Instruction: Testing a Local Theory Using Decimal Numbers"; (14) "Teaching Children to Use Schematic Drawings to Solve Addition and Subtraction Word Problems"; and (15) "Effect of Assignment Projects on Students' Mathematical Activity." The titles and sources of mathematics education research studies reported in CIJE and RIE, April-June 1988, are listed. (YP)
INVESTIGATIONS IN MATHEMATICS EDUCATION
Volume 21, Number 4 - Fall 1988

THE ERIC SCIENCE, MATHEMATICS AND ENVIRONMENTAL EDUCATION CLEARINGHOUSE
in cooperation with
The Office of Educational Research and Improvement, U.S. Department of Education and
The SMEAC Information Reference Center,
College of Education, The Ohio State University

BEST COPY AVAILABLE


Carpenter, Thomas P.; Fennema, Elizabeth; Peterson, Penelope L.; and Carey, Deborah A. TEACHERS' PEDAGOGICAL CONTENT KNOWLEDGE OF STUDENTS' PROBLEM SOLVING IN ELEMENTARY ARITHMETIC. Journal for Research in Mathematics Education 19: 385-401, November 1988. Abstracted by FRANCIS (SKIP) FENNELL


Abstracted by GEORGE W. BRIGHT ....... 31

Abstracted by DAIYO SAWADA .............. 35

Abstracted by JOHN GREGORY .......... 41

Abstracted by SIGRID WAGNER ........ 47

Abstracted by VIRGINIA HORAK ....... 53

Abstracted by JUDITH SOWDER .......... 59


Willis, Gordon B. and Fuson, Karen C. TEACHING CHILDREN TO USE SCHEMATIC DRAWINGS TO SOLVE ADDITION AND SUBTRACTION WORD PROBLEMS. Journal of Educational Psychology 80: 192-201; June 1988. Abstracted by JAMES M. SHERRILL ......................... 81

Zehavi, Nurit; Bruckheimer, Maxim; and Ben-Zvi, Ruth. EFFECT OF ASSIGNMENT PROJECTS ON STUDENTS' MATHEMATICAL ACTIVITY. Journal for Research in Mathematics Education 19: 421-438; November 1988. Abstracted by DONALD J. DESSART ...................... 88

Mathematics Education Research Studies Reported in Journals as Indexed by Current Index to Journals in Education April - June 1988. .......................... 92

Mathematics Education Research Studies Reported in Resources in Education April - June 1988 ....................... 94

Abstract and comments prepared for I.M.E. by HARRIETT C. BEBOUT, University of Cincinnati.

1. Purpose

The purpose of this study was to continue the authors' investigation of the development of numeration concepts in primary grade children. Specifically, this study was designed to investigate the effects of three years of continuous constructivist-based instruction on children's development of numeration representations and procedures.

2. Rationale

The study was based on a constructivist conception of learning and instruction in which the learner is the constructor, or elaborator, of mathematical concepts and the instruction is designed to correspond to the mathematical thinking of the learner. This study continued the authors' earlier investigations that had identified children's major difficulties and misconceptions during their development of numeration knowledge (Bednarz and Janvier, 1982) and that had introduced the authors' preliminary design for instruction (Bednarz and Janvier, 1985).

3. Research Design and Procedures

Sample. The sample consisted of three different groups of children as they moved through first, second, and third grades. Although the groups shifted somewhat during the three years of the study, essentially the constructivist group consisted of 23 children, the current comparison group consisted of 23 children, and the former comparison group consisted of 40-75 children.
Instructional Treatment. The study was designed to provide different instructional treatments to the constructivist group and to the two comparison groups over a three-year period as the children moved through grades 1-3. With the authors of the study as teachers, the constructivist group was taught concepts of numeration during constructivist-based instructional treatments; with regular classroom teachers, the comparison groups were taught numeration during traditional instruction focusing on conventional symbolizing and syntax.

Evaluation. Results obtained during individual interview with children at the end of each school year were used to compare the effects of the constructivist approach with those of the traditional approaches. The interview items included both a set of items directly related to instruction (specifically, children's reasons for their regrouping procedures and associated symbolism and their successes in operating with groupings of various configurations) and a set of transfer items (specifically, children's uses of the abacus, of groupings in subtraction and division situations, and of conventional symbolism in calculation).

4. Findings

The results of the interviews were displayed in several tables and graphs. For the set of instructional items, the results included children's rationales for regrouping and symbolizing, their success in working with single groupings, and their success in working with groupings of different orders; for each of these items the authors set forth a hierarchy of observable actions or behaviors. The constructivist group consistently exhibited behaviors that were more advanced in the hierarchy than did the comparison groups. On the transfer items the same pattern of results were noted: Children from the constructivist group performed more successfully than the comparison children with groupings on the abacus, in subtraction and division situations, and in conventional symbolism.
5. **Interpretations**

The authors interpreted the results of their study with specific observations. They observed that the constructivist-based instruction had helped children to develop understandings of grouping concepts at all three grade levels and had resulted in transfer to the standard curriculum items. They observed also in the constructivist group the evolution of numeration understanding from concrete representation to the more abstract. They attributed the gains of the constructivist group to specific contextual teaching environments that emphasized appropriate representations and procedures.

**Abstractor's Comments**

The major contributions of this study stem from its focus on the important issue of children's development of numeration concepts, from its longitudinal design, and from its descriptions of children's procedures for grouping numbers. The various interview tasks and the identification and description of children's strategies on these tasks are thought-provoking and may be potentially valuable as part of the research on children's mathematical thinking. But the potential value of this study is buried in a report that presents too much background information and too much data, that uses undefined term and imprecise language, and that in essence covers too many pages.

Another problem with the study is related to the use of hierarchies to evaluate the development of numeration concepts. Although these hierarchies are interesting and may be indicative of developmental levels, they have not been established yet as researched measures and thus should be treated as less definitive for evaluation than they are in this study.

And finally, in my opinion, the investigation of and reporting about of the constructivist group could stand alone without the comparison groups. Personally I have a problem with "plots" that compare instructional outcomes from mathematics education researchers with those of traditional classroom teachers and curriculum, probably because I guess the "villain" early in the story. With a single focus on the development of numeration concepts in the constructivist students, the report would be more readable and would illustrate better the potential contributions of the hierarchies.
References


1. **Purpose**

The purpose of this study was to investigate the ways students accurately and inaccurately represent fractions on number lines and the influence of instruction on those representations. The research goal was to attempt to identify the links between students' understandings and the representations of fractions on number lines.

2. **Rationale**

This study exists as a part of a larger research project. The greater project has as its major hypothesis: "translations between and within modes of representation facilitate learning." The rationale for this study seems primarily to be for the purpose of gathering information to be utilized, among other data, for purposes of the major research project.

3. **Research Design and Procedures**

This study was based on work done within (1) an 18-week clinical teaching experiment, (2) a 30-week clinical teaching experiment, and (3) a 30-week large-group teaching experiment. The subjects in the first (1) were five fourth-grade students, three male and two female, who received four days of instruction involving the association of fractions with points on the number line and the transformations of fraction representations on the number line. This instruction occurred "near the end" of the 18-week period, with a pretest and posttest immediately before and after the four-day period. The test,
developed by Larson (1980), is composed of 16 5-option multiple choice items, including "none of the above," and has six nonindependent subscales based on (a) fraction given with representation to be chosen versus representation given with fraction to be chosen, (b) number line showing 0 to 1 versus number line showing 0 to 2, and (c) representation on number line showing unreduced fraction versus representation showing reduced fraction. Additional information was obtained through interviewing each subject doing three equivalent fraction tasks using a number line (e.g. 5/3 = ?/12, 8/6 = ?/3, 8/6 = ?/12).

The second clinical experiment (2) provided eight days of instruction, an extended version of the first instruction to eight subjects (four male and four female) in ending fourth and beginning fifth grade. When this occurred during the 30-week experimental period was not indicated, although it was done in September 1982. In addition to the Larson test, a Number Line Test was administered immediately preceding and following the instruction. The interview sessions were also conducted.

The large-group teaching experiment (3) utilized the same procedures as those in the second experiment except that there were 34 subjects (20 male and 14 female) in an actual school setting and there was no interview component. Again, it was not stated when the instruction occurred within the 30-week period, but it was conducted during September 1983. In addition, a set of 28 items, 16 from the Larson test and 12 from "other tests," was selected for analysis. These data were analyzed by a two-way analysis of variance on test administration time (T) and on the characteristics of the items: (a) L, 0-to-1 number line versus 0-to-2 number line, (b) G, fraction given versus representation given, (c) P, repartition required versus not required, and (d) I, complete and precise information given versus extraneous marks included on number line.
4. **Findings**

With respect to the Larson pre- and posttests in experiment 1, the authors state that for five of the subscales, all scores increased or remained constant. On the sixth subscale the results indicated that students were unable to choose a reduced fraction name when an unreduced equivalent form was represented on a number line. An inspection of the incorrect answers on the pretest and posttest relating to this subscale indicated that 10 of the 31 on the pretest were answered "none of the above" and that the frequency of this incorrect response increased to 28 out of 30 on the posttest. During the interview sessions, two students solved all three correctly, but only one used the number line to do so. Two others did two correctly and the other only one. The interviews revealed that the students had considerable difficulty coordinating their symbolic work with number line representations.

The results from experiment 2 showed that performance on two subscales of the Larson test showed considerable improvement on the reduced representation subscale but not on the unreduced representation subscale, which was the difficult one in experiment 1. An analysis of the errors indicated that 25 of the 47 incorrect answers were "none of the above" on the pretest and that 30 of the 38 were this incorrect response on the posttest. The mean score on the Number Line Test increased from 0.75 to 7.75 according to pre- and posttesting. The interviews revealed that students again had great difficulty trying to solve the problems on the number line. Five of the eight subjects solved the problem symbolically and only tried to use the number line when asked to by the interviewer. Also, not all students who gave a correct symbolic solution could do so with the number line.

With respect to experiment 3, unfortunately the Larson test contained a misdrawn item making the response, "none of the above," a correct one. From the remaining items no information was given concerning the incorrect response, "none of the above," on the
unreduced representation subscale. The results of the ANOVA were the findings offered for this component of the study. Posttest scores were significantly higher than pretest scores, indicating that the instruction seemed to be effective. For the subjects, 0-to-1 number lines were easier than 0-to-2 number lines, fraction-given items were easier than representation-given items, no repartition required was easier than repartition required, and the relative difficulty of the four types of items shifted from pretest to posttest.

5. Interpretations

Instruction seems to have been effective. Error pattern shifts appear to have resulted from students becoming sensitized through instruction to the characteristics of the number line. The failure of the students to recognize unreduced representations may indicate an inability to unpartition, a lack of skill in reducing fractions, or an inability to translate between modes of representation. The data suggest that the students better understood the major characteristics of the number line model as a result of the instruction.

Implications of this study suggest the following possible hypotheses:

The need to coordinate symbolic and pictorial information with the number line model poses difficulty in matching fraction names with number line representations, and

As long as partitioning and unpartitioning are difficult for children, number line representations of fractions may not be easily taught.

In addition, the results of the student responses seem to indicate that instruction needs to involve more explicit steps in the sequence of the translations within and between modes of representation.
Abstractor's Comments

The work of this study does indeed appear to have value with respect to an investigation into students' comprehension of fraction concepts and their ability to translate a fraction numerical value onto a number line and to translate the representation of a fraction from a number line to a symbol. This, coupled with the impact of particularly arranged instruction, provides the components for a meaningful study.

Considering the study as an entity in and of itself, it must be noted that the authors stated that they chose the number line as a model for representing fractions because of "its pervasive use in school mathematics instruction" -- and because of "the large role the number line plays in elementary school mathematics instruction." There is a tacit assumption here that just because the number line is used to a large extend in the elementary mathematics instruction that it is used for the teaching of fractions. It has been my experience and observation that the number line is not generally used at the fourth- and fifth-grade level in the teaching of fraction concepts and operations.

With respect to the procedure of the study, the written article does not include certain significant information, nor is it always consistent in the information it does provide. First of all, there is no indication of who does the teaching in each of the three experiment components. Is it the same person, two different persons, or three different persons? Or is it a team of persons? Secondly, the article starts out stating that there is an 18-week and then two 30-week experimental periods, but goes on to indicate that the actual time is 4 days (plus testing) or 8 days within those given time frames. It is not until much later in the write-up that the reader is advised that this study is only a part of some larger one. This means that, in fact, the three components of the study being analyzed are not 18 and 30 weeks in duration, but rather more like 6 days and 10 days. A third criticism is that while some of the data are reported in detail
and depth, much of the information given is sketchy, incomplete, and inconsistent. To substantiate this statement the following list is given:

1. The time that the fraction/number line instruction was given was "near the end" of the 18-week period during 1981-82 in experiment 1, during September 14-24, 1982 for experiment 2, and during September 1983 for experiment 3.

2. The Larson test is discussed in some detail, but there are only three sample items shown and no reference to the origin or objectives of the "Number Line Test" used with experiments 2 and 3.

3. Table 1 is entitled "Errors on Number Line Test" but with no further information about this test, the table is cryptic at best.

4. The tests given for experiment 3 are stated as being the same as for experiment 2. There were the Larson and the Number Line Test. Yet, in the description of the 28 items selected for the ANOVA for experiment three, the authors state that 16 were from Larson and 12 from "other tests." This causes the reader to wonder if there were some other tests administered besides the Larson and Number Line Test.

In consideration of the interviews, it would seem that if part of the purpose of the study is to identify links between student understandings and representations on the number line, there should have been a more complete analysis and conjecture concerning why students responded as they did. For example, in experiment 1, when a student was confronted with $8/6 = ?/3$, the student reasoned that $3 \times 2 = 6$, therefore $8 \times 2 = 16$, so the answer would be 16 thirds. The statement is made, "This appears to be an instance of a well-known but not very frequent mistake in dealing with fractions (see Bright & Harvey, 1982)." If the study is exploring links, correct or incorrect, why not more of a discussion? In the case of the fourth student described, there was a nonchalant assessment of the student response:

One student solved only the last two problems correctly (there were only three—parenthetical statement that of the reviewer). For the second problem ($8/6 = ?/3$), he reasoned that $6 - 3 = 3$ so $8 - 4 = 4$ and then marked fractions appropriately. For the third problem, he plotted $8/6$ and then created markings for twelfths. He reasoned that $6 + 6 = 12$ so $8 + 8 = 16$. For the first problem, however, he said $4/12$ as the answer, perhaps thinking that $1/3 = 4/12$. 
That is the extent of the discussion about what the child may have been thinking. The fact is, the student used addition and subtraction strategies for two of the problems, so it would seem that he would be more apt to have stated 4/12 for the first problem (5/3 = ?/12) because he reasoned that 3 + 9 = 12, therefore 5 + ? = 9. The missing number would be 4 thus causing the student answer to be "4/12."

Because of the seeming lack of focus or clarity with respect to the work of this study, the question might be posed asking if the true intent of the researchers was only "to determine the ways students represent fraction on number lines and the influence of instruction on those representations." It appeared that the use of the number line was an obstacle rather than a help to most of the subjects. Why would a teacher want to use something confusing when fractions are difficult enough to teach? There is a school of thought that fractions should not even be formally introduced into the curriculum until the upper grades because of children's conceptual difficulty with them. Is it possible that the investigators were more intent on the major hypothesis of the research project which is that "translations between and within modes of representation facilitate learning" and, because of this, were trying to determine a successful way to teach students to use the number line with fractions so that use of this mode could be better accomplished for purposes of the major hypothesis? If so, this subtle difference of intent could account for the lack of clarity and focus which tend to characterize the report of this study.

The authors, however, do indicate that this study suggests the hypotheses (stated above) that it is difficult for students to coordinate symbolic and pictorial information in using the number with fractions and that, because children have difficulty with partitioning and unpartitioning, number line representations of fractions may not be easily taught.
References


1. Purpose

The purpose of this study was to assess the content knowledge of
first-grade teachers relative to their understanding of how children
think about addition and subtraction, and to examine the teachers'
knowledge of their own students' thinking.

2. Rationale

The current reform movements in teacher education describe the
importance of pedagogical content knowledge. Recently, state
departments of education and the American Association of Colleges of
Teacher Education (AACTE) have attempted to identify the "essentials"
of teacher preparation. The investigators indicate that "pedagogical
content knowledge includes knowledge of the conceptual and procedural
knowledge that students bring to the learning of a topic, the
misconceptions about the topic that they have developed, and the
stages of understanding that they are likely to pass through in moving
from a state of having little understanding of the topic to mastery of
it" (p. 386). The importance of teacher knowledge as it relates to
the ability of teachers to assess student understanding, diagnose
learner needs, and prescribe appropriate instruction is a critical
issue in elementary mathematics education.

Earlier research on students' learning of addition and
subtraction as reflected in their solutions to different types of word
problems provided the framework for examining teacher content
knowledge in this study. The research on addition and subtraction has
been well documented. Recent research on children's thinking and
problem solving has also indicated that the child brings a great deal of knowledge to almost any learning situation (Carpenter and Peterson, 1988). Using a classification of word problems involving joining and separating actions and combining and comparing relationships, the study dealt with the following teacher pedagogical content knowledge issues:

1. The extent to which teachers know about the distinction between different addition and subtraction problem types.

2. The extent to which teachers know about the strategies children use to solve different problems which involve addition and subtraction.

3. The level of success teachers have in predicting how their students will solve different types of problems and in identifying the strategies children use to solve problems of different types.

4. The relationship between different measures of teacher pedagogical content knowledge and student achievement.

3. Research Design and Procedures

Forty first-grade teachers from the Madison, Wisconsin area were involved in this study. The teachers had no prior background in recent research in addition and subtraction. All data collection measures were administered in the spring of 1986. Teacher measures were administered individually. Student performance measures were administered to all students in their classrooms. Teacher knowledge and student performance were measured using the following procedures:

Teacher Knowledge

**Distinction between problem types** - Teachers completed two measures to determine the extent to which they could distinguish addition and subtraction problem types.
In completing the Writing Word Problems measure, the teachers were asked to write six word problems that would best represent six given number sentences which corresponded to the six join and separate problem types as discussed in the earlier research by Carpenter, Moser, and Bebout (1988). These teacher-generated word problems were coded. The teachers also completed a measure entitled the Relative Problem Difficulty test, in which they reviewed 16 pairs of word problems determining which problem, for each pair, would be more difficult for first-grade children. The problem pairs included 6 pairs of separate-result-unknown problems, 6 pairs of join-change-unknown problems, and 4 pairs of problems in which the problems were of the same type with relatively minor context changes. After the teacher subjects had responded to all 16 word problem pairs they were asked to explain their response for 5 of the pairs.

**General Knowledge of Strategies** - Teacher knowledge of the strategies used by children to solve addition and subtraction word problems was assessed by showing the teachers a videotape of three first graders solving different problems. The teachers were asked to describe how these children would solve related problems. The strategies depicted by the children on the videotape included direct modeling of the problem and its solution, use of the counting on strategy, and deriving facts based on doubles. At the conclusion of each of the three videotape vignettes, the teachers were asked to describe how the child on the tape might solve additional problems. Teachers were expected to recognize how these common strategies, often discovered informally by students, might be useful or perhaps a hindrance to the solution of the additional problems.
Teacher knowledge of Their Own Students - The project teachers were asked, in an interview setting, to demonstrate how six randomly selected students from their classes would solve different addition and subtraction word problems. These students had solved the same problems prior to the teacher interviews. Interview scoring included a score based on whether the teacher could predict whether the student would solve the problem correctly, and a score based on the teacher's success in predicting the strategy the student would use in solving a particular problem.

Student Performance

Number Facts - The first-grade students completed a 2-minute number facts test which consisted of 20 addition and subtraction facts. Ten of the problems involved sums less than 10, and ten involved sums between 10 and 18. The operations were mixed and the problems were written horizontally (4) and vertically (16).

Problem Solving - The students also completed a problem-solving test which consisted of 17 word problems. The problems were presented orally, and represented a range of addition and subtraction problem types. Four of the problems involved several operations or included extraneous information, and four involved grouping and partitioning.

Data Collection

Teacher scores on the Writing Word Problems, Relative Problem Difficulty measures, and General Knowledge of Strategies (videotape analysis) were analyzed through a comparison of means, standard deviations and range.
Similarly, means, standard deviations, and range were used to analyze the extent to which the teachers could predict whether a student could solve problems and use specific strategies in solving the problems. Additionally, the actual means and standard deviations for the addition and subtraction problem types, provided to the randomly selected students, were presented. In the Teacher Knowledge of Student Success measure, percent correct was used to analyze teacher predictions of student success in solving the six word problem types and student performance on the problems. Percent correct comparisons were also used to assess the extent to which the project teachers could predict the strategies children would use to solve different problems. This was compared with the percentage of students who actually used the strategies of direct modeling, counting on, or deriving number facts in their solution. Correlation was used to gauge the relationship between the various measures of teacher knowledge and student achievement.

4. Findings

Most teachers could distinguish between the addition and subtraction problem types as measured by the Writing Word Problems and Relative Difficulty tests. However, join-change-unknown problems were overestimated in difficulty, with fewer than half of the paired comparisons, involving this problem type, being correct. The project teachers had difficulty articulating the differences between the problem types. Many of the project teachers tended to focus on problem features like key words and the sound of the problem rather than the actual difference in problem type.

In assessing the teacher's general knowledge of student solution strategies, most teachers could recognize direct modeling strategies used by the child. The teachers were less successful in recognizing
counting strategies used by children, particularly counting back as a strategy useful in solving separate-result-unknown problems, and counting on from a larger number when the smaller number appeared first in a problem. The project teachers were similarly less successful in identifying student use of derived facts.

The teachers were quite successful in predicting whether or not randomly selected students could solve particular addition and subtraction word problems (approximate success rate = 75%). They accurately predicted the strategy the students would use to solve a problem less than 50% of the time. The teachers were more successful in predicting student success for the more common join-result-unknown and separate-result-unknown problems than for the other four problem types. The teachers were very accurate in predicting the success of their students on particular addition and subtraction word problem types. They overestimated actual performance on the separate-result-unknown example and underestimated performance on the join-change-unknown example, but were very close to actual performance with the other problem types. In predicting the solution strategies that their children would use, the teachers consistently overestimated the use of direct modeling and recall of number facts and underestimated the use of counting on and counting down as counting strategies.

There was a significant degree of relationship between the teachers' ability to predict students' success in solving different addition and subtraction problems and student performance on both the number fact and problem-solving measures. As might be expected, there was also a significant correlation between student number fact and problem-solving performance.

5. Interpretations

For the most part, the teachers were successful in identifying distinctions between addition and subtraction word problems and the strategies that children use to solve them. However, this knowledge
generally was not organized by the teachers into a coherent network that related distinctions between problems, children's solutions, and problem difficulty to one another" (p. 398). The investigators did not find this overly surprising, since such relationships have only recently been specified and available within the research community.

Measures of teacher knowledge of problems, problem difficulty, or strategies did not correlate significantly with student achievement or teacher ability to predict success in solving problems or the strategies the students would use to solve the problems. "The lack of success in identifying expected relationships may have resulted from the lack of variability on the measures of teachers' general knowledge of problems and strategies" (p. 398).

Most of the project teachers were familiar with the basic strategies that children use to solve addition and subtraction problems and could successfully identify these strategies when observed on videotape.

Teacher ability to predict student success in solving addition and subtraction word problems was significantly correlated with student performance on number fact and word problem measures. One would expect teacher ability to predict the strategies students would use to have a similar relationship with the student measures. However, the investigators indicate some problems with measuring teacher knowledge of strategies students can use. This is attributed to the fact that students may not use strategies on a consistent basis. The investigators also speculate that the lack of relationship between predicting student use of strategies and the student measures suggests that teachers may be more apt to judge problems on their perceived level of difficulty rather than strategies students may use to solve them.
Finally, the results of the study seem to indicate that teachers do not have a systematic method for planning instruction based on the processes that students actually use to solve problems. If teachers understood how the research on addition and subtraction can be used to differentiate between problem types and the solution strategies used by their students, they might approach instruction related to addition and subtraction differently.

Abstractor's Comments

Comments on the Study

Essential content knowledge for the classroom teacher is a topic which is being both "bantered around" as an issue of discussion for teacher educators, and seriously studied. This issue is worthy of serious study. This investigation examines teacher content knowledge. Using prior research in addition and subtraction problem types as the basis for analyzing teacher content knowledge, the investigators dealt with important research issues (see Rationale). The findings of the study indicate that teachers can distinguish some of the basic differences between addition and subtraction problems, but their recognition of differences between the problem types seemed to be based on guess work or intuition, rather than a true frame of reference. The study also indicated that there was a significant relationship between the teachers' ability to predict success on addition and subtraction problems and measures of student performance, but the teachers' ability to predict the strategies that students would use to solve their problems did not show a significant degree of relationship with student performance. Why? Let's examine the following:

1. It is the reviewer's opinion that most teachers have limited awareness of the research related to addition and subtraction which was the content knowledge basis for this investigation. In short, the research
community is well aware of the research efforts of Carpenter, Moser, Romberg, Bebout and others relative to addition and subtraction, but most practitioners and, sadly, far too many elementary mathematics educators are not. A cursory review of several popular elementary mathematics methods and first-grade textbooks confirms this relative lack of exposure to powerful research-based ideas. Is this why so many primary grade classrooms over-emphasize the rote learning of addition and subtraction, using only the most common problem types?

Allow me to digress. I recently had an honors graduate come back to my methods class to visit and share ideas with my current students. To my horror, one of the things she mentioned was that her first-grade textbook did not have enough ideas for her students and perhaps went too far into the addition and subtraction algorithms for her students. I knew right there that I had failed this student! We had studied the research on addition and subtraction, which formed the basis for this study, but this teacher was led only by her first-grade teacher's manual. We had examined counting strategies and patterning as powerful early childhood mathematics activities, but the teacher's edition didn't provide them. Is this typical? Unfortunately, it probably is!

2. This study also indicated that the teachers were familiar with the basic strategies used by children to solve problems, but seemed more comfortable with predicting that their students would solve problems by direct modeling or retrieval of number facts rather than using counting strategies. Once again, it is this reviewer's belief that the research community is ahead (isn't it always) of the "front lines." Counting
strategies related to acquiring and using number facts and research regarding counting and counting back (or down) is widely accepted, but still has not reached many practitioners. Teachers recognize modeling as a strategy used by many children to solve problems because they either teach using manipulatives or remember that they should, at least according to some prior professor or mathematics methods text. The investigators indicate that teachers in this study consistently overestimated the use of modeling (p. 397). Teachers need to know when and how to use modeling; it's not enough to "trot out the manips" thinking that everyone will now learn, understand, or just get the example correct!

3. The study indicates that the teachers were more successful in predicting their students' overall success on given problems than predicting the strategies they would use to solve the problems. Are such teachers superficial in their daily diagnosis of their students? Can we expect teachers to provide appropriate instruction if they are only sure that the students can or cannot complete the assignment, but not sure how they will proceed? This finding, to me, is very important. This is what I should have provided for the student mentioned above, more depth and understanding of the addition and subtraction process.

Comments on Research Design and Procedures

Reviewing the design and procedures of the study, I would agree with the investigators that the Writing Word Problems measure was probably too easy for the teachers to really be of any use, comparison- or prediction-wise (the scores were quite high, and the variability low). The examiners may consider expanding this
instrument from 6 to 18 items, asking the respondents to write three examples for each problem type. With regard to the Relative Problem Difficulty pairs, it was not clear what was done with the explanations, given by the teachers, to the 5 problems. Coding of these responses may have altered the scores reported.

The videotape-response format used in the General Knowledge of Strategies measure was very interesting. I think this instrument has tremendous potential, not only for clinical and qualitative research, but also for mathematics education classrooms. I would see such a videotape process as being very helpful in determining the level of teacher understanding of a variety of mathematics concepts.

Concluding Comments

This study is important because it attempts to address the issue of teacher content knowledge, and while the results may not have been overly encouraging to the investigators, some important teacher education "essentials" become apparent. Teachers should know about the differences between addition and subtraction problem types, and not rely, instructionally, on the more common models for addition and subtraction. Balance is needed. Teachers should recognize student solution strategies and become familiar with how students solve problems. It is likely that many teachers approach instruction involving addition and subtraction with limited knowledge of the addition and subtraction process and a lack of understanding regarding the strategies actually used by their students to solve such problems. It is hoped that additional research will investigate the issue of teacher content knowledge in elementary mathematics education. At a time when states, school districts, and higher education are questioning teacher education, we need such "essentials."

References


Abstract and comments prepared for I.M.E. by THOMAS E. ROWAN, Montgomery County Public Schools, Rockville, Maryland.

1. Purpose

To investigate children's representation of addition and subtraction word problems with open number sentences.

2. Rationale

Studies have shown that children can solve a variety of addition and subtraction problem situations with direct physical representation before they receive formal instruction on addition and subtraction. However, such strategies are limited to simple problems with small numbers. Furthermore, representing problem situations with mathematical symbols is a major goal of mathematics instruction. Standard addition and subtraction open sentences (\(a + b = c\) and \(a - b = c\)) can be used to represent only a limited range of addition and subtraction situations. Relatively few studies have investigated the effect on young students of exposure to a full range of addition and subtraction open number sentences.

3. Research Design and Procedures

Twenty-two first- and 41 second-grade pupils were randomly separated into two instructional groups, 11 in each first-grade group, 20 and 21 in the second-grade groups. One group from each grade received 15 minutes of instruction on a full range of open sentences (\(a + b = c\), \(a - b = c\), \(a + c = b\), \(a - c = b\), \(c + a = b\), \(c - a = b\)) the first day, and 15 minutes to practice and discuss the association of open sentences with word problems the second day. Additional practice time was provided each day. The second group
received the same pattern of instruction, guided practice and practice, except that it was exposed only to the standard open addition and subtraction sentences \(a + b = \square\), and \(a - b = \square\). Exactly the same problems were used in both groups.

The two groups were then tested on their ability to write number sentences to represent addition and subtraction word problems using a 12-problem test representing a full range of addition and subtraction situations. The children were asked to write a number sentence for each problem and then to solve each problem.

4. Findings

No statistical analyses were used. Observation of percentages of students who answered correctly and who wrote various types of number sentences was used. First graders who were shown a variety of open number sentences tended to use a greater variety of sentences in expressing addition and subtraction word problems. They also gave more correct solutions for most types of problems. Many first graders who were shown only standard open number sentences during instruction tried to write direct model open sentences for word problems which were not standard form.

5. Interpretations

The authors state that "The semantic structure of word problems directly influences the number sentences that young children write to represent them." They felt there was evidence of a maturing in ability to use open sentences flexibly between first and second grade. They felt that instruction which encompassed open sentences in other than a standard form would take advantage of a natural progression in children's mathematical understanding. "The fact that young children tend to represent problems directly does not imply that instruction is not important." "...a more extended period of instruction appears necessary for most first graders to master the open sentence format." Finally, the authors felt that open sentences provide symbolic representations which children can relate to informal counting and modeling strategies, and this may help with solving word problems.
Abstractor's Comments

Although this study was somewhat informally carried out and involved a relatively small number of subjects, this writer found it to be very well controlled on the points which seemed vital to its purpose. Every effort was made to control for differences in instructors. The actual instruction (experimental treatment) was minimal. The length of time was very brief. The instrument used was simple and administered in a manner which would appear to give each student an equal opportunity. The fact that the data were observed without statistical manipulation seemed very appropriate in the context of the study and the interpretations expressed by the authors.

It is an interesting piece of work which bears replication, and perhaps should eventually be used by textbook authors to make some modifications to first-grade texts.

The only negative point that this reviewer found in the paper was a somewhat vague presentation of the data which explored the relation between "number-sentence category and answer." Both the table (Table 4) and the text could have been made much easier to read and interpret. This is obviously a minor complaint which may be seen as the reviewer's failure, rather than the researchers'.

Overall, this reviewer found this to be a useful, interesting, and well-reported study.
1. Purpose

The purpose of this study was to examine the effects of student cognitive entry behavior (CEB), as assessed by a test on knowledge of concepts of numbers and numerals; prescribed mastery level (ML); and information about criterion (INF) on mathematics achievement in a mastery-learning program. "Specifically, the research question asked was whether information about the prescribed mastery criterion had differential effects for high- and low-CEB children and for 90% ML and 70% ML instruction."

2. Rationale

Educators have assumed that mastery of basic skills and concepts is a part of learning mathematics. Some of this is hierarchical; mastery of simple tasks prepares the student for learning subsequent tasks. In a mastery learning setting a criterion must be realistic, "a tricky, sticky problem residing in one of those grey areas of education." Studies cited indicate that "setting an optimal mastery criterion is related to student aptitude, subject matter content, and other cognitive and affective factors." For relatively easy content high mastery criteria were more facilitative of learning than low mastery criteria. For more challenging content low, rather than high mastery criteria, produced better results.

3. Research Design and Procedures

Thirty-six grade 3 children in a primary school class were blocked into a high-CEB and a low-CEB group, and then randomly
assigned to four treatment groups (9 students each). A subset (concepts of numbers and numeration) of the Stanford Diagnostic Mathematics Test was used as the measure of CEB (22 and above - High, 21 and below - Low). Two of the treatment groups (informed-90% ML and informed-70% ML) were told the mastery levels they were required to achieve on each learning task. The other two treatment groups (not informed-90% ML and not informed-70% ML) were not told the mastery levels they were required to attain.

The teaching program of three sequentially organized tasks was conducted in five daily sessions of 45 minutes each. The content of the tasks was place value, largest and smallest number from a group of five, and greater-than, less-than relations. At the beginning of each session, the children were shown the worksheets for the day's task on an overhead projector. Instructions were given about how to complete the task, and a demonstration of a few items was provided. The children were given as many worksheets as were required until the prescribed mastery criterion was achieved. A summative test was designed as a culmination of the three instructional tasks and required that the children apply the skills they had learned in a problem-solving context.

4. Findings

The data for this aptitude-treatment interaction study were analyzed by means of a multiple regression analysis. The null hypothesis of no significant CEB x ML x INF interaction was rejected.

"The results indicated that the effects on achievement of informing the children about the prescribed criterion depended upon both the children's CEB and the mastery level they were required to achieve. For the high-CEB children, information about the prescribed criterion enhanced achievement in the 70% ML condition but not in the 90% ML condition. For the low-CEB children, information about the prescribed criterion enhanced achievement in the 90% ML condition and not in the 70% ML condition."
"The low-CEB children in the informed-90% ML group attained a mean summative test score that was at a level similar to that attained by the high-CEB children."

5. Interpretations

"The significant CEB x ML x INF interaction indicates that information about the prescribed mastery criterion had differential effects on mathematics achievement depending on level set. Specifically, information about the prescribed mastery level enhanced mathematics achievement for the low-CEB children receiving 90% ML instruction; but diminished achievement for the high-CEB children receiving 70% ML instruction."

"The present results appear to indicate a situation in which there was a mismatch for the high-CEB children in the not informed-70% ML group. The task might have been excessively easy for these children, so that they could attain the prescribed mastery level with minimum effort. As a result, they might not have been motivated to expend maximum effort when completing the learning tasks and taking the summative test."

"In the present study it appears that the informed-90% ML treatment was specifically facilitative for the low-CEB children, whereas the other three treatments failed to compensate the aptitude disadvantage of the low-CEB children. These results indicate that if a high mastery level is prescribed for low-CEB children, information about the prescribed level is critical to achievement."

**Abstractor's Comments**

A mastering learning program should fit the students relative to their background in the content of the instructional program. This study emphasizes this point. I have several questions about this 'fit.'
The skills and content of this study seem “easy” for grade 3 students. They have seen this material many times. Why not use intrinsically newer and fresher areas of mathematics, as known by grade 3 students, for the study, rather than 'old' material no longer having motivational appeal to the students?

Another question comes in the labelling of high-CEB and low-CEB students. Should this be based on an arbitrary cut point on a scale? The investigators did mention reservations about this point.

My major comment concerns the choice of mathematics content for the study. The investigators conducted a nice, detailed study. However, the study should have been done using important material, important from the point of view of the mathematics education of grade 3 students.

Abstract and comments prepared for I.M.E. by GEORGE W. BRIGHT, University of Houston.

1. Purpose

This study examined the effects of various mastery learning teaching methods (combinations of computer-based vs. teacher-based delivery modes and initial vs. remedial instruction) on accuracy of simple algebraic computations.

2. Rationale

Mastery learning techniques, though advocated by many researchers, are not supported, either theoretically or empirically, by other prominent researchers. Some suggest that group scores improve at the expense of high-ability students who are held back and that the technique works best for material that is not heavily dependent on prerequisite knowledge. The effectiveness of mastery instruction has often been benchmarked against one-on-one tutoring, which is a very expensive technique in the context of current school organization. However, computer-based instruction, which has consistently been shown to be more effective than regular instruction, offers the potential for providing cost-effective, one-on-one instruction. Examining combinations of computer-based instruction and mastery instruction then is an important area of investigation.

3. Research Design and Procedures

Subjects were 117 students from five sections of eighth-grade mathematics. The gender split was "approximately equal" and there were "representative proportions" of minorities.
Materials. The content was computing the area of a circle. First, concepts of circle and pi were presented, with stress placed on relationships to other geometric shapes and to everyday occurrences of circles. Second, "conceptual and procedural aspects" of how to compute the area given the radius were presented. Third, calculating the area given the diameter was presented. Ten-minute quizzes of 10 problems each were given at the end of the second and third sections of the instruction. Remediation materials, for students who did not master the instruction at the 70% level, consisted of 10 problems on the following day. Students who did achieve mastery were given a basic skills assignment on the following day.

The materials were presented either by the teacher or by a computer program. Identical examples and problems were used in the two delivery modes. In the computer remediation mode, four potential error types were coded along with the quiz problems, so that if one of those four errors were entered, the student received specific feedback designed to help overcome that error. If a response was unexpected, the problem was worked correctly.

Treatments. Five treatments were used. Four of these were the combinations of teacher vs. computer format for initial vs. remedial instruction. Prior to assignment to these treatments, subjects were blocked on prior achievement (HI vs. LO). The fifth treatment was a non-mastery, traditional teacher approach.

Measures. After the first day, subjects were given a 10-item quiz (reliability = .77). This quiz was used to determine the need for remediation. At the end of the second day, a mastery posttest (20 multiple-choice items, reliability = .96) was given. A 20-item attitude survey was also given.

4. Findings

One analysis was on mastery versus non-mastery. The mastery group scored higher (p < .0005), with an effect size of 0.6, with HI students outperforming LO students by about the same amount in both mastery and non-mastery conditions.
Within the mastery groups, neither computer nor teacher methods influenced performance within either initial or remedial instruction. There was, however, an interaction \((p < .05)\). The groups that received both computer and teacher presentations across initial and remedial instruction outperformed the groups that received the same type of presentation for both instructional periods. Again, HI students outperformed LO students.

The only attitude difference was on initial instruction. Students in the computer groups had a higher attitude \((p < .05)\) than those in the teacher led groups.

5. Interpretations

The mastery groups outperformed the non-mastery groups, apparently without putting the HI students at any disadvantage. Further, the importance of varying instruction seems clearly supported. This suggests that it is more appropriate for researchers to identify effective combinations of instructional methods rather than to look for a single "best" method.

The lack of attitude effects was somewhat surprising. It was perhaps related to the fact that students in this study were already computer-experienced. If so, the frequently encountered novelty effect was not operative.

Abstractor's Comments

This is a very clean study with a good theoretical rationale and clear direct analysis. The results are thus quite believable in terms of what the researchers set out to do.

There are two primary concerns. First, that the instruction lasted for only one day raises considerable question about the application of the results to longer periods of instruction. For example, would the same interaction effect have been found for a week of initial instruction followed by three days of remediation?
instruction? It is of course impossible to investigate all possible combinations of times, but given that classroom teachers are faced with planning for months at a time, research built around a few minutes of instruction does not seem very realistic.

Second, the mathematics topic seems of marginal importance. It is important, of course, to view the content as an instance of learning a procedure; procedures are important in mathematics. But the authors seemed to conceptualize it as an instance of "algebra" instead. I personally doubt that one day's instruction can do much toward developing an understanding of "algebraic computations," even simple ones. Learning to use an area formula does not seem to me to be central to algebra.

The most important implication of this work for me is its support of the notion that variation of instructional delivery is important. Like the authors, I am convinced that technology is primarily useful not as the sole medium of instruction but rather as it is integrated with other instructional techniques (e.g., manipulatives, cooperative groups) that are known to be effective. It is critical to view technology as one tool for communication of ideas; there are many other equally valuable tools.

In trying to interpret this research for teachers, however, I am left with several unaddressed (much less unanswered) questions: If combinations of instructional techniques are more effective than any one technique, does it really matter what techniques are combined? Is there anything about technology that makes combining it with mastery learning techniques particularly effective? My guess is that the careful planning of instruction by these researchers is at least as important as the particular formats within which the instruction was delivered, but I am not sure. I would have appreciated more discussion of this point. The conclusion that "traditional and computer-based delivery systems...are of greatest value when complementing one another" (p. 32) is too simplistic; clearer specification of the terms, traditional and computer-based, is needed for proper interpretation and application of the conclusion.

Abstract and comments prepared for I.M.E. by DAIYO SAWADA, University of Alberta, Edmonton, Canada.

1. Purpose

The study is an experiment to determine the effects of two different methods of teaching problem-solving processes using microcomputers.

2. Rationale

During the last decade, there has been a focus on the process aspect of mathematics, particularly as encountered in problem solving. With the use of appropriate software, the microcomputer has provided new avenues for the open-ended exploration of mathematical processes. This study compared and contrasted two different methods of using microcomputers to foster problem-solving strategies in order to understand more clearly how microcomputers could help students explore mathematical processes.

3. Research Design and Procedure

The study employed a typical pretest-posttest design using two groups of 13-year-old students, with each student assigned non-randomly to one of two treatments each lasting for a duration of two-and-a-half hours spread over nine days.

a. Treatments

In Treatment I, 12 students (reduced to 9 due to absences) were assigned in pairs to work directly with microcomputers at their own pace through a sequence of computer-generated problem situations, taking turns operating the computer. They were encouraged to discuss
and plan their actions together but to record their work individually. The teacher was primarily a facilitator, providing initial instructions, giving additional assistance as requested, and loading new software as appropriate.

In Treatment II, 12 students were grouped around a single computer with a large monitor close to a chalkboard with both teacher and students operating the computer. The teacher demonstrated the software-generated problem situations and adopted a very open non-judgmental attitude to encourage class participation and the free generation of conjectures. The students worked in small groups of two or three, with only a single student doing the recording on paper.

Students were assigned to treatments in pre-existing friendship groups in order to enhance group interaction during instruction. The content objectives (functions and relations) and the process objectives (using trial and error, looking for patterns, making and testing hypotheses, using symbols to express relations) were the same for each treatment and were clearly defined in the study.

b. Hypotheses

No hypotheses were specified.

c. Instruments

Based on the content objectives, two "equivalent test forms" were administered, one as a pretest and the other as a posttest. No information was provided concerning the psychometric properties of the tests. In addition, the written work of students was examined and informal analyses were made of video recordings of the instruction.

4. Findings

Although the design of the study was an experiment, no tests of statistical significance were performed. Instead, descriptive statistics were presented showing that the pretest and posttest means for the two groups were almost identical (52.6 and 57.8 for treatment I and 52.8 and 57.5 for treatment II). These scores "show that both groups made, on average, very small overall gains" (p. 814).
5. **Interpretations**

As shown by the gain scores, "the two modes of computer-assisted instruction (CAI) employed seem to have had negligible overall effect on student achievement" (p. 814). Faced with these unanticipated results, the researcher devoted his discussion and most of his conclusion analyzing and conjecturing as to the possible causes of the experimental failure. In doing so, he presented a rather exhaustive and cogent analysis of the likely sources of invalidity that plagued the study (e.g., unreliability of the tests, very small sample size, extremely short duration of treatments, decline in student motivation, failure to match tests to treatments, etc.).

**Abstractor's Comments**

This study presents us with another example of an instructional experiment which fails to show any significant differences. This is a rather common phenomenon in educational research and has given rise to a rather voluminous commentary, much of it expressed with a very guarded conservatism (see, for example, Tobias, 1982). Ernest's study adds to this growing commentary. However, in choosing to devote all of his discussion and most of his conclusion to the intricacies of internal and external validity with little focus on substantive matters, the study contributes little to our understanding of mathematical processes. This raises the question as to what warranted both its initial publication and then its selection for inclusion in IME?

The study most certainly is not a model of a well-designed experiment: on the contrary, and upon its own analysis, it is plagued with design problems as is obvious in the following points:

a. The study has the form of, and claims to be, an experiment and yet enunciates no hypotheses. Furthermore, although treatments were specified and implemented, no theoretical framework was presented in support of the treatments, or to indicate why they might be expected to produce differences.
b. As an experiment that failed to show positive results, the study cannot even be said to have produced 'nsd' (nonsignificant differences) because, despite being an experiment, no inferential tests of significance were done.

c. The pre-post achievement tests used are presented as valid and equivalent. However, no psychometric evidence is presented to support either claim. Indeed, the researchers 'post-mortem' analysis suggests that the tests are lacking on both counts.

d. The very small sample size is constraining enough, but when combined with the nonrandom assignment of cohort groupings to treatments, any treatment differences become difficult to interpret at best, and completely confounded at worst. Perhaps it was fortunate that treatment differences were not detected.

e. The very short duration of the latments (150 minutes) totally ignores (shows no acknowledgement of) the classroom as a sociocultural entity with identity, stability, and integrity of its own.

This listing could be easily extended. Again, what then is significant about this study? I personally believe that it was fortunate that this study was published, and more fortunate that it was selected for review in IME, and even more fortunate that I had the chance to review it. I say this because it presents an opportunity to examine more critically how studies of this kind are received and interpreted. To clarify this point, let me begin with the interpretation that the researcher himself gives to his negative findings (p. 816):

These results do not justify an assertion that the forms of teaching, organization and modes of microcomputer use are inappropriate for attaining the aims of the study. A more valid conclusion is that the specific applications of the forms of teaching, organization and modes of microcomputer use employed in the experiment proved ineffective relative to the measures of achievement adopted.
This is a very eloquent way of expressing the classic dilemma of a scientist who has not been able to supply experimental support for his theory: he always chooses to demonstrate that it was the experiment itself that was in error, not the theory. I suggest that, in this statement (and in this study in general) we have a textbook example of what Hempel (1966:28) has called the search for "ad hoc auxiliary hypotheses" to save the central (and unexamined) tenets underlying an experiment, in this case that the forms of teaching implicit in this study and the modes of computer use (CAI) that they gave rise to are valid: it can't be that CAI is wanting; it must be that unanticipated experimental defects invalidated the study. In this way, the central beliefs undergirding the experiment are always shielded from negative evidence; the fundamental presumptions concerning teaching are immune from experimental refutation.

The researcher therefore, in order to maintain his fundamental pedagogical beliefs, was driven to show that his negative results were likely due to defects in his experiment. He was most successful in this endeavor. However, this success rendered his study impotent in furthering our understanding of CAI as a way of facilitating the development of mathematical processes. This is a serious, long-standing problem. It is a variation of what is known in the philosophy of science as the Quine-Duhem Thesis (Quine & Ullian, 1970). In these days of the quantitative-qualitative controversy in educational research, it is perhaps both wise and timely to recognize the operation of the Quine-Duhem Thesis in the interpretation of 'nsd' findings, rather than sweeping them into the dust pan of experimental ad hoc hypotheses.

What meta-conclusions might be drawn from this situation? Some advice from Hempel (1966:28) may be appropriate: "But science is not interested in thus protecting its hypotheses or theories at all costs - and for good reason." In line with Hempel's remark, I suggest that "for good reason" the fundamental pedagogical beliefs underlying studies of this kind need to be seriously critiqued and alternatives constructed. More particularly, I suggest that the following positions need serious study:
a. That the fundamental presumptions concerning teaching that undergird this study are essentially invalid.

b. That the mode of inquiry accepted in this study is inappropriate to the phenomenon investigated.

c. That the epistemological stance adopted in this study is incongruent with the study of living things.

I am sure that these meta-conclusions are controversial and need to be elaborated and critically examined in order to be properly understood. Indeed, these statements may be no more valid than the presumptions underlying this study. This, however, is not the issue here. I only wish to draw our attention to how routinely we use faulty experimental design as the scapegoat for our own failure to understand more profoundly what we are doing. The good news is that we do have an alternative: to reexamine our fundamental postures in regard to teaching and learning and in so doing to orient ourselves in perhaps profoundly different ways to education and to inquiry in education. Some alternatives can be found in Doll (1989), Sawada (1986, 1989), and in many others.

References


Abstract and comments prepared for I.M.E. by JOHN GREGORY, University of Florida.

1. Purpose

The purpose of the study was to analyze key factors associated with mathematics performance of women and their selection of quantitative fields of study. Specifically the study sought to determine:

a. if a pattern of differences that had been found for Graduate Record Examination (GRE) mathematics performance of women who had majored in quantitative fields existed for performance on the quantitative portions of the Scholastic Aptitude Test (SAT-M) of women intending to major in different quantitative fields.

b. if measures found to be associated with women who completed degrees in quantitative fields were also influential for women intending to major in quantitative fields.

2. Rationale

The investigation stemmed from the bases of: a) previous investigations have identified factors influencing women to choose between quantitative and nonquantitative fields; b) other research has found differences in measures existent between men and women who have chosen to major in quantitative fields of study. In an exploratory study the investigator had found differences between women majoring in different quantitative fields. The two category classification of quantitative fields were majors in engineering or physical sciences, and majors in mathematics, statistics, or computer science. The question of the reported study was simply whether or not these same differences existed for women at the stage of their career when they were selecting their intended major in either of these two fields.
3. Research Design and Procedures

A sample (n unreported) of the population of college-bound high school seniors who took the SAT in 1982-83, and who responded to the Student Descriptive Questionnaire (SDQ) when registering for the SAT, was divided into seven groups according to identified intended undergraduate majors of (a) mathematics, (b) statistics, physics, computer science; (c) engineering, (d) other physical sciences, (e) biological sciences, (f) social sciences, and (g) humanities. The investigator presents this grouping scheme and sequence as an indication of "general level of mathematics required" in the undergraduate curricula.

A preliminary analysis of the mean SAT-M scores of the sample was achieved via a "median polish (Tukey, 1977)" technique that decomposes the data into "a common value, row effects, column effects and residuals" which leads to the development of hypotheses to be subsequently tested by more rigorous statistical methods.

A proposed causal model explaining SAT-M performance entered variables found in previous investigations to be related to women's selection of quantitative fields of study and mathematics performance. The 15 variables entered were: race, each parent's level of education, family income, number of years and level of high school math/science courses, level of extracurricular involvement, average year-end or midyear grade in math/science, self-ratings of math/science abilities and leadership abilities, high school rank, and intended major (clustered as identified above). The model was estimated using the mean performance SAT-M scores for 314 women who indicated their intentions to major in one of two categories of quantitative fields: engineering and physical science (n=95) or mathematics statistics, physics and computer science (n=219).
4. Findings

When polished across gender and intended undergraduate major, a substantial gender effect (favoring men) on SAT-M performance was found. A positive effect for mathematically related majors was found to exist except for the statistics, physics, and computer science group.

The pattern of residuals for this sample of intended majors was consistent with those that had been found previously for GRE scores of college seniors. Women intending to major in engineering and the physical sciences had scores higher than would be anticipated from the overall data.

For the two groups of women intending to major in quantitative fields, each of the variables in the model was tested separately, finding no differential effects on SAT-M performance. This suggests that the effects of influential variables in the model were the same for women regardless of their group membership.

In the entire model, the variables explained 17.9% of the variance in choice of one of the two intended major groups, which the investigator reports as twice as high as percentages from previous studies. For SAT-M performance, 58.7% of the variance is explained by the model. Of the variables having a significant effect upon SAT-M scores for these women, self-rating of leadership ability was the only one found to be negative. Nonsignificant variables included mother's education and undergraduate major group. Mathematics and science self-ratings had the largest direct and total effects on SAT-M performance in the model.

Variables differentiating between the two groups of women included parental income and years of science studied in high school. Mathematics and science self-rating again was the most influential variable. Each of the variables exhibiting influence on the undergraduate major selection had significant effects on SAT-M performance as well.
5. **Interpretations**

The investigator interprets the results as showing that women who intend to major in engineering and the physical sciences exhibit the same patterns/qualities as women who completed majors in these fields (as per previous investigations). As found in studies of women in quantitative fields versus nonquantitative fields, this study found that taking more science courses, having higher self-ratings in math/science, and having higher indices of family background differentiate women in the two groups within quantitative fields.

Self-rating in math and science, as the strongest variable in both selection of major and SAT-M performance for the sample, emphasizes the importance the investigator sees for mathematically capable women perceiving themselves as having good mathematical and scientific abilities -- referred to as "appropriate attributions" in a related research domain. The investigator also suggests that the results support the use of intervention strategies for developing positive attitudes of women toward the study of mathematics and science beginning in the middle school years.

**Abstractor's Comments**

This investigation once again supports the body of research showing that, although capable, women are reluctant to choose mathematically related fields in college. There are two significant findings, in the view of this abstractor, offered by this investigation: (1) the influential variables on selection of undergraduate major are social, and (2) the influence takes effect even prior to achievement of any level of success needed to attain college senior status within a mathematically-related field. Probably what is being evidenced is a "contraryness" of certain women who (thank goodness) are so strong in their conviction to prove society wrong (i.e., that women should be in scientific fields), that the conviction must begin to develop earlier than might have been thought.
before. We all need to concur with the investigator that intervention strategies really do need to be in place earlier than high school in current curricular practice.

The investigation does raise some questions, too. For example, do the variables having effect upon selection of major within engineering and physical sciences versus math/stat or computer science have anything to do with perceived differential levels of income generated from employment in either of the two fields? Family income was found to be influential in the choice of major. Do mathematically capable women perceive the opportunity to more easily maintain their usual level of economic condition by being engineers instead of mathematicians?

The National Council of Teachers of Mathematics has recently released a description of standards to be sought from mathematics instruction in K-12 education. As these standards are implemented, will more minorities, including women, be encouraged to enter quantitative fields? The concept of mathematics currently held by all students will certainly change if the standards are adopted. No longer will mathematics be thought of as the subject in which you merely memorize and grind out. Problem solving, applications, and modeling of the world as emphasized by the standards will potentially "socialize" mathematics, thus making it at least look more like the majors typically sought by women.

More specific to the study, might the statistical differences between the two women's groups be due more to sample sizes than actual effects? Granted, each variable was found to have equal variance for the two groups prior to statistical testing of the entire model, but the favored group of engineering and physical science majors had only half as many subjects as the math/stat and computer science group. Is the application of the Tukey "median polish" as sensitive to unequal n as some other statistics have been found to be, thus constituting effects that are not real?
And lastly, not only have there been differences found to exist between men and women in quantitative fields, but differences exist between men in quantitative versus nonquantitative majors. If one were to compare males selecting engineering and physical science versus males selecting math/stat and computer science, would the same patterns of effects be found as was reported for the women in this study? In other words, does a gender gap continue to exist once a student has selected a major within a quantitative field of study? An answer to this question might suggest searching for variables that are not as closely related to gender as the investigator wishes to have us believe.

Reference

1. **Purpose**

The purpose of this study was to investigate differences between skilled and unskilled mental calculators in (a) the strategies they use to calculate products of multidigit whole numbers mentally, (b) their recall of basic multiplication facts, (c) their knowledge of whole-number products beyond basic facts, and (d) their short-term memory capacity.

2. **Rationale**

Mental calculations are useful not only in everyday activities like estimating, but also in enhancing students' general number sense. Yet, NAEP results show widespread difficulties in performing such simple mental calculations as $90 \times 70$. Previous research suggests that mental calculative skill is a function of the strategies employed, the knowledge of useful numerical equivalents, and the capacity to process numbers. A better understanding of which strategies and numerical equivalents are most useful can help teachers improve students' mental calculative skills.

3. **Research Design and Procedures**

A screening test consisting of 20 mental multiplication tasks -- 10 easy ($30 \times 200, 8 \times 99$) and 10 difficult ($32 \times 64, 24 \times 24$) -- was administered to 286 mathematics students in Grades 11 and 12. The test was group-administered by audiotape, with one item presented every 20 seconds. Students were instructed to use pencils only to
record their answers and not as a calculative aid. Responses were scored right or wrong, and scores were used to identify the 15 most skilled and 15 least skilled mental calculators -- the sample for the study.

A probing test of 30 mental multiplication tasks was individually administered to each of the 30 students in the sample. For each task the factors were read orally and students were asked to calculate the product and explain the method they used to get the answer. Interviews were audiotaped. The 900 responses to items on the probing test were categorized according to the dominant strategy used, regardless of whether the answer was right or wrong.

Each student in the sample was also given an oral test of the 100 basic multiplication facts, as well as the forward and backward digit-span subtests of the Wechsler Adult Intelligence Scale.

4. Findings

Analysis of the interview tapes led to the identification of 4 general methods of solution, comprising 12 specific strategies, as follows:

a. Pencil-and-paper mental analogue

Both partial products calculated
One partial product obtained by recall
Both partial products obtained by recall
Stacking (a special strategy applied to repeated-digit factors)
b. Distribution

Additive distribution
Fractional distribution
Subtractive distribution
Quadratic distribution

c. Factoring

General factoring
Half-and-double
Aliquot parts
Exponential factoring

d. Retrieval of a numerical equivalent

Unskilled students used a paper-and-pencil analogue 86% of the time and distribution about 12% of the time. Skilled students, on the other hand, used a pencil-and-paper analogue only 22% of the time, distribution 54%, factoring 14%, and straight recall 10% of the time.

Mean scores on the basic facts test were 99.9 (SD = 0.35) for the skilled students and 96.7 (SD = 3.1) for the unskilled. Skilled students outscored unskilled on both the forward digit span (7.8 vs. 6.3) and backward digit span (6.2 vs. 4.8). The correlation between scores on the screening test and each measure of memory capacity was +.3 (p < .01).

5. Interpretations

The findings suggest that efficient strategies for mental calculations may be those that (a) do not require "carrying," (b) proceed from left to right, and (c) incorporate interim calculations into a single result. Skilled mental calculators rarely used
strategies that required carrying, often performed calculations in a left-to-right fashion (beginning with the most significant digits first), and kept running totals as they went along.

Mastery of basic facts does not seem to contribute substantially to differences in performance on mental calculations, since both skilled and unskilled calculators showed high levels of mastery of basic facts. However, knowledge of combinations beyond basic facts may be a factor: Skilled students solved 44 problems by recalling rather than calculating a product, while unskilled students solved only 2 problems by recall. Most of the recalled combinations were squares of numbers that students said they had learned outside of the mathematics classroom.

Although the correlation between performance and short-term memory was weak in this study, memory seemed to be a factor, especially in the forgetting of interim calculations.

Abstractor's Comments

The current curricular emphasis on enhancing estimation and other mental computational skills makes this study an especially timely one. Its focus on products of multidigit whole numbers seems appropriate, given the relative frequency of multiplicative mental computations and the opportunity that multiplication provides for identifying varied computational strategies.

One thing that would have been helpful to include in the report is a listing of the 30 items used in the probing test. Pilot data probably suggested some of the strategies that students would use to complete the multiplication tasks, and the authors may have counterbalanced the items according to "likely" strategy. In any case, the nature of the items would surely affect the frequencies of strategy use, although it would presumably not change in any
fundamental way the unskilled calculators' strong preference for the pencil-and-paper mental analogue versus the skilled calculators' facility in using a variety of strategies.

It would also have been interesting to see some item and/or student contrasts. For example, were there particular items for which the differences between strategies used by skilled and unskilled mental calculators were noteworthy? Were there particular students for whom the choice of strategy was surprising or highly varied? A contrast of skilled versus unskilled calculators' strategy preferences for some typical items (or perhaps for some categories of items) would have helped to satisfy this reader's curiosity about the effect that item type had on choice of strategy.

As the authors point out, this study focused on efficient versus inefficient students and thus did not permit the analysis of efficient strategies per se. Nevertheless, the authors conjecture that efficient strategies "may be those that (a) eliminate the need for a carry operation, (b) proceed in a left-to-right manner, and (c) progressively incorporate each interim calculation into a single result" (p. 108). It would be interesting, and especially helpful to teachers, to conduct some follow-up studies that focus on the strategies themselves. For instance, are the strategies that seem efficient for the skilled calculators also efficient for unskilled calculators (once they have been taught the strategies)? Or, must unskilled calculators rely on a pencil-and-paper analogue in order to succeed at all with mental computations?

Studies that focus on the efficiency of strategies would necessarily have to take into account the accuracy of answers, perhaps disregarding errors in basic facts, since these seem to be rather negligible. Even in the present study (where accuracy was not pertinent to the identification of strategies), it would have been informative to see some analysis of performance vis-a-vis the
strategies used, or at least a report of whether students who were categorized as skilled/unskilled on the screening test performed similarly on the probing test.

Two small suggestions, in closing: (a) It seems as though a student's learning modality (in this case, visual versus aural) could be a factor in mental calculations; the unskilled students' propensity for writing in the air certainly calls into question the effect that the mode of presenting the task (written versus oral) has upon performance; and (b) the logistical problem that the authors allude to with respect to students' using pencils to record answers, but not to calculate, could be circumvented by using computers to administer the screening test.
1. Purpose

The purpose of this study was to investigate the extent to which differences in motivation profiles prompt differences in children's performance in CAI settings. To this end, two research questions were examined: (1) Do intrinsically motivated students conceive learning with the new technology as a challenging task and thus use it more efficiently than extrinsically oriented children? (2) Do students who possess self-criteria for success and failure perform better in CAI settings than students who are dependent on external feedback?

2. Rationale

This study investigated these questions within the framework of Harter's model (1981) of intrinsic orientation, in which children's intrinsic motivation for learning in the classroom is determined by two orientations: the motivational component composed of curiosity, preference for challenge, and tendency to gain mastery; and the cognitive informational component constructed of independent judgment and self-criteria for success and failure.

The author hypothesized that the CAI system would exert a greater effect on the intrinsically motivated students (because of their tendency to work on their own) than on those who possess an extrinsic motivational orientation (because of their tendency to rely on external guidance). It was further hypothesized that the cognitive informational orientation would not be related to achievement at the computer, since self-criteria for success and failure are important for learning in a traditional classroom and may be less important in CAI settings.
3. Research Design and Procedures

The subjects for this study were 257 fourth- through sixth-grade pupils in two Israeli schools which had been using CAI since 1977. Pupils received two 20-minute sessions of a CAI program called TOAM (the Hebrew acronym of computer-assisted testing and practice) and three 45-minute sessions of traditional mathematics instruction per week. The CAI program presented students with a random mixture of problems from all topics of elementary school arithmetic.

The independent variable of intrinsic-extrinsic orientation was measured at the beginning of the school year with the self-report scale developed by Harter (1981). The scale contained items addressing the following five motivational factors: curiosity, challenge, mastery, judgment, and self-criteria for success and failure. The scores on the first three subscales were combined to give the motivational orientation component and the scores on the latter two were combined to give the cognitive informational orientation component.

The dependent variable of mathematics achievement was composed of two parts, the Arithmetic Achievement Test (AAT; Israeli Ministry of Education, 1976) and the TOAM Computer Testing Procedure (TCT; Osin, 1981). These were administered at the end of the school year. Both tests covered the same content, but differed in: (1) mode of presentation, (2) quality of feedback, and (3) rate of success.

To analyze the data, the author used a $2 \times 2 \times 3$ multivariate analysis of covariance (MANCOVA) with two dependent variables, the AAT and TCT scores, which was followed by a series of univariate ANCOVA to determine the effects of the intrinsic orientation profiles on these variables. The between-subjects factors included two levels of...
motivational orientation (above and below the mean), two levels of
cognitive informational orientation (above and below the mean), and
three grade levels. Standardized mathematics achievement scores from
the beginning of the year were used as a covariate in all analyses.

Main effects were found to be significant for motivational
component (p < .006) and grade level (p < .001). The follow-up ANCOVA
found that while the achievement levels of the intrinsically motivated
pupils were higher than those of the extrinsically motivated subjects,
the difference was significant (p < .005) only on the TCT. Although
the MANCOVA cognitive informational component main effect was not
statistically significant (p > .05), the author did a follow-up ANCOVA
which found a marginally significant (p < .06) main effect on the AAT,
possibly indicating that the intrinsically cognitive oriented pupils
tended to score higher on the AAT than the extrinsically cognitive
oriented pupils. No other interactions were found to be statistically
significant.

5. Interpreations

The author of this study made several summarizing statements,
which support the hypotheses based on Harter's model. One of these
was that the intrinsically motivated students performed better than
the extrinsically motivated students at the computer but not on the
paper-and-pencil achievement test. Furthermore, pupils who were not
dependent on the teacher's judgment and who had self-criteria for
success and failure did better on the paper-and-pencil test, but not
at the computer, than the extrinsically cognitive oriented children.
Students who had intrinsic orientations on both components gained the
highest scores on both measures.
This study should be of interest and importance to educators who utilize computer-assisted instruction in their classrooms. As more computers are finding their way into schools, concern about who might or might not benefit the most from their use is an important question. Certainly differences in students' motivational profiles would seem to affect differences in achievements in CAI settings, where the student must assume more responsibility for his/her own performance.

The investigator in this study carefully carried out the statistical analysis and adequately reported the pertinent data. It was a technically sound study.

As I studied the article and carefully considered various aspects of it, several questions and concerns became evident. I will briefly address each of them. First, the author appears to have been rather vague in his/her description of parts of the study. The length of the study was never explicitly stated, which may not have been necessary due to the nature of the investigation. The subjects were given no different instruction than they normally would have received and have received since first grade. The traditional classroom instruction was not described or discussed. The fact that the intrinsic-extrinsic orientation instrument was given at the beginning of the school year and the achievement tests at the end of the year is probably irrelevant. The author also neglected to provide information about the students in the study, with respect to their socioeconomic status. Referred to in the article were studies which have shown a correlation between such status and achievement and motivation profiles, but these were not then related to the present study. Lastly, the author reported the KR(21) reliability coefficients for the AAT and TCT for all three grade levels, but did not indicate if these values were based on the scores for the subjects in the current study. The Osin reference for the TCT test seems to indicate that it has been previously used in disadvantaged schools.
I realize length is always a limiting factor in the reporting of a study. However, my understanding of the research and its implications would have benefitted from more description of the intrinsic vs. extrinsic motivation of Harter's model. At times in the article, the terminology and labeling of types cause the reader to get bogged down and miss the salient points of the findings. A reader who has a firm grasp of Harter's model will draw the most information from this article.

In analyzing the data, the author chose to form the two levels of motivational orientation and cognitive informational orientation by those scores above and those below the mean. Another possibility would have been to form those groups as the top 1/3 of the scores and the bottom 1/3 of the scores, thus eliminating the middle third. Although this reduces the number of subjects in the analysis, it focuses the study on students who exhibit strong tendencies toward the extremes of the orientations.

I have one question regarding the author's presentation of the statistics for the study. One table in the article contains the multivariate and univariate analysis of covariance F-values. While six of these are significant at or beyond the .05 level, six of them are also reported as being < 1.00. I would like the author to have reported what these F-values actually were. An F-value close to 1.00 is to be expected if there is little difference between the groups. However, an F ratio much less than 1.00 can serve as a signal that something may be amiss in the experimental situation itself. It can indicate that another unanalyzed factor may be affecting the results of the experiment. By having the F-values reported the reader would be given a clearer indication if this could have happened.

My major concern with this study questions the appropriateness of the emphasis it places on "...Learning Mathematics in CAI Settings," as the title indicates. Was mathematics achievement of learning
really investigated? It appears that mathematics was only incidentally the content area utilized. Could it be that the variable of interest was the students' ability to perform in a paper-and-pencil setting vs. a computer setting? Assessment measures and the effect a student's intrinsic orientation profile has on his/her performance were the focus of the study. If, in fact, mathematics "learning" were the focus of the study, some discussion should have addressed the effect of differing testing modes on outcome measures and how that affects our understanding of student achievement. It would seem that the present study needs to replicated in other content areas. Would similar results occur in science, geography, or reading?

References


Abstract and comments prepared for I.M.E. by JUDITH SOWDER, San Diego State University.

1. Purpose

This research was designed to examine whether or not word problems that were personalized with information gathered from a biographical questionnaire would affect learning and attitudes.

2. Rationale

Research has shown that people learn by relating information to what they already know. Student difficulties with word problems seem to be due primarily to lack of understanding of the problems and inability to translate the problem into a mathematical expression. Since concrete problems describing realistic situations are easier to solve than abstract problems, it would seem that problems that have been personalized would increase task motivation and improve comprehension because information in the problem would be easier to interpret.

3. Research Design and Procedures

Study 1: Seventy-two fifth and sixth graders were randomly assigned to three treatments in which personalized, concrete, or abstract contexts were used in word problems. Examples of the three types of problems taken from the article are:

Abstract context: There are 4 quantities of fluid to pour into containers. We pour 2/3 of a quantity into each container. How many containers can we fill with the 4 quantities of fluid?
Concrete Context: Mike had 4 bottles of juice to pour into cups. He pours 2/3 of a bottle of juice into each cup. How many cups can Mike fill with the 4 bottles of juice?

Personalized Context*: Joe, Chris, and other friends visit Steve on a weekend. Steve has 4 bottles of cola in his refrigerator. He pours 2/3 of a bottle of cola into each cup. How many cups can Steve fill with 4 bottles of cola?

*Italicized word indicate personalized referents that were varied for each student.

Students completed a biographical questionnaire that included such items as homeroom teacher's name, birthdate, household pets, favorite food, and friends' names. In the example of personalized context given above, the italicized words varied according to the biographical information provided by the individual working the problem.

The instructional unit was based on the topic of dividing integers by fractions. The 45-minute lesson, presented on a computer, began with some instruction on using the computer, a review of prerequisite mathematics facts and terminology, and five problems demonstrating a four-step solution process. All five problems were presented in the context of the assigned treatment, but all numerical values and types of measurement units were constant across treatments.

Students then took an 11-item achievement test which included a six-problem context subtest, with two problems from each of the three contexts, all of which called for dividing an integer by a fraction. There was also a three-problem transfer subtest: One numerical problem required dividing an integer by a fraction and two verbal
problems; another required dividing a fraction by an integer; and the third required dividing a fraction by a fraction. The final two-item subtest called for recognition of the rule statement and the four-step procedure taught in the instructional unit. Students were also given an attitude questionnaire to measure reactions to the tasks.

**Study 2:** The design and procedures of Study 1 were replicated, except that instead of having students work individually with a computer, the computer was used to generate print versions of the lesson.

4. **Findings**

**Study 1:** The authors used multivariate analysis of variance to compare context treatment on the three achievement subtests (context, transfer, recognition) and attitude. Only the context treatment main effect was significant. The personalized-context group performed significantly better than the other groups on transfer items with fractions as divisors. Attitude scores were significantly higher for the personalized-context group than for the concrete-context group, but the abstract-context group did not differ from either. No gender differences were found. Aptitude-treatment interaction effects were found between total achievement scores and the California Achievement Tests on mathematics and on reading. High achievers performed well with either personalized or concrete contexts, while low achievers performed better with personalized context than with the other contexts.

**Study 2:** This time the personalized-context group surpassed both other groups on the context and transfer subtests, and surpassed the concrete-context group on the recognition subtest. The attitude questionnaire results showed that the personalized-context group found the problems more understandable.
Effect sizes for the combined studies indicated that the impact of the personalized context was substantial. A quantitative analysis on open-ended reactions showed that the mean for the personalized-context group was higher than the means for the other groups.

5. **Interpretations**

Presenting problems in familiar contexts made problems more interesting and understandable for students in either presentation mode. The authors noted, however, that individualized lessons require extra time and effort to prepare. Moreover, the novelty effect of personalized problems might diminish with time. With these limitations in mind, the authors presented six suggestions, varying in sophistication and complexity, in which personalized problems can be used in classroom settings: Problems can be personalized in tutoring sessions; problems can be embedded in themes of common interest to a class; students can construct examples; problems can be personalized with a word processor; a programming language can be used to personalize problems; a data-base program can be developed.

**Abstractor's Comments**

Although teachers have believed for many years that personalizing problems makes the problems more interesting to students, there has been little research to support this belief. This well-designed study gives strong support for personalizing problems.

I was surprised to see so few differences between the concrete and abstract treatments, since previous work (e.g., Carraher, Carraher, & Schliemann, 1987) has shown concrete problems to be easier to solve than abstract problems. However, abstract problems in other studies are usually numerical only. One can see that the abstract and concrete problems here had little to distinguish them from one another.
My major dissatisfaction with the study has little to do with the design, outcomes, or interpretations. Rather, I was disturbed by the instructional unit itself. It was extremely rule-oriented and contrary to the spirit of the new Curriculum and Evaluation Standards for School Mathematics (NCTM, 1989). Even though the Standards may not have been available, there has been some effort on the part of researchers, both within and outside of mathematics education, to move away from basing research on instruction that is antithetical to what we advocate should be happening in classrooms. Research based on poor instruction makes results difficult to interpret to teachers who are attempting to move away from rule-based instruction. This is less true in the present study than in some others, but I urge the investigators to consider this point when designing future studies on mathematics learning.

References


Abstract and comments prepared for I.M.E. by THOMAS C. GIBNEY, The University of Toledo.

1. Purpose

This study investigated the influence of cultural practices in Brazil on the mathematical understandings of 10- to 12-year-old children with little or no schooling. Children's performances were analyzed on three types of mathematical problems: representation of large numerical values, arithmetical operations on currency values, and ratio comparisons.

2. Rationale

Research on mathematical understandings has increasingly documented many competencies that children acquire outside formal education. Young children know principles underlying counting, number conservation and the ability to use a variety of arithmetical problem-solving procedures that are not directly taught in school. Research with both non-Western populations and populations in developing countries where schooling is not universal points to the wide variations that mathematical procedures can take.

The view adopted in the present study is that the particular form that informal mathematical knowledge takes is the result of an interplay between the character of the mathematical problems with which individuals are engaged in everyday practices and the prior knowledge that they bring to bear on those problems.

In the present research, Brazilian candy sellers were selected for study because of the unique nature of the knowledge they bring to bear on their practice as well as the complexity of the mathematical
problems with which they are engaged. Monetary inflation in Brazil has resulted in situations where all Brazilian children identify and use large numerical values as a part of their everyday activities. In the candy-selling practice mathematical problems involving arithmetic with multiple bills and ratio comparisons emerge. Problems requiring adding and subtracting multiple bills of large values emerge when candy sellers must determine whether they have enough cash to purchase a wholesale box of candy or when sellers make an effort to keep track of the amount of cash they have during the course of their sales. Ratio comparisons emerge as a result of a pricing convention that has evolved to reduce the complexity of the arithmetical problems in retail sales transactions.

The primary goal in this study was to understand both the nature of children's prior knowledge of large-number representation and the emergence of mathematical understandings as a function of practice participation. Two additional secondary goals of the research were to examine if Brazilian children’s use of large as contrasted with small currency values in everyday activities would lead them to greater facility with larger values in numerical comparison and to examine if adding and subtracting multiple bills of multiple denominations may vary as a function of problem complexity.

3. **Research Design and Procedures**

The sample for the present study included 23 candy sellers (mean age = 10.8 years, SD = 1.0 years; mean grade level = 1.6, SD = 0.6), 20 urban nonsellers (mean age = 10.8 years, SD = 0.8 years; mean grade level = 1.8, SD = 0.4), and 17 rural nonsellers (mean age = 10.6 years, SD = 0.8 years; mean grade level = 1.3, SD = 0.5).

Candy sellers were included in the study only if they had sold candy for at least a three-month period, were between 10 and 12 years of age, and had not progressed beyond grade 2. Urban nonsellers were
recruited from first- and second-grade classes at public schools. The schools were similar to those attended by sellers or schools sellers would have attended if they had enrolled in schools. Urban nonsellers were included only if they had no selling experience, were between 10 and 12 years of age, and were enrolled in school during the time the study was conducted. The rural nonsellers were 10- to 12-year-olds recruited in a small remote town who had not progressed beyond second grade.

The following tasks were used in an interview to measure children's mathematical abilities. Three measures were used to identify and compare large numerical values:

1. **Identifying numerical values using the standard orthography.** -- Children were presented with 10 pairs of numbers written in the standard orthography, pair by pair, and asked to read and compare them.

2. **Identifying numerical values on the basis of the orthography versus figurative characteristics of bills.** -- To assess children's ability to determine the numerical value of bills by relying on bills' figurative characteristics, as contrasted with the printed orthography, children were presented with 12 trials requiring them to identify numerical values in each of three conditions: one in which they had to identify the numerical values of bills; another in which they had to identify the numerical values of identical bills, but with the printed numbers on the bills occluded with tape; and another in which children had to identify the printed number on the bills without referring to the bills' figurative characteristics.
3. Identifying numerical relations between currency values. -- The purpose of this task was to assess children's knowledge of the structure of the currency system. Children were presented with 14 pairs of bills and/or coins pair by pair (the order of presentation was randomized for each subject) and asked to compare them to determine which was the greater value and how many of the lesser value were equivalent to the greater value. Two measures were used to assess children's ability with arithmetic operations on currency values. To assess children's ability to solve addition problems involving multiple bill values, children were presented with two addition problems that differed in the number of bills to be added and the sum of those bills. To assess children's ability to solve subtraction problems, children were presented with two problems that differed only in the price of a box to be purchased and the amount of money available to purchase the box. The measurement of ratio comparisons was presented to children as follows. To introduce the ratio-comparison problems, the interviewer presented a bag of Pirulitos (another type candy) and told the child the following: "Suppose that you bought this bag of Pirulitos, and you must decide the price you will sell the pieces for in the street." The child was then administered three problems successively that varied in pricing ratios to be compared.

Results were analyzed by using ANOVA's on proportion of correct responses and Duncan's multiple-range test.
4. Findings

None of the population groups displayed a high level of competence in reading numbers. The urban nonsellers with the greatest mean proportion correct only had about one-half correct responses.

Virtually all children from each of the three groups were able to correctly identify the Standard Bills. This indicated that children across population groups had developed an ability to use bills and did not need to rely on their imperfect knowledge of the standard number orthography to represent and identify large values.

Children's ability to produce ordinal comparisons between currency units were analyzed with no significant effects.

Analysis of cardinal relations revealed that both the urban and rural nonsellers achieved more correct answers to the problems involving the larger as contrasted to the smaller values, whereas there were no difference in the sellers' performance across the problems. The sellers achieved more correct identifications on the smaller-value problems than the rural nonsellers.

Analysis of the three groups adding currency values revealed significant differences. The sellers achieved more accurate sums than the urban and rural nonsellers. None of the children used the number orthography to represent the problem in paper-and-pencil solution strategies.

Analysis of the three groups subtracting currency values also revealed a significant effect. Rural nonsellers achieved significantly less accurate scores than both sellers and urban nonsellers. Children did not use paper-and-pencil solutions for the subtraction problems.
Ratio comparisons were analyzed to see how accurate children were in their judgments about which ratio would provide the greatest profit. Candy sellers achieved more appropriate responses than both groups of nonsellers. There were no significant differences between the urban and rural nonsellers' performances.

5. Interpretations

It was concluded that the present study provides insight into the relation between children’s engagement with everyday practices and their developing mathematical understandings. Children generate mathematical problems as they participate in cultural practices such as candy selling. These problems are linked to both larger social processes, such as the inflated money system, and to local conventions that arise in the practice, such as the ratio conventions for retail pricing. It was suggested that in addressing these problems, children create new problem-solving procedures and understandings well adapted to the exigencies of their everyday lives.

Abstractor's Comments

The reported fact that all of the nonsellers were enrolled in first or second grade at the time of the study while many of the sellers were not currently enrolled in school or were never enrolled was an interesting difference that needed more explanation than it received in the article.

Most, if not all, of the data collection came from an interview process with the three groups of children. The questions used in the interviews are detailed in the article but the implementation of the interviews is not clear. Was each child interviewed separately with similar time frames or were there differences in the interviews among the three groups of children?
The rural nonsellers were selected from a small remote town 100 miles from the town where the other two groups of children were selected. The author had used children from this small town for another study on topological tasks. Readers are recommended to read other references listed in the article by the author to obtain a more complete feeling of his research efforts.

Various ANOVA's, Duncan, and Whitney U Tests comparisons were reported. It was difficult to determine if the reported statistical analyses and tables added or distracted from the reportings in the article.

The reason for this research was to find if children construct novel understandings as they address problems that emerge in their everyday cultural practices. The design for this research was clearly reported, as were the procedures employed and the findings reached.

Abstract and comments prepared for I.M.E. by CLYDE A WILES, Indiana University Northwest.

1. **Purpose**

   This study tests three implications of a theory of mathematics learning relating to the meaning and use of decimal fractions.

2. **Rationale**

   The authors have been developing a semantic and syntactic learning theory applied to decimal numbers and numerals and the symbolic manipulations of those symbols.

   Briefly, students become competent with decimal symbols when they develop four (actually five) processes in the following sequence. The first two are semantic processes, and the last two are syntactic processes.

   Two semantic processes:
   a) connecting process - the connecting of individual symbols with referents (concrete referents that may be manipulated)
   b) developing process - the developing of symbol-manipulation procedures with reference to the concrete referents

   Two syntactic processes:
   c) elaborating (1) and routinizing (2) the rules for the symbols
   d) using the symbols and rules as referents for a more abstract system.

   It is also postulated that an alternate sequence of acquisition (syntactic than semantic processes) will be more difficult and may prevent the development of competence.
The hypotheses were:

1. Students who acquire and use the connecting and developing processes (semantic meanings based upon decimal blocks) will solve tasks correctly and
2. will transfer these processes to novel tasks.
3. Students who have already routinized syntactic rules (symbolic manipulation skill) without establishing connections between symbols and referents will be less likely to engage in the semantic process than students who are encountering decimal symbols for the first time (in the recommended sequence).

3. Research Design and Procedures

Subjects. Five students were selected from each of two classes at each of grade levels 4, 5, and 6. Their teachers selected them as: below average (1 student), above average (1 student), and average in achievement (3 students). One fourth-grade student dropped out of the study, leaving a total of 29 students.

Five of the fifth graders and all of the sixth graders were classified as "exposed previously to a syntactic approach for dealing with decimals." This instruction was assumed to be empty of semantic processes.

Instruction. Nine 25-minute lessons were provided for all subjects by one of the authors. The first five focused upon establishing connections between decimal numerals through .01 and decimal blocks (cubes, flats, longs, etc.). The last four lessons focused upon developing addition and subtraction of decimal numbers, with explicit reference to the blocks which were combined or separated as appropriate. Problems were presented with written symbols, and students were asked to determine the answers by manipulating the blocks. Students were eventually asked to "work the problems on paper, but to think about the blocks before deciding what to do."
Assessment. All assessment was by interview. The interviews were standardized with respect to tasks and sequence, but follow-up questions varied as needed to get students to describe how they solved each problem. The interviews were taped and transcribed. Tasks were "direct measures" of tasks that had been taught in the nine lessons, and "transfer measures," tasks that were novel to the instructional experience.

The tasks included the following:

Direct:

\[2.3 + .62 = 5 + .3 =\]

Transfer:

1. Represent 1.503 with blocks. [This was possible using blocks that were present, but not used during instruction.]
2. Represent 1.0623 with blocks. [This could not be done using the blocks provided.]
3. Choose the larger of .5 and .42.
4. Write .7 as a common fraction.
5. Write 6/100 as a decimal fraction.

These tasks were selected on the basis of prior inquiry as to what kind of tasks made discrimination between student use of a semantic analysis or syntactic rule most evident.

Items were scored as correct or incorrect, and the method of solution classified as semantic or non-semantic. The process was identified as semantic if the student referred to the values of the numerals, using the decimal blocks as referents, or any other verbal description of the quantities. Inter-rater agreement was 59 of 60 decisions based upon the interview recordings and scripts for three randomly selected students.

Procedure. Children were pretested on both the direct and transfer measures except that the two transfer tasks involving the use of blocks were not given at the pretest. Children were taught in grade-level groups ranging in time from 7 to 9 days. The post-instruction interviews took place about 6 weeks after instruction.
4. **Findings**

The data were described in narrative form and were organized into three tables. Table 1 reported the strategy employed before and after instruction for the five transfer tasks, Table 2 reported the process used and the correctness of the response related to each process for two direct tasks for the last three transfer tasks, and Table 3 reported the frequency of the use of semantic analysis before and after instruction by those who had prior (syntactic-based) instruction and those who had no prior instruction. No statistical tests of significance were reported.

With respect to the "connecting process":

1. Prior to instruction 3 of the 29 subjects could represent a decimal numeral with the decimal blocks; 24 of 29 students could do this after instruction.

With respect to the "developing process,"

2. Prior to instruction, 2 of 29 students appealed to the values of the digits in their explanations of how to solve one of $2.3 + .62$ and $5 + .3$; 19 of the 29 did so after instruction.

With respect to transfer:

3. After instruction 27 of the students could represent 1.503 with the decimal blocks. (Note that instruction did not explicitly involve the thousandths place, and this task was not assessed at the pretest.) And 19 of the 29 could represent 1.0623 with blocks. (The ten-thousandths position, however, could not be represented by the available blocks.)

4. For the other three transfer tasks, 6 of 87 responses involved a semantic analysis before instruction, while 40 of the 87 responses involved a semantic analysis six weeks after instruction.

With respect to correctness:

5. For the direct items, 31 of 31 semantic analyses produced correct responses, while 23 of 27 nonsemantic processes produced correct responses.
6. For the transfer items, 13 of 40 semantic analyses produced correct responses, while 9 of 47 nonsemantic process responses were correct.

With respect to sequence:

7. For the two direct items, 13 of the 30 responses for those with previous instruction employed semantic analysis. Of those with no previous instruction, 18 of 28 responses used semantic analysis.

8. For the three transfer items, 23 of the 45 responses for those with previous instruction employed semantic analysis. Of those with no previous instruction, 18 of 42 responses employed semantic analysis.

5. Interpretations

Almost all of the students had acquired the connecting process. And the following findings were considered to be most interesting:

i. Most students acquired and used the developing process.

ii. The process was relatively robust in that it remained in place six weeks after instruction.

iii. About one-half of the students used the process flexibly, i.e., in transfer tasks.

iv. The process was highly related to correct performance on both the direct and transfer measures.

With respect to the sequential prediction from the theory, "Prior instruction that encouraged the routinization of syntactic rules seemed to interfere with and prevented the adoption of semantic analyses of the affected tasks. ...However, prior instruction that provides factual or conceptual knowledge may support the acquisition of semantic processes."

The findings were seen as quite consistent with the theory proposed, and with the findings and theory of others, i.e., "it is preferable to develop meanings for symbols before practicing syntactic routines."
The theory of semantic learning via connecting and developing processes was contrasted to mapping instruction (as discussed by several others). While both mapping and semantic analysis aim to help students make sense of symbolic rules through connections with (concrete) referents, mapping intends to produce algorithms and rules that are consistent with the use of the referents. Furthermore, a semantic analysis makes many rules unnecessary in a mature syntactic system.

The semantic processes facilitate problem solving in that they generate conceptual representations of symbolically presented problems. In short, these connected "representations have rich associations that guide the development and selection of symbol procedures (for problem solving)."

The semantic processes constitute the definition of understanding in contexts of written symbol expressions. It is further noted that "as one might expect, understanding (as defined by the processes) played a greater role in solving problems that were relatively novel than in solving problems for which a symbol rule already had been learned."

Abstractor's Comments

1. This study is part of a larger set of studies that have been conducted by these authors and coordinated with the work of others. Much of this work to date is summarized in articles in the NCTM publication Number Concepts and Operations in the Middle Grades (1988). Anyone interested in understanding the scope of what is intended should read this book.

2. The disciplined nature of this study is impressive. It is well done and fits logically and necessarily into the long-range goal indicated by the title. The data-gathering techniques were most appropriate for this study.

The study is directed toward the development of a theoretical framework of mathematics learning of the type that we must have if we
are to make more complete sense of old and new inquiries. The fact that it is part of a continuing effort makes it much more valuable than it would be by itself.

3. The strongest findings/interpretations of this well-done and carefully reported study concern the ability of students to acquire semantic and then syntactic learning. Students did acquire semantic meanings provided in instruction, and did apply them (transfer them) to solve novel situations. And furthermore, students who acquired these semantic meanings were able to solve problems that students who used syntactic reasoning did not.

The fact that students did acquire the ability to use the models - use the semantic processes - is an important matter and is not obvious. While it seems clear that opportunity to acquire a working knowledge of any semantic is largely lacking in schools, one might suppose that this represents a state of curricular evolution that corresponds to what children can and do learn. To demonstrate that learning does not have to be purely mechanical is important. The "payoff" is real in that students who were able (willing?) to employ a semantic did solve problems that those who did not use a semantic did not solve.

4. The hypothesis of sequential interference, however, (i.e., that students who acquired syntactic learning before semantic meanings would resist the learning of semantic meanings) is not so strongly supported.

While no statistical tests were reported for the data, they could have been done. The transfer tasks, for example, could be set up in a 2 x 2 contingency table as Semantic Analysis (used, not used) versus Items (number correct, or incorrect). When the data provided by the study were so analyzed they yielded a Yates corrected Chi Square of 34.89, with p < .001 against the null hypothesis. This strongly supports the hypothesis that semantic learning greatly facilitates transfer.
However, the contingency table for Prior Instruction (prior syntactic instruction, no prior instruction) versus Analysis Used (semantic, not semantic), for the transfer items produced a corrected Chi Square of only .607. If we restrict the data to those "possible changes" in the sample - those students that could have changed from a syntactic approach to a semantic analysis requiring cubes - we get a Chi Square of 3.066 with p < .10. Thus the hypothesis that students who have acquired syntactic learning resist semantic learning is appealing, and squares with personal informal observations, but the data do not provide unequivocal support. The authors use the language "seems to interfere with...the adoption of semantic processes" in their description of the data, but from the data one might just as well have said there seems to be not much of an effect, if any. Furthermore, the meaning of "resist" is not clear. If we suppose students do resist this semantic, is this because they have great difficulty rearranging their thinking, or do they resist as a matter of choice? Experience indicates that persons who have confidence in a method (even if misplaced) will often not attend to something that their personal agendas do not appear to require. Such persons may well suppose that since they can already do a task, they don't need another - perhaps more cumbersome - way of doing it. If this conjecture is true, the apparent weak interference in actually due to an unwillingness to attend to instruction rather than an inability or difficulty in doing so. Motivation, not cognitive structure, would be the issue.

The whole matter or retroactive inhibition in this study is cloudy for several reasons. First, the .10 level of significance is at least equivocal. Second, how did those disqualified students acquire a semantic of decimal cubes or whatever without any instruction with the cubes or other referent? I'm surprised that apparently none of these students used a money semantic. If dollars, dimes, and pennies had been in the visual field, would none of the subjects offer explanations for decimal numbers in terms of them?
5. The authors' attempts to define "understanding" and "competent
with the written symbols of the decimal fraction system" are admirable
efforts. However, the definition of understanding is thin at present
in that it does not account for levels of understanding. For example,
a person who can use (connect with and develop operations with respect
to) one referent obviously has less understanding than a student who
can do these same tasks with two, three, or more referents. How many
referents are necessary before a student "understands" or is
"competent", or "fully competent"? Is one enough? It seems
doubtfully valid to say a student who meets all the semantic criteria
for decimal blocks and for syntactic processes but who cannot apply
the knowledge to length, area, weight, and time measures, truly
"understands." How many referents are enough? Why is the volume
model of the decimal cubes so important? Why not choose length as the
first model? Why not use money with its pervasive presence in the
child's environment as the first model? I wonder if we would find
that students who had achieved competence and understanding with
respect to say volume, money, and length might easily (however that
might be measured) apply their understandings to new applications and
other referents.

The authors also note that students with syntactic understanding
at the outset transferred at least as well to the transfer tasks
involving alternate numeration symbols. Now suppose that the nine
lessons had been devoted to developing skill - syntactic skill - which
are in the authors' model of understanding as elaborating/routinizing
processes. Might we then find that students who had acquired a
semantic meaning would resist the syntactic learning? Personally, I
would expect the semantic then syntactic sequence would prove to be
more efficient than the reverse, but if we take a student through all
five of the processes (connecting, developing, elaborating,
routinizing, and abstracting), I wonder if the efficiencies would be
at all significant or important?
6. I liked the authors' insistence upon the terms semantic and syntactic even though one might argue that "the meaning" and "rote procedures" could be used. These terms keep our focus on these two readily distinguishable aspects of children's and adults' performances with respect to practically all mathematics, and force us to think more carefully about what we mean by either meaning or rote. Also, the idea that one or another sequence may be more efficient in either the long range (e.g., through secondary schooling?) or short range (e.g., the content presented through grade 5) is important.
1. Purpose

(a) To determine if second graders can be taught to use schematic drawings to represent differing categories of addition and subtraction word problems.

(b) To determine which of the 12 problem types are within the zone of proximal development of second graders when multidigit numbers are used.

(c) Given that second graders can be taught to use schematic drawings to represent word problems, then determine if such an ability improves their ability to solve word problems.

2. Rationale

The most common method of teaching addition and subtraction word problems ignore children's need to represent the problem situation and instead focuses only on the solution strategy. The disadvantage of this approach is particularly strong for the more complex kinds of word problems, for these require not only that children represent a problem but also that they reflect on that problem representation and modify it in some way in order to select a solution strategy. A teaching method that helps children to represent the problem situation would be more helpful than the prevalent solution sentence method.
3. Research Design and Procedures

The subjects were in two second-grade classes. One class (HA) contained 24 subjects and was categorized by the teachers as containing children with high mathematics ability; the other class (AA) contained 19 subjects categorized by teachers as children of average mathematics ability.

There were 12 categories of word problems and four schematics taught to each subject. The HA subjects were first introduced to a general category of word problems and shown the drawing for that category. The subjects were taught the labels for that category. The subjects were taught the labels for each schematic. The subjects were taught for any given problem they should first write on the problem the letter for the verbal labels used in the particular schematic. After the labelling of the problem, the subjects were taught to make the appropriate schematic drawing. The numbers from the specified problem were then written in the appropriate parts of the drawing. Finally, the subjects selected the solution procedure by determining how to obtain the unknown given the two filled parts of the drawing. The drawing facilitates the selection either by the relative sizes of the sections of the drawing or the temporal ordering of events in the drawing.

The HA subjects were given worksheets used not only for student practice, but to provide data on problem difficulty. There were three types of worksheets: Pure - 18 problems from one category distributed equally over the three possible positions for the unknown; Compare - 27 problems from the Compare category distributed equally over the three positions of the unknown and problems in which more or fewer appeared; Mixed - 24 problems equally distributed over the four categories and the three positions of the unknown.
Teaching was divided into units, each based on one major problem category. One mathematics period was spent on instruction for a particular problem category and associated schematics; two to four periods were spent on practice for that category. After all four categories had been introduced and practiced, a unit on problems of mixed categories was presented.

It was intended that the same procedures would be used for both the HA and the AA groups. However, the AA group was taught near the end of the school year and numerous unexpected activities cut down the amount of time spent on practice, especially of mixed problem categories.

Percent correct data were provided for the HA posttest and mixed worksheets and the AA posttest for each of the four problem categories based on the subject selecting the correct drawing for the problem type, filling in the draw correctly, whether the drawing was an accurate representation of the specific problem, and whether the student selected the correct solution strategy. A test of correlated proportions was used to determine the significance of any pre-posttest improvements in selecting the correct strategy and getting the answer.

4. Findings

(a) The subjects were able to make all the schematic drawings used in the study.

(b) The range for the percent correct in selecting the appropriate drawing to make was from 61% to 92%. The students confused the "Compare" category problems and the "Put-Together" category problems.

(c) The HA students averaged 66% correct for filling in the drawings on the worksheets, but improved to 82% correct on the posttest. Since the AA group had so little time to spend on the worksheets the comparison was not made.
(d) The percent correct for the drawing adequately representing the elements of the problem ranged from 69% to 93%.

(e) The percent correct for selecting the correct solution strategy ranged from 69% to 95%. There were some statistically significant differences to support the finding that even with the schematic drawing children have difficulty interpreting problems that "sound" as if they are subtraction problems when they actually require an addition solution strategy.

(f) There was little evidence to support the proposal that children use a part-part-whole schema to solve all of the word problem types presented.

(g) There was a significant pre-posttest improvement for correct use of the solution strategy and getting the answer.

5. Interpretations

Second graders of average and above-average mathematics ability can make the schematic drawings used in this study, can reliably distinguish among the different semantic word problem categories and make the correct drawing for a problem for the category, can usually insert the numbers from the problem into the drawing correctly, and can select the correct solution strategy for most problems. In addition there was significant improvement in selecting the correct solution strategy and getting the answer.

Overall, the posttest scores indicate that all 12 of the problem types are well within the "zone of proximal development" of American second graders of average and above-average ability. The almost total omission of the more difficult subtypes of addition and subtraction word problems from United States textbooks thus does not seem warranted.
Abstractor's Comments

As with all abstracts only the highlights have been presented. Just as I had to decide what to include and what to exclude and still be able to give the reader the "flavour" of the article, the authors had to pick and choose what to include in the article. How much time was spent by the authors considering the topics listed below is unknown.

(a) If one accepts the interpretation of the findings by the authors, then it is established that the subjects could be taught the use of schematics to solve the specified problems. There is very little discussion as to why one would want to do such a thing or the cost for having done so.

There is no discussion that such an approach generalizes to other problem types. The whole concept appears to be a version of teaching "upstream/downstream" problems in the secondary schools. In this age of emphasis on problem solving there is an attempt to move away from teaching techniques that are specific to certain types of problems. Considering that the instructional period was from 16 - 26 days, it is a very expensive technique in today's cramped curricula.

(b) I generally do not agree with the statement, "The strategy scores generally improved significantly on all problems for which there was sufficient room for improvement except on the 'Change-Get-Less: missing start' problem." I would categorize 12 of the 20 comparisons as having "room for improvement" and only three of these yielded significant differences. Perhaps the authors' definition of "room for improvement" was different from mine; no definition is given. I used
those comparisons with a pre-test score less than 70%. Of the eight comparisons with little room for improvement (pre-test scores > 70%), four show no improvement at all.

(c) The authors feel that their study shows that there is little evidence to support two hypotheses stated in other research: (i) children use a single part-part-whole schema for the representation of all four kinds of addition and subtraction word problems and (ii) children use the part-part-whole schema to solve the more difficult kinds of Change problems. My understanding of both of those proposed hypotheses is given normal instruction and left to their own devices children will tend to do as hypothesized. The current study shows only that it is possible to provide children alternatives which they will use in lieu of blindly staying with the part-part-whole schema.

(d) One of the changes in procedures which occurred as the study progressed was dropping the requirement that each child write an equation for each problem. At the end of the article the authors state: "The use of equations that semantically model a problem as intermediate representations during problem solution seems to be superfluous in conjunction with the schematic drawings." Superfluous or not it would appear to be the ideal time to introduce such equations. It would have been interesting to read when such equations should be introduced according to the authors. Again, one cannot put a discussion of every issue in one article.

As with any article there are picayune items such as part of the problem is left out of the example for Change-Get-More: missing start in Table 1, use of
thinking strategies as a "method of adding or subtracting," and each ability group being defined as simply a class categorized by teachers as containing children of the specific ability.
Zehavi, Nurit; Bruckheimer, Maxim; and Ben-Zvi, Ruth. EFFECT OF ASSIGNMENT PROJECTS ON STUDENTS’ MATHEMATICAL ACTIVITY. Journal for Research in Mathematics Education 19: 421-438; November 1988.

Abstract and comments prepared for I.M.E. by DONALD J. DESSART, The University of Tennessee, Knoxville.

1. Purpose

The purpose of this study was three-fold: (a) to evaluate the effectiveness of problem-solving assignment projects with high-achieving ninth-grade students; (b) to determine links between student difficulties in problem solving and certain guidance techniques; and (c) to study the applicability of the guidance procedures in problem-solving situations.

2. Rationale

Problem solving has held the attention of teachers and researchers for at least a decade. It is highly cherished goal of most mathematics teachers. Unfortunately, the most effective ways to teach and evaluate student problem solving have evolved very slowly. The overall problem has three facets: (a) the development of true problem-solving materials that are within the grasps of students, requiring mathematical techniques that are familiar to them; (b) the development of guidance techniques to help students solve nonroutine problems; and (c) the development of evaluation methods to ascertain the degree of student success in problem solving. The mathematics education community has made reasonable progress in all three facets of the overall problem, but much future work will be needed.

3. Research Design and Procedures

The complete study consisted of three substudies. First, problem-solving assignment projects were developed which were open-ended and were restricted to the curriculum of ninth-grade students. Assignment tests were developed which were used as pretest and
posttest instruments in the initial investigation. Evaluation techniques were developed which included attention to the quality of the procedure (trial, incomplete argument, and complete argument) and the quality of the result (primitive, partial, and full). A quality-of-solution score (Q) combined the quality of procedures and quality of results into one score which varied from one through six. An achievement test and an attitude questionnaire were also constructed.

Complete data were available for 191 students in the experimental group and 194 in the control group. The students in the experimental group worked on three or four problem-solving projects during a four-month period. Written and verbal guidance of the students were given by the investigators as well as the teacher. The pretest and posttest versions of the assignment test, the achievement test, and the attitude questionnaire were administered to the students of the experimental and control groups.

In the second part of the complete study, conducted one year later, the pretest version of the assignment test was given to a ninth-grade class. Their unassisted responses were classified by three levels of result and three levels of procedure. All students were also interviewed to determine obstacles that may have impeded their progress. A hierarchy of hints: instrumental (I), instrumental-relational (IR), relational (R), relational-logic (RL), and logical (L) were developed. In the interviews, most of the students were initially provided a logical hint, followed by a relational-logical, and so on, with progressively lower-order hints, until student progress was observed. The order of hints and results were recorded.

In the third and final study, six teachers from the first study were selected with their ninth-grade students. The six classes worked during one class period on an assignment test. Student papers were reviewed independently by the teachers and investigators and were returned to the students with hints for further work. The students spent another class hour on the problems using the hints. The responses were categorized as instrumental (I) when Q was less than two, as relational (R) when Q was three or four, and as logical (L) when Q was 5 or 6. Profiles of responses were constructed for each student.
4. Findings

In the first study, it was found that the experimental group did significantly better (p < .01) on the achievement posttest, whereas, the control group did not show significant improvement. There was a significant decline in the attitudes of the control group but not the experimental group. A significant increase was found for the experimental group and a nonsignificant decrease for the control students in the assignment test.

In the second study, one-third of the responses was classified as primitive with incomplete arguments or as a partial result in a trial procedure. The hints which were effective were noted, and it was seen that the quality of the effective hint was related to the quality of the original response. The findings for all assignments indicated a regularity among procedures, results, and hints that resulted in a model of eight categories, e.g., incomplete argument, partial result, and relational-logical hint. These eight categories accounted for 72 percent of the cases. The remaining cases were equally divided among those who responded to a hint at a higher level than expected or at a lower level than expected.

In the third study, the findings were related to the triple profile patterns, e.g., LRI (logical, relational, instrumental), on the three assignments. In general, it was found that for "constant" profiles, e.g., RRR, higher hints than suggested by the model could be offered; but for those cases in which the profile showed an increase, lower-level hints generally helped.

5. Interpretations

The investigators concluded that the first study demonstrated that the assignment projects were effective in promoting student problem-solving activity. However, the treatment was far too demanding for both the teachers and the students. The researchers concluded that limited treatments at repeated intervals were better than the full treatment tried with the experimental groups.
The second study led to a viable diagnostic classification of qualities of mathematical activity and an associated hierarchy of guidance procedures.

The third study resulted in a further elaboration of the guidance model based upon a study of profile clusters. The result was an organized framework for qualitative evaluation of open-ended problem-solving activities.

Abstractor's Comments

This study represented a major step forward in the development of viable problem-solving materials, in the development of guidance techniques for helping students progress in problem solving, and in the evaluation of student problem-solving activities. These researchers are to be highly commended for their fine work. In spite of the progress made by this study, much has to be done before we will see problem-solving activities being properly evaluated by common, standardized measures so that teachers will be motivated to emphasize problem-solving activities instead of low-level skills.

True problem-solving activities (e.g., the Polya Model) are applicable to many non-mathematical situations, too. The physician diagnosing an illness and the mechanic finding a problem in a faulty engine use problem solving. The ultimate step in research will be an evaluation model that will be equally applicable to mathematical as well as non-mathematical problem-solving situations. This is certainly a worthy, future goal for researchers.

The study reported in this abstract has shortcomings that the authors will probably readily admit. The samples, particularly in the second and third studies, were quite small, so that valid generalizations are elusive. The investigators had to play vital and significant roles in the study. This was a necessity because many of the teachers were apparently not readily knowledgeable in the intricacies of the evaluation and guidance techniques. This is understandable; however, the experimenter-bias detracts seriously from the internal and external validity of the studies. It is clear that the techniques of evaluation and guidance developed in this study will have to receive much greater refinement before they will be ready for application in most mathematics classrooms.
MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN JOURNALS AS INDEXED by CURRENT INDEX TO JOURNALS IN EDUCATION
April - June 1988


ED 288 723 Blank, Rolf K. Results of an Inventory of State Indicators and Data on Science and Mathematics Education. Science and Mathematics Indicators Project. 17p. MF01/PC01 Plus Postage. Available from EDRS.


ED 289 390 Ettrington, Corinna A.; Wolfle, Lee M. The Selection of Quantitative Undergraduate Fields of Study: Direct and Indirect Influences. 27p. MF01/PC02 Plus Postage. Available from EDRS.


ED 289 692 Linn, Mary McMahon. Effects of Journal Writing on Thinking Skills of High School Geometry Students. 72p. MF01/PC03 Plus Postage. Available from EDRS.
ED 289 693 Lichtenheld, Peter James. Descriptive Study of a Prescription Learning Basic Mathematics Skills Multi-Media Lab in an Elementary School Setting. 120p. MF01/PC05 Plus Postage. Available from EDRS.

ED 289 703 Harvin, Virginia R. Problem Solving Skills of Selected Fourth Grades. 16p. MF01/PC01 Plus Postage. Available from EDRS.

ED 289 704 Ben-Chaim, David; And Others. The Aronberg Project: An Intervention Program To Enhance Teacher and Student Performance in Mathematics in Socially Deprived Junior High Schools. Final Report. 247p. MF01/PC10 Plus Postage. Available from EDRS.


ED 289 727 Matsuda, F.; And Others. Concepts about Relations among Time, Distance and Velocity in Children. 17p. MF01/PC01 Plus Postage. Available from EDRS.


ED 289 907  O'Neal, Maria R.; And Others. Predicting Success and Measuring Math Skills Improvement in an Introductory Statistics Course. 26p. MF01/PC02 Plus Postage. Available from EDRS.

ED 290 467 Flaherty, Robert John. A Determination of the Relationship between Various Factors and Success on the Advanced Placement Computer Science Examination. 46p. MF01/PC02 Plus Postage. Available from EDRS.

ED 290 628 Hardiman, Pamela Thibodeau; Mestre, Jose P. Understanding of Multiplicative Contexts Involving Fractions. 43p. MF01/PC02 Plus Postage. Available from EDRS.