Pearson
Research Overview

Pearson Research Overview

Pearson Education is committed to using scientific, evidence-based methods in the development of its educational curricula. A research team, comprised of educational research methodologists, has been working with Pearson for seven years to integrate scientific research practices into the development of its curricula. Pearson also collaborates with regional education laboratories, universities, and private research companies to independently evaluate the effectiveness and usability of its curricula. These studies are designed to meet the rigorous standards of the *What Works Clearinghouse*.

Four phases of research are incorporated into the development of each new curriculum. The goal of establishing such extensive research methods is to ensure that every program enables all children to learn the skills and concepts they need for academic success. During the first phase of the research process, previous editions of the curricula are evaluated to determine best instruction and practices as demonstrated by scientific evidence. These practices will be incorporated into the current curricula to begin establishing a scientific research base.

During the second phase, the authors and researchers conduct extensive literature reviews on content, instructional practices, and education standards. The data is synthesized and embedded into the curricula.

During the third phase, formative research is conducted on the curricula under development. Classroom field tests investigate usability, teacher and student feedback, and preliminary curricula effectiveness. School administrators, content specialists, and classroom teachers systematically evaluate the curricula in development.

The final phase of research examines the implementation and effectiveness of the curricula. Independent, randomized control trial studies are conducted to provide scientific evidence of student achievement on standardized assessments. Implementation and best practices are documented throughout the study period to further contribute to the effectiveness of the curricula. Pearson believes that research needs to be ongoing with continual feedback to inform product revisions to meet student and teacher needs.

Investigations in Number, Data, and Space®: Program Background

Investigations in Number, Data, and Space® (referred to hereafter as Investigations) was developed at TERC by a team of curriculum developers and mathematics educators. The development of *Investigations* was funded in part by the National Science Foundation. The program was developed over an eight-year period (1990–1998) and field-tested in a variety of schools during the development period. Dr. Susan Jo Russell was the Principal Investigator of the National Science Foundation grant that funded the development of *Investigations*.

Investigations is designed for use as the primary source of mathematics instruction in the K–5 classroom. The four major goals of the curriculum are to: offer students meaningful mathematical problems, emphasize depth in mathematical thinking rather than superficial exposure to a series of fragmented topics, communicate mathematics content and pedagogy to teachers, and substantially expand the pool of mathematically literate students.

The *Investigations* program embodies the vision of the rigorous national standards for mathematics developed by the National Council of Teachers of Mathematics. The program is designed to help all elementary children understand the fundamental ideas underlying number and arithmetic, geometry, data, measurement, and algebraic thinking. The *Investigations* curriculum also embodies an approach based on years of research about how children learn mathematics. Each grade level consists of a set of separate units, each offering two to eight weeks of work. These units of study are presented through investigations that involve students in the exploration of major mathematical ideas, and may revolve around related areas, such as, addition and subtraction or geometry and fractions.

Approaching the mathematics content through investigations helps student develop flexibility and confidence in approaching problems, fluency in using mathematical skills and tools to solve problems, and proficiency in evaluating their solutions. Computational fluency of basic number combinations develops through understanding of number relationships and practice in games and activities. Students also build a repertoire of ways to communicate about their mathematical thinking, while their enjoyment and application of mathematics grows.

Investigations in Number, Data, and Space: Summative Research–2nd Edition

The *Investigations* program underwent a revision in 2007, and Pearson felt it was important to continue documenting the effectiveness of the program. Pearson contracted with Gatti Evaluation to conduct a longitudinal study of the *Investigations* program. The study commenced in 2007–08 with first and fourth-grade students, following them into second and fifth-grade respectively in 2008–09. This report summary presents the evaluation design and methods, an assessment of program implementation, student performance results, and a discussion of findings.

Study Design and Research Questions

The general purpose of the study was to rigorously evaluate the effectiveness of the *Investigations* program in helping elementary students attain critical math skills, to document curriculum implementation, and to explore how teachers and students reacted to the *Investigations* program.

The study employed a longitudinal, randomized controlled trial (RCT) design with the random assignment of teachers to treatment (*Investigations*) and comparison (non-*Investigations*) groups. That is, teachers within each participating school were *randomly assigned* to use either *Investigations* materials or their current school math curriculum. This study design was utilized in order to address all quality standards and criteria described in the What Works Clearinghouse (WWC) Study Review Standards (2008).

Specifically, the study addressed the following overarching evaluation questions:

1. Do *Investigations* students demonstrate significant learning gains in math during the study period?
2. How does the math performance of *Investigations* students compare to that of students using other math programs?
3. How do teachers implement the *Investigations* curriculum?
4. How did the math curriculum assigned to students impact their attitudes?

Participants and Settings

Gatti Evaluation recruited eight schools across five districts to participate in the first year of the study, including sites in AZ, MA, OR, and SC. Three schools from one district had to withdraw from the study in the second year (2008–09) to comply with a district-wide mathematics adoption. In addition, three other schools allowed teacher participation in the second year of the study to be voluntary, reducing the ability to maintain treatment and comparison conditions across both years of the study. Sixty percent of first grade students and 72% of fourth grade students remained in their assigned study conditions for the duration of the longitudinal study. Since the focus of this study was to evaluate the longitudinal impact of *Investigations* on student achievement, students who changed the study conditions, were not included in the final analytic sample. The final analytic sample was comprised of 39 teachers and 400 students. Table 1 provides demographic information on the participating schools.

Table 1. Demographics of Schools in Final Analytic Sample

District	School	Grades	State	Student Count	Percent Reduced Lunch	Percent Caucasian	Percent Hispanic/ Native American	Percent African American/ Caribbean	Percent Asian American
1	1	K-4	MA	478	64.6%	43.5%	43.9%	8.2%	4.4%
1	2	5-8	MA	441	63.4%	47.1%	31.4%	10.1%	11.4%
1	3	5-8	MA	658	57.9%	50.5%	36.0%	6.2%	7.3%
2	4	PK-5	SC	751	37.2%	50.2%	6.4%	41.3%	2.1%
3	5	K-5	AZ	630	40.6%	38.1%	42.2%	12.9%	6.8%
3	6	K-6	AZ	353	36.5%	43.6%	40.2%	11.0%	5.1%
3	7	PK-5	AZ	354	31.4%	47.5%	39.3%	7.9%	5.4%
4	8	K-6	OR	547	36.1%	79.2%	13.1%	3.5%	4.2%

Measures

Multiple measures were used to assess student achievement and program implementation. In order to measure program implementation and teacher perceptions, evaluators collected data through observations and interviews with math teachers. Math teachers also completed weekly self-report implementation logs. This background information provided researchers with a detailed data source on what was occurring in treatment and comparison classrooms in terms of math instruction and allowed researchers to identify areas of overlap in terms of content taught and activities. Evaluators also conducted biannual classroom observations and interviews with classroom teachers. The observation data provided critical insight into the nature of use and the effectiveness of the math materials used with treatment and comparison students.

Evaluators employed two student measures to assess changes in students' math skills over the course of the study. Teachers administered each assessment in fall 2007, spring 2008, and spring 2009. Evaluators selected the Group Mathematics Assessment and Diagnostic Evaluation (*GMADE*) as a norm-referenced assessment, and an evaluator constructed multi-step performance tasks. The *GMADE* assessment has broad visibility and acceptance in the field and demonstrates high technical merit. The assessments were given to all treatment and comparison students.

The *GMADE* is a standardized, nationally norm-referenced mathematics achievement test published by Pearson. The *GMADE* includes 9 levels that span Grades K–12. Each level has two parallel forms. Level one and level four form A was administered at baseline with form B administered at the end of the school year. The *GMADE* is not a timed test but generally takes between 60 and 90 minutes to administer. The *GMADE* offers a total math score, as well as three subtests. The subtests are *Concepts and Communication* (28 questions), *Operations and Computation* (24 questions), and *Process and Applications* (28 questions). These subtests address students' knowledge of mathematics facts and language, use of basic computational algorithms and operations, and the ability to solve problems presented in written form, respectively.

Students were also administered one multi-step performance task as part of the assessment battery. The performance tasks were constructed by the evaluator to assess the students' ability to reason and formulate strategies to solve unfamiliar real-world problems based on developmentally appropriate content. The baseline and end-of-second-year tasks were similar in format and context of problem posed to the students. However the end-of-second-year tasks were more advanced and involved in the content presented.

Student Performance Results

Results for *Investigations* Students

Gatti Evaluation determined that students who used *Investigations* demonstrated statistically significant gains in math achievement over the two-year study period. Moreover, significant gains in mathematics achievement were evidenced after just one year of implementation as well. Specifically, students using *Investigations* significantly improved in the areas of math concepts and communication, math computation and operations, and math process and applications.

Figures 1 and 2 present *Investigations* students' longitudinal gains in mathematics achievement, presented in grade equivalents. The average growth expected per year is one grade equivalent. The students in this study were post-tested one month prior to the end of their school year, so we would expect them to gain 1.8 grade equivalents over the course of two years. The early elementary (1st to 2nd grade) cohort did gain 1.8 grade equivalents in two years of implementation. The late elementary (4th to 5th grade) gained 3 grade equivalents during the same period, exhibiting more growth than is typically expected.

Figure 1. *Investigations* Early Elementary: *GMADE* Grade Equivalence

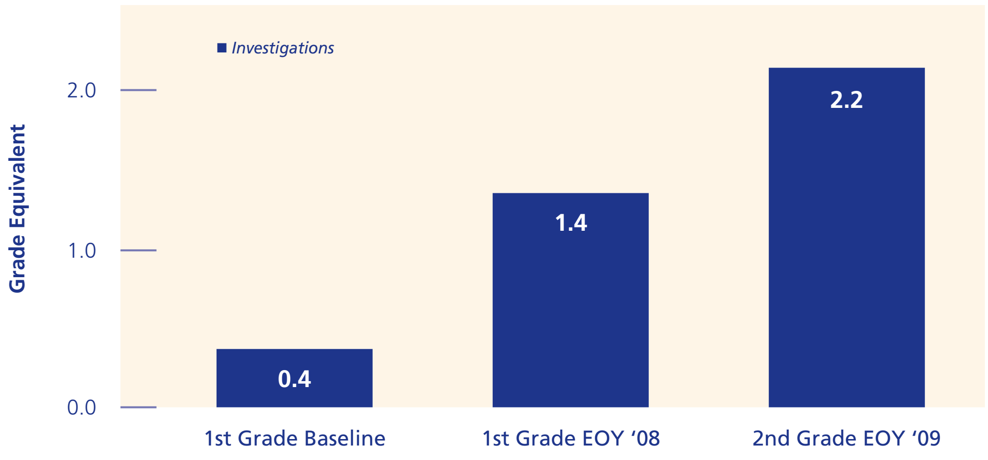
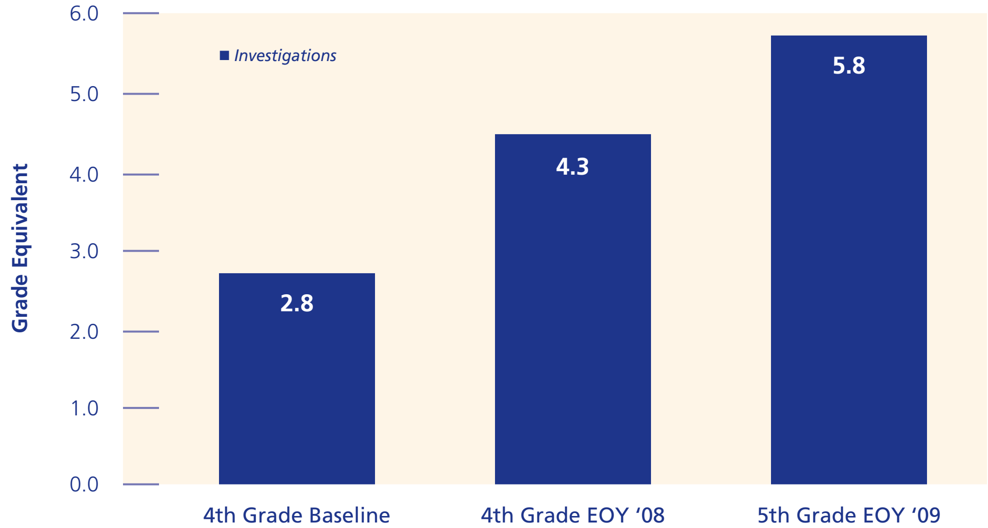


Figure 2. *Investigations* Late Elementary: *GMADE* Grade Equivalence



Subgroup Results: *Investigations* Students

Gatti Evaluation disaggregated the data by five key variables: English proficiency, ethnicity, Supplemental Educational Service (SES) as represented by eligibility for free/reduced lunch, gender, and mathematics ability (determined by the median for the national spring norm group). Evaluators found that all subgroups experienced significant gains in mathematics achievement, including: English proficient and non-English proficient, African-American, Caucasian, and Hispanic students, students receiving free/reduced lunch and those not receiving this aid, males and females, and students of low and high math ability. Tables 2 and 3 present the subgroup analyses, represented in grade equivalents, for the early and late elementary *Investigations* students.

Table 2. Subgroup Analyses for Early Elementary *Investigations* Students

GMADE Subpopulation	Baseline 1st Grade	EOY 1st Grade	EOY 2nd Grade	Growth
Lower achieving	0.3	1.3	2.0	1.7
Higher achieving	1.5	2.6	4.1	2.6
Reduced priced lunch	0.3	1.2	2.0	1.7
Full priced lunch	0.5	1.6	2.2	1.7
Not English proficient	0.1	1.2	2.2	2.0
English proficient	0.5	1.4	2.2	1.7
African American	0.2	1.2	1.8	1.6
Hispanic	0.2	1.3	2.2	2.0
Caucasian	0.5	1.5	2.3	1.8

Cell values represent grade equivalent scores transformed from sample or unadjusted *GMADE* total raw score means.

Table 3. Subgroup Analyses for Late Elementary *Investigations* Students

GMADE Subpopulation	Baseline 4th Grade	EOY 4th Grade	EOY 5th Grade	Growth
Lower achieving	2.4	3.5	5.0	2.6
Higher achieving	5.3	12.2	12.8	7.5
Reduced priced lunch	2.5	3.5	4.8	2.3
Full priced lunch	3.8	6.2	8.2	4.4
Not English proficient	3.0	4.3	6.2	3.2
English proficient	2.8	4.3	5.8	3.0
African American	1.8	3.5	4.8	3.0
Hispanic	2.5	3.6	4.8	2.3
Caucasian	3.3	4.5	6.2	2.9

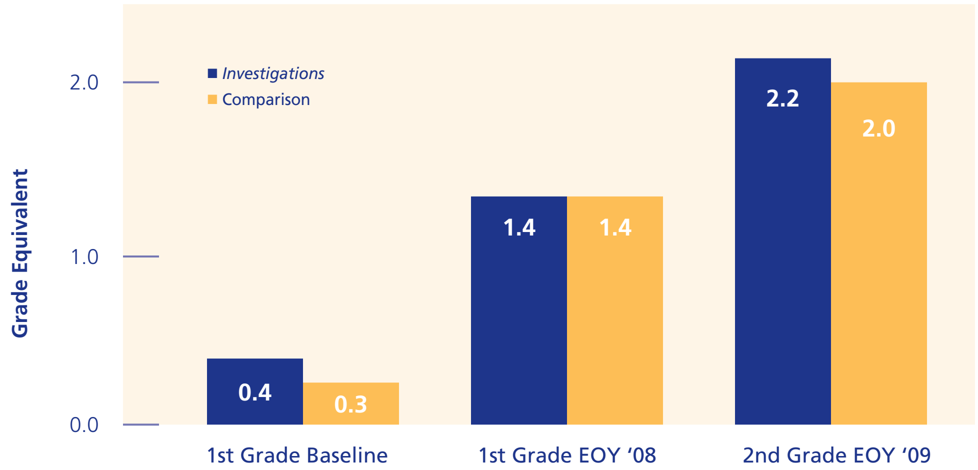
Cell values represent grade equivalent scores transformed from sample or unadjusted *GMADE* total raw score means.

Investigations vs. Other Math Programs

Evaluators conducted analyses comparing how *Investigations* students performed in comparison to students using other math programs. Results indicate that early elementary *Investigations* students performed similarly to their peers using other math programs on the *GMADE* and the performance task. However late elementary *Investigations* students statistically significantly outperformed their peers using other math programs on the *GMADE* and the performance task. The *GMADE* results can be seen in Figures 3 and 4.

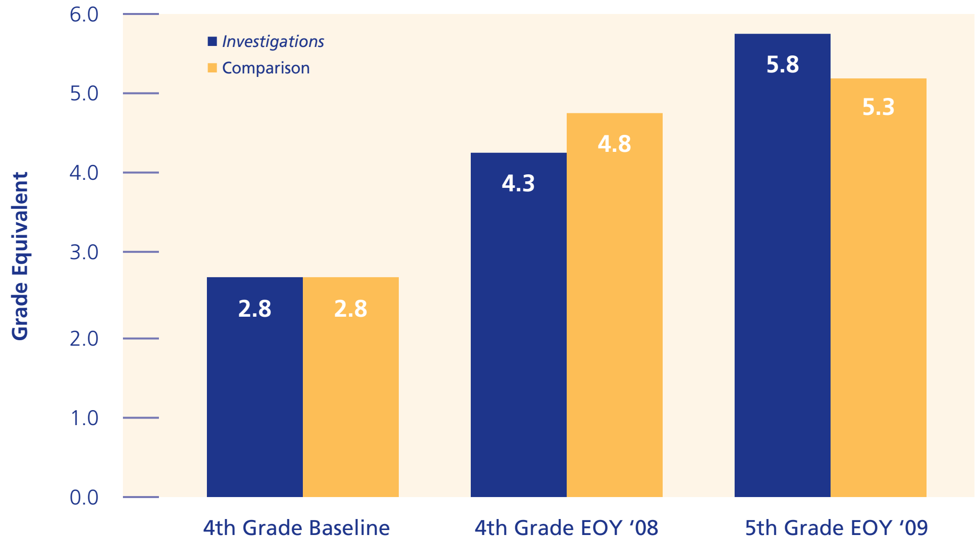
The results can be further analyzed by examining the *GMADE* subtests. Both the early and late elementary *Investigations* students statistically significantly outperformed their peers using other math programs on the Concepts and Communication subtest. The early elementary comparison and *Investigations* groups performed similarly on the Process and Applications subtest, and the comparison group significantly outperformed the early elementary *Investigations* group on the Operations and Computation subtest. The late elementary *Investigations* students significantly outperformed the comparison group on the Process and Applications subtest, and both groups performed similarly on the Operations and Computation.

Figure 3. Early Elementary *Investigations* vs. Comparison Group: *GMADE* Grade Equivalence



1. Bars represent grade equivalence corresponding to sample group means.
2. 1st grade baseline means are not statistically significantly different.
3. 1st grade end-of-year means are not statistically significantly different.
4. 2nd grade end-of-year means are not statistically significantly different.

Figure 4. Late Elementary *Investigations* vs. Comparison Group: *GMADE* Grade Equivalence



1. Bars represent grade equivalence corresponding to sample group means.
2. 4th grade baseline means are not statistically significantly different.
3. 4th grade end-of-year means are statistically significantly different, Cohen's d = –0.29, $p < 0.001$.
4. 5th grade end-of-year means are statistically significantly different, Cohen's d = 0.25, $p = 0.053$.

Subgroup Results: *Investigations* vs. Other Math Programs

The differences in subgroup populations were also calculated for group differences between the *Investigations* and comparison groups. The majority of analyses were not significant for the early elementary subpopulation, as evidenced by Table 4. However, the late elementary group demonstrated multiple significant findings, with the *Investigations* group outperforming the comparison group in each case, as evidenced by Table 5.

Table 4. Subgroup Analyses: *Investigations* vs. Other Programs

Early Elementary

2nd Grade Subpopulation	GMADE effect size[1,2]
Lower achieving	***
Higher achieving	***
Male	***
Female	***
Reduced priced lunch	***
Full priced lunch	-0.16
Not English proficient	***
English proficient	***
African American	0.28
Hispanic	***
Caucasian	***

Late Elementary

5th Grade Subpopulation	GMADE effect size[1,2]
Lower achieving	0.31
Higher achieving	0.47
Male	***
Female	0.42
Reduced priced lunch	0.20
Full priced lunch	0.35
Not English proficient	***
English proficient	0.27
African American	0.31
Hispanic	***
Caucasian	0.30

*** Indicates group means are not statistically significantly different.
1. effect size = estimated adjusted group difference / sample standard deviation
2. The average effect size for studies with small samples (i.e., less than 250 students) has been recently estimated at 0.27 standard deviations.

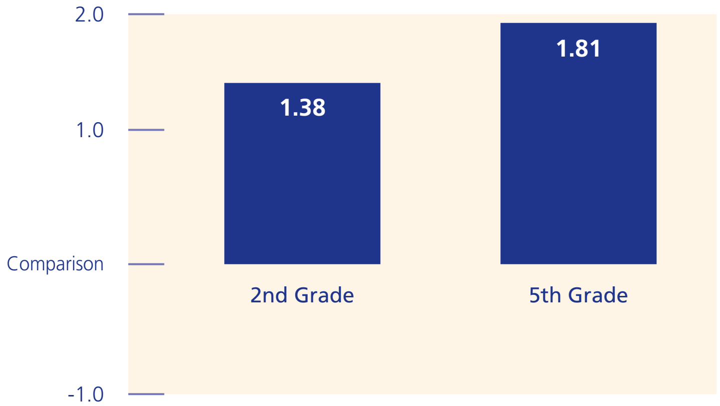
Program Fidelity of Implementation

Gatti Evaluation conducted biannual site observations and teacher interviews to document implementation. In addition, teachers were required to submit weekly self-report implementation logs. The data was triangulated to rate the level of implementation fidelity, with a quantitative scale ranging from 1 (inadequate) to 10 (excellent). The quantitative data collected from the site observations indicated that the *Investigations* teachers implemented the program at a high level of fidelity, with the vast majority of teachers receiving ratings in the *Good* to *Excellent* range. The site observation data also indicated that the study groups were instructing students and directing student learning at very similar levels, except at 5th grade where the *Investigations* teachers averaged in the *Excellent* range while the comparison teachers averaged a little lower in the *Good* range.

Student Attitudes

In addition to evaluating the effects of *Investigations* on student achievement, Gatti Evaluation also examined the program's effect on students' attitudes toward mathematics. Students' math attitudes were collected using a short self-report attitudinal survey. Analyses of the attitudinal surveys indicate that *Investigations* students demonstrated significantly more positive attitudes toward math and their mathematics curriculum as compared to their peers using other math programs. Figure 5 demonstrates the difference in student attitudes. The "comparison" line represents the comparison group adjusted mean score, and the bars above indicate the number of points higher that *Investigations* students rated their attitudes above their peers using other math programs.

Figure 5. *Investigations* vs. Comparison Group: Math Attitude Survey Score Difference



1. Bars represent end of school year 2008-09 model adjusted raw score group means. Groups are equated by baseline score, student demographic, and classroom information.
2. 2nd grade means are statistically significantly different, Cohen's d = 0.34, $p = 0.038$.
3. 5th grade means are statistically significantly different, Cohen's d = 0.34, $p < 0.001$.

Conclusion

The full study report, entitled *Pearson's Investigations in Number, Data, and Space Efficacy Study*, is available on the Pearson Education web site (www.pearsoned.com/new_research.htm). The breadth and depth of research that supports this program proves that *Investigations* is truly a scientific, evidence-based program with empirical data to prove its effectiveness in increasing student math achievement. In addition, teachers and students using *Investigations* found it to be engaging and motivating. The combination of these factors demonstrates the ability of *Investigations* to help students attain critical math skills in the elementary grades and prepare them for the skills needed in middle school mathematics.

Appendix

Foundational Research on 1st Edition

The following descriptions of studies from 1994–2000 are based on this original mathematics content and pedagogy. The research described used a variety of methodologies, including measures of student achievement and learning, student and teacher interviews, and research gathered from studies published in peer-reviewed journals. These studies form the theoretical support for the curriculum.

Impact Evaluation

TERC conducted an evaluation of the impact of *Investigations* on third- and fourth-graders' mathematical understanding (Mokros et al, 1994) during the 1992–1993 school year. The research was supported in part by the National Science Foundation. A comparison group design was used with a pre-and post-implementation assessment. The assessment included a written test stressing calculation skills, as well as hour-long individual interviews that examined students' understanding of concepts involving number, data, and geometry. The comparison group consisted of students who were participating in standard mathematics classes. Students from seven *Investigations* classrooms, along with students from five non-*Investigations* (also referred to as comparison) classes, were involved in the study. The comparison classrooms were matched on socioeconomic and student population characteristics to the *Investigations* classrooms. The classrooms were located in the state of Massachusetts with diversity in geographic setting (urban, suburban, rural), race, gender, background characteristics, first language, and mathematical skills. Eight students from each classroom were randomly selected to participate in the study.

The written assessment included items similar to those found on many standardized tests at these grade levels. A TERC interviewer administered the untimed test to small groups of students in each school. The interview tasks were also constructed by TERC assessment specialists and curriculum developers to reflect the goals stressed by the NCTM Standards and the *Investigations* curriculum. The interviews were conducted by TERC staff members. The tasks included content in the areas of number, geometry, and data. As part of solving these problems, students were asked to estimate, predict, use manipulatives and calculators, explain their strategies, and show how they were thinking through the use of drawings and constructions.

A three-point scoring rubric was designed to assess the interview items. The factors considered in analyzing students' responses included choice of an appropriate strategy to solve the problem, completeness and accuracy of the solution, clarity of explanation and/or demonstration of how to solve the problem, and flexibility in dealing with new constraints introduced into the task. The written test was scored correct or incorrect. The assessment members were blind to the student groups.

Results

The results of this study indicated that comparison and *Investigations* students increased significantly from pre- to post-implementation. There was no statistically significant difference between the two groups. However, the *Investigations* students made significantly greater gains than the comparison group in solving the interview problems. There were no interview tasks where only the comparison group showed significant gains, while there were six interview tasks in which the *Investigations* group showed gains.

The researchers concluded that the *Investigations* curriculum did not disadvantage students in solving "traditional" mathematics problems, and they were at an advantage when it came to solving tasks that involved more conceptual work. If a task needed a student to choose and carry out strategies based on a good understanding of number relationships and operations, the *Investigations* group had an advantage. The researchers concluded that as teachers make *Investigations* their own and become more familiar with how mathematics teaching and learning are changing, the impact of the curriculum on student understanding will be even greater.

Additional Research

A second study was commissioned for the 1994–1995 school year in several second-grade classrooms in the Boston area (Mokros et al, 1996). The researchers were particularly interested in examining how children using a different approach to learning number operations would make sense of a variety of problems involving number operations. How would *Investigations* children, who had not learned any standard algorithms, perform on number tasks relative to children who had learned standard algorithms and also had invented some of their own learning strategies? At the second-grade level, *Investigations* offers three month-long units that focus on numerical operations. Students were not taught algorithms at all, but were encouraged to develop their own strategies based on sound number sense and understanding of the meaning of the problem.

The researchers used a post-implementation, comparison group design to evaluate the effectiveness of this approach to learning/teaching operations. They used end-of-year individual interviews with 50 students. Thirty students from three different classrooms were in the *Investigations* group, and twenty students from two classrooms using a different curriculum were in the comparison group. The comparison group used a curriculum that emphasized both invented procedures and learning traditional algorithms.

The interview posed a series of work problems tapping children's understanding of subtraction, combining, and comparing, as well as their ability to decompose numbers and find factors. Children were encouraged to use manipulatives and representations and to explain how they were thinking as they solved the problems. They were asked to solve the problem using a different method, if they were having problems with the first strategy. The interviewer recorded children's actions, use of paper and pencil, use of manipulatives, and their verbalizations.

The problems were scored using an analytic scoring rubric with scorers kept blind to the identity of the students. The following questions were addressed:

- Was the child's answer accurate or inaccurate?
- If inaccurate, was it off by 1 or 2, off by tens or groups of tens, or off by some other number?
- What was the child's primary strategy?
- Did the child use manipulatives, mental strategies, standard algorithms, or a combination?
- Was the child able to explain his/her thinking clearly through talking, writing, or drawing?

Results

Students in the *Investigations* group achieved a higher accuracy level on the problems than the students in the comparison group. Out of 7 points total, the mean for *Investigations* students was 4.6 items correct, compared with 3.25 for the comparison group. *Investigations* students were significantly more accurate on three of the addition/subtraction problems. On factoring problems, there were not significant differences between the groups. Use of the standard algorithm, on problems where it could have been useful, was higher among students in the comparison group, however neither group was particularly successful using it. The findings also indicated that the comparison group did not tend to use manipulatives to solve problems, and when they did, they had a low success rate in comparison to the *Investigations* group.

The researchers believe that *Investigations* students choose strategies that work for them, that make sense to them, and with which they can be successful. As a group, the *Investigations* students showed greater flexibility than comparison group students in being able to choose from a wide range of strategies. The process of inventing strategies that make sense for the learner within the context of a given problem is an emphasis of the *Investigations* curriculum, and this emphasis was clearly reflected in the findings.

Third Study

A third study was conducted by Goodrow (1998) from TERC. The study examined: (a) the development of number sense and number representation by children in traditional, transitional, and constructivist second-grade mathematics classrooms; and (b) how different teaching approaches influence the way children deal with computation exercises.

Results

Results on tasks designed to explore number sense and number representation show that children in constructivist classrooms used a wider variety of representations to express their mathematical thinking. Their ways included the liberal recording of one's counting to solving a problem, the invention of number strings, and the use of negative numbers. When asked to solve addition and subtraction computation problems presented in vertical form and in horizontal form, children in traditional classrooms nearly always chose to use algorithm procedures taught at school. Some of these children found it necessary to rewrite problems presented horizontally in vertical form. In contrast, children in constructivist classrooms, who had not learned algorithmic procedures for addition and subtraction but, instead, relied on their own number sense, presented a higher number of correct responses through use of varied strategies which revealed understanding of number relations and of the properties of the decimal system.

Fourth Study

During this same period, Judith Flowers (1998) investigated elementary students' multiplicative and proportional reasoning as part of her dissertation work. Flowers found that much research indicates that the concepts of ratio and proportion are difficult for students to acquire. In conventional classrooms, the typical focus is on the repeated addition model for multiplication and the development of computational skills. That treatment of multiplication, however, does not always lead to a full understanding of the operation. This study explored the understanding of and reasoning about multiplication, division, and proportion tasks among students in three different instructional programs.

One reform program encouraged the use of reasoning procedures, another taught a variety of conventional procedures, and finally, a traditional program taught standard algorithms. The intent of the study was not to make comparisons among programs nor imply causal relationships. Rather, it was situated in three different contexts in order to increase the likelihood of including students who had different ways of thinking about the operations and divergent thinking patterns on proportional reasoning tasks.

The subjects were 137 fourth-grade students from three schools. A sample of eighteen students was individually interviewed on tasks involving reasoning about multiplicative and proportional relationships. Data were collected from January to June 1997, from pretests, posttests, interviews, and attitude surveys.

Results

It was hypothesized that students who were encouraged to use reasoning procedures for multiplication and division, that is, perform computational tasks based on an understanding of the operations and of number relationships, would have a more developed understanding of the operations and would be more successful at extending the knowledge to proportional reasoning tasks than students who were taught conventional algorithms only. The data analyses suggest that this hypothesis is true.

Students who were encouraged to use reasoning procedures for multiplication and division used more sophisticated reasoning and had greater success on proportional reasoning tasks. These students also used a wider variety of strategies to solve problems and displayed greater flexibility in strategy choice. Findings also suggest that students who used reasoning procedures had a more developed understanding of multiplication and division and of the relationship between the two operations. Finally, an emphasis on learning and practicing standard procedures did not appear to result in greater skill on computation.

The research described above demonstrates that scientific research was utilized to construct the theoretical foundation of *Investigations*. The *Investigations* program has been built upon best-practices and rigorously evaluated with multiple field tests.

Research on 1st Edition

While much of the research outlined above does support the scientific research base for the *Investigations* curriculum, and in some instances provides evidence of efficacy, the ARC Center Tri-State Student Achievement Study (2003) was the first formal efficacy study involving the complete *Investigations* program. In 2000, the ARC Center received funding from NSF to carry out a large-scale study of the effects of *Everyday Mathematics (EM), Investigations (IN),* and *Math Trailblazers (MT)* on student performance using state-mandated standardized tests in Massachusetts, Illinois, and Washington state.

The study collected ex-post facto standardized achievement data from schools that were documented using the three curricula listed above. The study also utilized survey data from the participating schools. The districts and schools were identified as curricula users by the program publishers. The survey collected 1999–2000 school year data for the grades in which state test data were available. Each school was determined to be eligible for participation using a set of standards developed by ARC. The final sample contained 742 schools. These schools were then matched to comparison schools that were not using any of the three study curricula on demographics such as race, gender, and SES.

Results

The principal finding in this study was that the students in the NSF-funded reform curricula consistently outperformed the comparison students: All significant differences favored the reform students; no significant difference favored the comparison students. These results held across all tests, all grade levels, and all strands. The data from this study indicated that the reform curricula improved student performance in all areas of elementary mathematics, including both basic skills and higher-level processes. Use of these curricula resulted in higher state test scores.

A follow-up to the ARC study resulted in Gatti Evaluations (2004) reanalyzing a portion of the data from the original study. The full report, titled *ARC Study Data Analyses*, can be found on the Pearson web site (www.pearsoned.com). This project analyzed data specifically for fourth-grade students in Massachusetts schools that used the *Investigations* program and their matched schools. Specifically, total math scaled scores, total raw scores, short-answer and multiple-choice items scores, as well as strand scores, were analyzed for subgroups of free or reduced lunch and ethnicity. The goal was to determine whether *Investigations* users were more likely to see greater effects on student math achievement than those comparison groups using a basal math program. The effect sizes indicated that the *Investigations* group over the comparison group was small to moderate, and the effect sizes for the sub-samples were of similar magnitude. If these results are generalizable, they would indicate that schools that adopt the *Investigations* reform mathematics program are likely to see higher mathematics achievement test scores than those using regular basal mathematics textbooks for all their students.

In summary, the 1st Edition Summative Research Results studies on *Investigations* indicated that:

- *Investigations* students do as well or better than students using other curricula in straight calculation problems involving basic facts and whole number operations.
- *Investigations* students have a better understanding of number and number relationships than students working with more traditional programs.
- *Investigations* works equally well with students at different achievement levels in mathematics.
- Students who use *Investigations* achieve greater accuracy on word problems and on more complex calculations than students in comparison classrooms.
- Students in schools fully implementing *Investigations* outperform students in schools not using *Investigations* (or other NSF-funded elementary mathematics curricula) on a high-stakes standardized test administered in Massachusetts.

References

Flowers, J.M. (1998). *A study of proportional reasoning as it relates to the development of multiplication concepts*. Unpublished doctoral dissertation, University of Michigan, Ann Arbor, MI.

Gatti, G. (2004). *Investigations in Number, Data, and Space ARC Study Data Reanalysis*. Gatti Evaluation

Gatti, G. (2009). *Pearson's Investigations in Number, Data, and Space Efficacy Study*. Gatti Evaluation

Goodrow, A.M. (1998). *Children's construction of number sense in traditional constructivist, and mixed classrooms*. Unpublished doctoral dissertation, Tufts University, Medford, MA.

Mokros, J., M. Berle-Carmen, A. Rubin, and T. Wright, 1994. *Full year pilot grades 3 and 4: Investigations in Number, Data, and Space*. Cambridge, MA: TERC.

Mokros, J, M. Berle-Carmen, A. Rubin, and K. O'Neill, 1996. *Learning operations: Invented strategies that work*. Paper presented at the Annual Meeting of the American Educational Research Association. New York, NY.

PearsonSchool.com
800-848-9500

ADV: 978-1-4182-9457-1 **(FOR REP ORDERS)**

PEARSON