**The *Investigations* Curriculum and Children’s Understanding of Whole Number Operations**

by Jan Mokros, TERC, Cambridge, MA (2000)

The research described in this paper was supported, in part, by the National Science Foundation Grant No. ESI-9050210. Opinions expressed are those of the authors and not necessarily those of the Foundation.

**Introduction**

This paper describes four important studies concerning the impact of the *Investigations in Number, Data, and Space* curriculum. The goal of these studies was to examine the effects of the curriculum on children’s understanding of number and number operations. While other areas of mathematics also are emphasized in the curriculum and have been addressed by other researchers, numerical understanding is a critical piece of the curriculum’s mathematical agenda. It also is the most visible and recognized part of mathematics education in the eyes of the American public.

Of the four studies discussed here, two (Mokros et al., 1994; Mokros et al., 1996) were conducted by researchers during the time the curriculum was under development. The other two (Goodrow, 1998; Flowers, 1998) are comprehensive studies that examined the impact of the published curriculum. The studies have several commonalities. Each poses calculation problems as well as word problems involving whole number operations, and each examines the accuracy and effectiveness of participants’ methods of solving problems. Each of the studies involves individual interviews to provide in-depth information on children’s mathematical thinking. In addition, each study is based on an experimental design that involves both an *Investigations* Group and one or more Comparison Groups. Finally, unlike some studies of curricular impact, the studies presented below are based on a considerable amount of data concerning the actual implementation of the *Investigations* or comparison curricula.

*Note: All figures appear at the end of this article.*
Research Study 1

Learning Operations: Invented Strategies That Work

Background and Method

The major question this study (Mokros et al., 1996) addressed was how students would perform on computation tasks and word problems involving basic arithmetic operations. The study took place during the 1994–1995 school year. It involved 50 second-grade students, 30 of whom were participating in Investigations classrooms and 20 of whom were from classrooms using another curriculum. The Comparison classrooms used a curriculum that emphasized both invented procedures and traditional algorithms. Both groups were from suburban Boston schools, with the Comparison students having a smaller class size and being from a somewhat more affluent community. The Investigations classrooms were fully implementing the curriculum in its draft version.

Children were given end-of-year interviews which involved a series of word problems. (Problems are shown in Figure 1.) The problems were read aloud and the numbers involved were presented on cards (with no operational signs attached). Children had access to counting materials and pencil and paper. They were asked to solve each problem and show how they solved it. Scoring involved an analytic rubric, which addressed accuracy, type of inaccuracy, use of materials, quality of the strategies used, and strength of the mathematical explanation. Scoring rubrics were determined to be highly reliable, with 80% to 100% reliability achieved on the five criteria listed above. All interviews were scored blindly by coders who had no knowledge of the group to which each student belonged (Investigations or Comparison).

Results and Discussion

Students using Investigations achieved greater accuracy on the problems than students in the Comparison Group. (See Figure 2 for item analyses.) Of seven possible points for accuracy, the mean score for Investigations students was 4.6 items correct, compared with 3.25 for the Comparison Group ($t = 2.13, p < .05$). Item analyses showed that Investigations students were significantly more accurate than Comparison students on all three addition/subtraction problems. The problems involving multiplication, division, and the use of factors showed the same trend, but no significant differences. (Further exploration of classroom data revealed that these concepts were not yet covered, or were only briefly addressed, by the time of the interviews.)

In terms of the strategies used by students, there were interesting differences between the groups. Investigations students were more likely to use mental arithmetic strategies for addition and subtraction than were students in the Comparison Group. Investigations students were accurate in their use of mental arithmetic strategies 80% of the time, compared with a 30% accuracy rate among the Comparison Group. Furthermore, students who relied on their own strategies, regardless of which group they belonged to, were more accurate than those who relied on the standard algorithms for addition and subtraction involving regrouping.
Research Study 2

Full-Year Pilot Grades 3 and 4: Investigations in Number, Data, and Space

Background and Method

This study, conducted in the 1992–1993 school year, was an early examination of the impact of Investigations on third- and fourth-graders' understanding of number, as well as data and geometry (Mokros et al., 1994). A total of 96 students, randomly selected from seven Investigations classrooms and five Comparison classrooms, were involved in the study. The classrooms were located in a range of urban, suburban, and rural communities in Massachusetts, which collectively provided substantial diversity with respect to race, first language, and background characteristics of students. The five Comparison classrooms were using a variety of math textbooks and were not using reform math curricula.

Students were assessed at the beginning and again at the end of the school year, using a pre-post test design. There were two components to the assessment: 1) A non-timed, written test of calculation skills (about 20% word problems and 80% number problems). This test consisted of 20 items (3rd grade) and 22 items (4th grade) which were based on textbook exercises at these grade levels. 2) Individual interviews consisted of more complex word problems involving numerical reasoning. Students had access to manipulatives and pencil and paper as they worked on the interview problems. The interview problems are displayed in Figure 3.

To increase methodological rigor, the researchers who scored written tests and interviews were kept blind to student identity and to the curriculum being studied. The written test was scored for accuracy, and the time taken by each student to complete the test was noted. There were no differences between groups with respect to amount of time used to complete the test. Interview scoring involved holistic three-point rubrics, which encompassed the following characteristics: accuracy, completeness of the solution, use of an effective strategy, explanation of how to solve the problem, and use of tools and representations. Inter-rater reliability on all items was at least 90%.

Results and Discussion

As can be seen in Figure 4, students in both third and fourth grades performed significantly better on the calculation posttest than on the pretest. The amount of growth on calculation skills was comparable for Investigations and Comparison students. In other words, Investigations students were at no disadvantage with respect to acquiring solid calculation skills.

Item analyses were conducted (pre- to posttest) for the interview tasks. In examining the results for the five items across grades, we found that Investigations students made significant gains (p < .05) on all of the problems. Comparison students, on the other hand, showed no significant gains on four of the five number tasks. The word problems presented in the interviews were generally more challenging than the pencil and paper test of calculation skills because the word problems require students to figure out the mathematical goal and to choose appropriate arithmetic operations. One student who was interviewed immediately recognized this complexity and asked the interviewer to “Just tell me whether to add, subtract, multiply or divide. Then I can do it.” On problems like these, in which the path to the goal was not provided, Investigations students outperformed students who were using more traditional math programs.
Research Study 3

Modes of Teaching and Ways of Thinking

Background and Method

Goodrow (1998) examined children’s understanding of number composition, place value, basic addition and subtraction facts, and two-digit addition and subtraction. Goodrow worked with 46 children, drawn from three groups: 1) an Investigations Group, in which the curriculum was taught in the manner recommended by developers; 2) a Mixed Group, in which Investigations was taught in a teacher-directed way along with the demonstration of traditional algorithms; and 3) a Traditional Group, which used a traditional textbook. The study examined the manner in which each curriculum was taught.

Children were interviewed individually, using a variety of tasks involving constructing number sentences, performing addition and subtraction, and grouping by tens and ones. All testing took place at the end of the school year. Each child in the study was interviewed several times, with each interview lasting approximately 20 minutes. In addition, computational worksheets were administered.

Results and Discussion

Children’s strategies and level of accuracy were recorded. There were no significant differences between the three groups in terms of accuracy on basic addition and subtraction facts. Accuracy scores ranged from 89.1% to 97.5%. Similarly, there were no differences between groups on two-digit addition problems. However, on two-digit subtraction problems involving regrouping, children in the Investigations Group showed significantly greater accuracy (80%) than those in the other groups [60.8% accuracy for the Mixed Group and 69.2% accuracy for the Traditional Group (p = .023)]. Further analysis showed that students in the Investigations Group used their own strategies rather than using traditional computational algorithms (i.e., “borrowing”).

A major significant difference between groups was also found in the construction of number sentences. Students were asked to choose a number and write as many number sentences as they could (using addition and/or subtraction) that would result in that number. Children in the Mixed Group and Investigations Group produced an average of 6.71 and 5.44 correct number sentences, while those in the Traditional Group produced only 2.88 correct sentences (p = .0025).
Research Study 4

Abstract: A Study of Proportional Reasoning As It Relates to the Development of Multiplication Concepts

Background and Method

This study (Flowers, 1998) involved fourth graders from six classrooms in a suburban school district in the Midwest during the 1996–1997 school year. One hundred thirty-seven students participated. These students were in classrooms which were classified as either: 1) an Investigations Group using Investigations; 2) a Mixed Group using a reform curriculum, which included learning traditional algorithms; or 3) a Traditional Group using a 1992 textbook. Flowers administered a 24-item written test at the middle and again at the end of the school year. The test included 14 computational problems, five word problems involving multiplication and proportional reasoning, and five conceptual understanding problems. A typical calculation item on this test was to solve $112 \div 8$. A more difficult word problem was to figure out how many tigers there are in a zoo if there are 36 lions and there are 3 lions for every 4 tigers. An example of a conceptual problem was to use the multiplication fact $3 \times 8 = 24$ to determine the answer to a related problem, such as $6 \times 8$. In addition, Flowers interviewed six students from each of the three groups, giving them more complex proportional reasoning problems and assessing how they used a known multiplication or division fact to solve a more difficult problem. For example, students were asked how many words a person can read in 18 minutes if she reads 540 words in 4 minutes.

Results and Discussion

On the written test, students in the Investigations Group made the greatest gains overall, with particularly high gains in conceptual understanding. (See Figure 5.) The Mixed Group performed least well, and actually showed a decrease from pre- to posttesting in terms of their conceptual understanding.

On the interview tasks, those in the Investigations Group again performed better than students in other groups. For example, one interview item asked students to determine the distance covered by a bicyclist in 4 hours if she travels at a steady speed and has traveled 20 miles in 3 hours and 40 miles in 6 hours. One-third of those in the Investigations Group presented a correct solution, whereas none of those in the Mixed or Traditional Groups presented a correct solution. Ratio problems such as these are beyond the scope of most fourth-grade curricula, but can be successfully solved by students who have a solid understanding of the nature of multiplication and division.

Flowers notes that the pretest scores of the three groups were not comparable. She dealt with this methodological problem by conducting further analyses to determine the extent to which initially low, medium, and high-scoring children made gains over the course of the study. These data showed that all children in the Investigations Groups—those who initially scored low, medium, or high on the pretest—made significantly greater gains in conceptual understanding than students in other groups. These findings suggest that Investigations has a comparable impact on students across the ability spectrum.
Synthesis of the Findings

*Investigations* students do as well on mastery of basic facts as students using other curricula.

In three of the studies (Mokros et al., 1994; Goodrow, 1998; and Flowers 1998), students completed computational worksheets in which “basic fact” problems were presented in the standard vertical manner. There were no differences between groups with respect to accuracy. Furthermore, all groups made comparable and significant gains over time. One caveat: While Mokros et al., (1994) reported no differences between groups in the overall amount of time taken to complete the calculation test, the other studies offer no data concerning calculation speed.

*Investigations* students do as well or better than students using other curricula in calculation problems.

Collectively, these studies provide a great deal of data about students’ mastery of all four number operations. Goodrow (1998) discovered that second graders using *Investigations* performed better than their peers on subtraction problems involving regrouping. Mokros et al. (1994) report no differences in a written test of computational skills. Likewise, Flowers (1998) shows no differences between groups with respect to performance gains on multiplication and division items among fourth graders.

*Investigations* students achieve greater accuracy than students using other curricula on word problems and on more complex calculations.

In all four studies, *Investigations* students performed better in solving word problems and complex calculations embedded in word problems than their counterparts using other curricula. They achieved greater levels of accuracy on these problems and showed more flexibility in their thinking. For example:

- *Investigations* students generated more solutions and more complex solutions (e.g., use of larger numbers and a longer series of operations) when asked to write number sentences resulting in a given number (Goodrow, 1998).

- *Investigations* students were more successful on word problems in which the use of a particular arithmetic operation was not specified (Mokros et al., 1994).

- *Investigations* students showed a better understanding of proportional reasoning (Flowers, 1998).

- *Investigations* students in second grade performed more accurately on a variety of word problems (Mokros et al., 1996).

*Investigations* students are less likely than other students to use standard algorithms and more likely to use their own strategies and algorithms to achieve accuracy.

All of the researchers noted differences between *Investigations* students and other students with respect to the degree to which they successfully used standard U.S. algorithms versus other algorithms and strategies. Consistently, *Investigations* students developed and used strategies that were well-matched to the particular problems they were solving. *Investigations* students rarely used standard algorithms, while children from other groups used standard algorithms often and with inconsistent accuracy. Mokros et al. (1996) found that students who used more varied strategies achieved greater accuracy.

Implementation of the *Investigations* curriculum has the greatest impact when students are encouraged to develop their own strategies and when teachers do not combine *Investigations* with a more traditional approach to teaching algorithms.

In the studies reported above, there were different levels of use of the *Investigations* curriculum, which was sometimes used in conjunction with other materials or programs. Goodrow (1998) concluded that a mixed approach, that of teaching traditional algorithms along with *Investigations*, was not as effective as using an *Investigations* approach in which students develop a variety of algorithms. Like Goodrow, Flowers (1998) and Mokros et al. (1996) found that children in the mixed conditions did not perform as well as students who were consistently using *Investigations*. This finding, and the reasons behind it, need further study.

References


Figure 1

Synopsis of Word Problems for Second Graders (Mokros et al., 1996)

Note: Actual problems used language appropriate for second graders.

- Determine how many pencils remain when you start with 66 and lose 29.
- Determine how many more beans in a group of 33 than in a group of 17.
- Combine 33, 17, and 5 beans to determine the total.
- Distribute 52 fish among 3 fish tanks (in any possible way).
- Distribute 52 fish among 3 tanks when one tank contains 13 fish.
- Determine ways of “jumping” on a number line by equal jumps that will allow you to land exactly on 24.
- Determine ways of “jumping” on a number line by equal jumps that will allow you to land exactly on 100.

Figure 2

Item Analyses of Accuracy on Word Problems (Mokros et al., 1996).
Figure 3
Synopsis of Interview Problems for Third and Fourth Graders (Mokros et al., 1994). Note: Actual problems used language appropriate for third and fourth graders.

Grade 3 Number Problems
• Determine how much candy there will be altogether if three people each get 39 pieces of candy.
• Determine the possible ages of four people in a family, if their ages together are 101 years. Then do the same problem given the mother is 37 years old.

Grade 4 Number Problems
• Estimate about how much 349 X 3 will be. Then figure it out exactly.
• Determine how many 60-pound dogs it would take to weigh the same as a 3,000-pound elephant.
• Determine ways of “jumping” on a number line, using same-sized jumps to land exactly on 20, to land exactly on 72, and to land exactly on 400. (You are starting at zero.) Determine how many jumps are needed in each case to arrive at the goal number.

Figure 4
Mean Number Correct on Written Calculation Test (Mokros et al., 1994)

Figure 5
Research on Fourth Graders’ Understanding of Multiplication (Flowers, 1998)