A Synthesis of Empirical Research on Teaching Mathematics to Low-Achieving Students

Scott Baker; Russell Gersten; Dae-Sik Lee

The Elementary School Journal, Vol. 103, No. 1. (Sep., 2002), pp. 51-73.

Stable URL:
http://links.jstor.org/sici?sici=0013-5984%28200209%29103%3A1%3C51%3AASOERO%3E2.0.CO%3B2-8

The Elementary School Journal is currently published by The University of Chicago Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ucpress.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org
Fri Mar 23 14:25:43 2007

A Synthesis of Empirical Research on Teaching Mathematics to Low-Achieving Students

Scott Baker
Eugene Research Institute/University of Oregon

Russell Gersten
Instructional Research Group
Long Beach, CA

Dae-Sik Lee
Inchon National University of Education

The Elementary School Journal
Volume 103, Number 1
© 2002 by The University of Chicago. All rights reserved.
0013-5984/2003/10301-0003$05.00

Abstract

The purpose of this study was to synthesize research on the effects of interventions to improve the mathematics achievement of students considered low achieving or at risk for failure. Meta-analytic techniques were used to calculate mean effect sizes for 15 studies that met inclusion criteria. Studies were coded according to 5 categories of mathematics interventions, and effect sizes were examined on a study-by-study basis within each of these categories. Results indicated that different types of interventions led to improvements in the mathematics achievement of students experiencing mathematics difficulty, including the following: (*a*) providing teachers and students with data on student performance; (*b*) using peers as tutors or instructional guides; (*c*) providing clear, specific feedback to parents on their children's mathematics success; and (*d*) using principles of explicit instruction in teaching math concepts and procedures.

Recently, the National Research Council (Kilpatrick, Swafford, & Findell, 2001) convened a panel of experts to "synthesize the rich and diverse research" on mathematics learning in the elementary and middle school years, to "provide research-based recommendations for teaching . . . and curriculum for improving student learning," and to "identify areas where research is needed" (p. 3). The panel examined all types of research: experimental interventions, quantitative studies linking observed classroom interactions to growth in mathematics achievement, qualitative studies of classroom practice, comparative international studies of mathematics achievement, and the vast array of qualitative studies of the development of mathematical concepts and reasoning in students. In that formidable report the panel attempted to cut across the wide array of data sources and disciplines in order to draw conclusions about how mathematics instruction in U.S. schools can be improved.

Our goal in this article is far more modest. We synthesize data from one source: well-controlled experimental and quasi-experimental studies that assess the effects of interventions designed to improve the mathematics achievement of students considered low achieving or at risk for failure. All of the studies in this research synthesis meet standards of scientific rigor similar to those used in the recent synthesis of beginning reading research conducted by the National Reading Panel (2000).

As the National Research Council (Kilpatrick et al., 2001) aptly noted, "Experimental rigor often requires narrowing one's focus to a single feature of an instructional method or to a limited amount of mathematical content" (p. 25). The focus of many of the studies we reviewed was narrow, limited to the effects of an assessment system or a classroom organizational structure. In some cases, fairly subtle aspects of curriculum design were manipulated in order to assess effectiveness. Our objective was not to create a new vision of mathematics instruction for students with learning problems but rather to provide a dispassionate, systematic look at what has been learned over the past 20 years through controlled research in classroom settings.

We are not aware of previous quantitative syntheses investigating the effects of instruction on the mathematics achievement of students at risk of mathematics failure. Swanson, Hoskyn, and Lee (1999) used meta-analysis to investigate the effects of a variety of instructional interventions on students with diagnosed learning disabilities. Students with difficulties in mathematics were not included in Swanson et al.'s meta-analysis unless they also had an identified learning disability.

Swanson et al.'s investigation is relevant to our task, however, because it addressed instructional effects with a population of students with significant achievement problems. Overall, Swanson et al. found that the set of 18 instructional intervention studies that addressed mathematics most specifically had a mean effect size of .40 on measures of mathematics performance, which is considered a moderate effect (Cohen, 1988). In the same analysis, effect sizes on measures of reading comprehension, word recognition, and writing were all somewhat higher: .72, .57, and .63, respectively. And, in fact, effect sizes in mathematics were among the lowest reported. Another interesting finding was that only 10% of the intervention studies had a primary focus on mathematics.

An important methodological feature of the Swanson et al. (1999) meta-analysis is that the authors categorized studies primarily on the basis of the types of dependent measures used to determine effects. Consequently, there is a rich source of information on what effect general aspects of instruction had on a range of outcomes, but there is less information about the details of the instructional interventions that produced those effects.

The National Research Council (Kilpatrick et al., 2001) recently summarized the knowledge base on helping students learn mathematics. Although they did not focus on students experiencing serious difficulty learning mathematics, many of the suggestions they provide constitute sound instructional recommendations for students struggling with mathematics. A central recommendation in the report is that teachers should play a more active instructional role in helping their students build mathematical proficiency than they currently do. Active instruction is critical to "engaging students in the mathematical work, maintaining their focused involvement in it, and helping them take advantage of instruction to learn" (p. 331). Use of multiple instructional methods to achieve this goal is clearly endorsed. For example, the report suggests that there are times when math content should be constrained in ways that focus students' attention on specific learning goals, as well as times when problems should be represented in multiple ways with a variety of strategies for solving them. Thus, the report argues for a blend of focused, explicit instruction with the more open-ended problem-solving approach advocated, for example, in earlier versions of the National Council of Teachers of Mathematics (NCTM) (1989) standards.

Segments of mathematics instruction should target teaching students to generate explanations of math concepts in their own words and to justify the methods they use to solve problems. Focusing on student errors and misconceptions can also be an effective instructional method, especially when teachers anticipate predictable student errors and prepare in advance to use those errors to help students understand correct solutions.

Knowing how to teach math well to students with differing abilities seems to be much more important than having math teachers who possess strong backgrounds in mathematics (Ball, Lubienski, & Mewborn, 2001). What is less clear from the National Research Council Report (Kilpatrick et al., 2001) is how teachers can balance different instructional approaches in a comprehensive program. Also, unlike the current research synthesis, the council's report does not address how instructional methods are best adjusted for students experiencing serious difficulty learning mathematics.

We discovered a body of controlled experimental and quasi-experimental research on teaching mathematics to low-achieving students that includes reliable, valid outcome measures. We believe a synthesis of the findings from this research can shed light on effective instructional approaches for students with low mathematics achievement. Although the number of studies is small, the quality of the research is, in general, high. Well-conceptualized instructional approaches, measures of treatment fidelity, students randomly assigned to treatment and comparison conditions, and outcome measures that tapped possible effects in both the treatment and comparison conditions characterized the set of studies we reviewed.

Many of the studies have addressed well-defined problems in teaching mathematics to students who struggle with learning mathematical concepts and procedures. The specificity of these problems and how they were examined in controlled investigations can have great practical utility, especially if commonalties among the problems, as well as the solutions, can be discerned.

One purpose of synthesizing empirical research is to examine a given body of studies, searching for commonalties and ways to summarize them accurately and succinctly. For example, in their analysis of 16 studies on reciprocal teaching, Rosenshine and Meister (1994) concluded that, across the studies that used measures aligned with the intervention focus, results "were generally the same regardless of the number of strategies that were taught" (p. 507). The authors of the synthesis also began to articulate common features of effective reading comprehension approaches.

A second purpose of research syntheses can be to search for important differences among studies that share a common focus. Analyzing differences typically requires a more detailed explication of studies than summarizing commonalties. In Rosenshine and Meister's (1994) synthesis on reciprocal teaching, for example, effects were much greater when experimenter-developed comprehension measures were used than when standardized tests were used. These discrepancies led Rosenshine and Meister to compare details of the most commonly used standardized test with the most commonly used experimenter-developed test. Their analysis led to a hypothesis that the two types of tests differed on six important dimensions. The authors then provided plausible reasons for how students in reciprocal teaching may have benefited over comparison students on each of these dimensions when completing experimenter-developed comprehension measures, but not necessarily on standardized tests.

In the current synthesis, we used meta-analytic techniques (Cooper & Hedges, 1994) to present mean effect sizes for each study within certain categories of mathematics interventions. Because the small number of studies in each category virtually precluded the search for moderator variables, we examined variation in effect sizes on a study-by-study basis, without applying homogeneity tests or other statistical techniques. Using effect size as a common metric helps the reader easily discern the relative effectiveness of each approach. Rosenshine and Meister (1994) relied on meta-analytic techniques in a similar way to examine the empirical research base on reciprocal teaching. Our objective was to analyze findings from experimental research conducted in schools to improve the mathematics achievement of students struggling to learn math.

Method

All studies published from 1971 to 1999 that included specific instructional mathematics-based intervention strategies to improve the mathematics performance of low-achieving school-age students were included in the review. Our basic source for identifying relevant studies was a doctoral dissertation by Lee (2000), later contributing to a technical report (Lee, Kame'enui, & Gersten, 2000).

The following search procedures were used to locate mathematics intervention studies (Lee, 2000). Computer searches of the ERIC and PsycINFO databases were conducted to locate studies published from 1971 to 1999 that addressed mathematics interventions with students who were low in mathematics achievement. The following combinations of descriptors were used in this search: mathematics achievement, mathematics education, mathematics research, elementary education, secondary education, slow learners, underachievement, academically disadvantaged, math anxiety, low achieving, at risk, and learning problems. We examined bibliographies of research reviews in the area of learning disabilities (i.e., Maccini & Hughes, 1997; Mastropieri, Scruggs, & Shiah, 1991; Miller, Butler, & Lee, 1998; Swanson & Hoskyn, 1998; Swanson et al., 1996) for studies published during the same period but that may have been omitted from the computerized databases. Finally, we conducted a manual search of major journals in special, remedial, and elementary education.

This procedure resulted in the identification of 599 studies. Of this total, we selected 194 studies for further review based on the title, key words, and abstracts. From these 194 studies, 17 (9%) met the following criteria for inclusion in the analysis:

1. Only studies that provided math instruction, or structured opportunities for students to practice or apply classroom mathematics lesson objectives, were included. Math-related studies that examined the effects of test-taking strategies, or taught students computer programming, logic, or assessed the effects of inclusion and mainstreaming on mathematics achievement, for example, were excluded.
2. Math instruction must have lasted for a minimum of 90 minutes during the course of the intervention.
3. Only experimental or quasi-experimental intervention studies that employed group-design methods with a control group were included (i.e., no single-subject studies or case study research reports were included).
4. Quasi-experiments were included as long as one of three conditions was met: (*a*) posttest performance could be adjusted statistically by factoring in pretest performance on relevant outcome measures (Wortman & Bryant, 1985), or (*b*) the researchers in the original study adjusted posttest performance using appropriate analysis of covariance (ANCOVA) techniques. In addition, when posttest scores could not be adjusted statistically for pretest differences in performance, the original study documented that there

were no significant differences between groups at pretest on relevant measures of mathematics achievement. (Table 1 reports which studies were experimental and which were quasi-experimental. In some cases, teachers rather than students were randomly assigned to condition. These studies were considered quasi-experimental, and all were included because posttest performance could be adjusted to factor in pretest differences.)

5. Each study needed to include at least one mathematics performance or achievement measure. Studies that measured computation skill, math problem solving, understanding mathematics concepts, and other activities where students had to demonstrate mathematics proficiency in some way were included in the analysis. When experimenter-developed measures were used, reliability information on the measure needed to be reported. Studies that only measured students' attitude toward mathematics or self-concept were excluded from analysis.
6. Studies must have reported means and standard deviations, or *F*-values, so that effect sizes could be calculated (Cooper & Hedges, 1994). (Only one study was excluded based on this criterion.)

We excluded studies for the following reasons: (*a*) an experimental or quasi-experimental design was not used (33.3%), (*b*) there was insufficient information documenting that students were low achieving in mathematics (32.2%), (*c*) a mathematics intervention was not implemented or the intervention was not described with enough clarity for coding (17.0%), (*d*) outcome data for calculating an effect size were not reported (14.1%), (*e*) other reasons (3.4%) (e.g., the intervention was not of sufficient duration, only experimenter-developed measures were used and the reliability of those measures could not be verified).

Definition of Low Achieving

Students in these studies were identified as low achieving in mathematics on the basis of their performance on standardized or informal tests or by their placement in remedial mathematics classes. In some studies students were receiving Title I services in mathematics.

All studies provided operational definitions of low achieving. Typically the researchers relied on both teacher nomination and a measure of math performance. For example, Fuchs et al. (1997) asked teachers to select "two students whose mathematics performance was at or near the bottom of the class . . . but who had never been referred for special education" (p. 519). Next, the researchers administered a pretest, which systematically sampled math problems from the Tennessee mathematics framework for grades 1–6. The student with the lower score was considered low achieving.

Woodward and Baxter (1997a) included all students who scored below the thirty-fourth percentile on a standardized mathematics test (i.e., the Iowa Tests of Basic Skills, ITBS). A somewhat broader net was applied in the research of Fantuzzo and colleagues (Fantuzzo, Davis, & Ginsburg, 1995; Ginsburg-Block & Fantuzzo, 1997; Heller & Fantuzzo, 1993). In one case, Heller and Fantuzzo (1993) defined low achieving as "(*a*) scores below the 50th percentile on standardized mathematics achievement scales (based on the School District of Philadelphia's citywide norms), and (*b*) poor performance in mathematics as rated by classroom teachers" (p. 519). In this case, teacher nominations were used to confirm student performance data.

Students with identified learning disabilities were included in one-third of the studies in the analysis but in those studies only constituted a small percentage of the entire sample. Some researchers (e.g., Fuchs & colleagues) presented separate data for low-achieving students and students with learning disabilities. In those cases, we excluded students with disabilities from our analysis. In other cases, researchers did not disaggregate the data so it was impossible to separate the performance of students with and without identified disabilities.

Coding of Studies

We coded studies on a set of standard variables common in synthesizing instructional intervention research. Initial coding variables included the number of students in each condition, the procedures used to assign students to conditions, and the length of the intervention. In the second coding phase, we analyzed studies according to the dimensions of the interventions that are unique to this body of studies.

Describing the studies: Phase 1 coding. A total of 17 studies met the criteria for inclusion in the synthesis. These studies and major descriptive information are presented in Table 1. Initial coding involved determining the following information: (*a*) whether the study randomly assigned students to conditions or was a quasi-experiment, (*b*) the number of students per condition, (*c*) grade level(s) of students, (*d*) the ethnicity and income of students, (*e*) the length of intervention, and (*f*) how low achieving was operationally defined. We also listed all dependent variables and noted all reliability and validity data provided. Finally, we noted the pages in the article where the interventions were described so that we could use this information in the next phase of the coding process.

Unfortunately, most studies did not report student ethnicity. Of the studies that did report ethnicity, both studies by Cardelle-Elawar (1992, 1995) involved primarily Hispanic students. Three studies reported involving primarily African-American students (Fantuzzo et al., 1995). Our sense is that the remaining studies concerned primarily European-American populations.

Identifying independent variable(s): Phase 2 coding. The second phase of coding was our attempt to identify precisely the research question or questions addressed in each study. The senior authors developed the coding scheme for the set of studies over several months. The process was iterative. During the first reading of an article, we coded features of the intervention according to a developing set of broad categories (e.g., curriculum design, providing ongoing performance feedback to teachers and students, using data to generate specific instructional recommendations, the use of technology). We reviewed these codes and our notes and reread relevant sections of the study to pinpoint the precise research questions being addressed. This required rereading major sections of all the studies. Five of the 17 studies included multiple intervention groups and thus addressed multiple research questions. The senior authors confirmed all coding during the second phase of the coding process.

In our final analysis, we settled on five major categories that characterized the research questions and intervention approaches for the set of studies. These were (*a*) providing data and ongoing feedback to teachers and/or students about mathematics performance, (*b*) peer tutoring / peer-assisted mathematics instruction, (*c*) use of parents to support classroom mathematics instruction, (*d*) the use of explicit or teacher-facilitated instructional approaches, and (*e*) computer-assisted instruction.

We discuss each category in detail as we present the findings. In the appendix, the studies are listed by category along with the associated effect sizes. A few studies are included in more than one category because it was possible for a study to explore more than one research issue. For example, some studies examined the effects of peer tutoring as well as any "value-added" effects of strategies that encourage parental involvement. All of the studies included in more than one category involved three or more groups (e.g., two treatment groups and a comparison group). When studies included more than two groups in the overall analysis, we used orthogonal contrasts (Keppel & Zedeck, 1989) to calculate effect sizes. This was done to ensure that no statistical assumptions of independence were vio-

Table 1. Information on Studies Included in the Meta-Analysis

Study	Students Randomly Assigned to Conditions?	Grade Levels	Length of Intervention	Fidelity Measured?	Low Achieving (%)	Special Education (%)
Bottge & Hasselbring (1993)	No	9	Five 40-minute sessions	Yes	83	17
Cardelle-Elawar (1995)	Yes	3–8	Daily 50-minute lessons throughout school year	Yes	100	0
Cardelle-Elawar (1992)	Yes	6	6 weeks (daily)	Yes	100	0
Clairiana & Smith (1989)	Yes	11	Five 3-hour sessions over 5 weeks	No	100	0
Fantuzzo et al. (1995)	Yes	4 and 5	Two 45-minute sessions per week over 10 weeks	Yes	100	0
Fuchs et al. (1994)	No (but random assignment of teachers)	2–5	25 weeks (3 times/week)	Yes	100	0
Fuchs et al. (1995)	Yes	2–4	25 weeks	Yes	100	0
Fuchs et al. (1997)	Yes	2–4	22 weeks	Yes	100	0
Ginsburg-Block & Fantuzzo (1997)	Yes	4–5	10 weeks	Yes	100	0
Heller & Fantuzzo (1993)	Yes	4–5	Two 45-minute sessions/week for 8 months	Yes	100	0
Henderson & Landesman, (1995)	No	7	1 school year	No	100	0
Moore & Carnine (1989)	Yes	9, 10, 11	14 or 13 days	Yes	79	21
Schunk (1982)	Yes	4–5	Two 45-minute sessions on consecutive days	No	100	0
Woodward & Baxter (1997a)	No	3	1 school year	No	94	6
Woodward et al. (1999)	No	8–9	15 days	No	54	46

lated. Data-analysis procedures are described in the next section.

Data Analysis

Computation of effect size. Standard techniques for effect-size calculation are intended for studies with one experimental group and one comparison group. For studies with two groups, we used standard procedures for determining effect sizes (Cooper & Hedges, 1994). The basic index we used to calculate an effect size was Cohen's *d*, defined as the difference between the experimental and comparison group means divided by the pooled standard deviation (Cooper & Hedges, 1994). For calculating effect sizes when studies were combined for comparisons across categories, we followed procedures of Shadish and Haddock (1994), which weight each effect size by the number of students in each study.

For studies that reported both pretest and posttest scores, we calculated posttest effect sizes adjusting for pretest performance, using the first and the second equations (Wortman & Bryant, 1985).

$$\text{Adjusted Effect Size} = \text{Unadjusted } d - \text{Pretest Correction}, \quad (1)$$

$$\text{Pretest Correction} = (M_{\text{E[pretest]}} - M_{\text{C[pretest]}})/\text{SD}_{\text{pooled [pretest]}}, \quad (2)$$

where $M_{\text{E[pretest]}}$ = the mean of the experimental group at pretest, $M_{\text{C[pretest]}}$ = the mean of the comparison group at pretest, and $\text{SD}_{\text{pooled [pretest]}}$ is the pooled standard deviation at pretest. This technique is especially useful for quasi-experimental studies, or studies where the sample sizes are small and there are small differences in pretest scores between samples.

A somewhat controversial issue in calculating effect sizes is how to determine the appropriate number and types of comparisons to make when there are more than two groups in one study. Although standard procedures (Cooper & Hedges, 1994) indicate that each study should contribute only one effect size for each relevant category, we did not deem this approach appropriate for this set of studies. One-third of the studies addressed research questions that fit more than one category.

Use of orthogonal contrasts. Five studies had at least two groups receiving an experimental intervention, as well as a comparison group. (One of these studies had three intervention groups and a comparison group.) The experimental interventions were often subtle variations of one substantive intervention. For example, in one group the teachers received data on student performance, whereas in the other group the teachers received data on student performance as well as ideas on curriculum to use with particular students. In these cases it made sense to compute one effect size for each research question asked. If a study had three independent groups (i.e., two treatment groups and a comparison group), one could easily calculate three effect sizes: comparing the first group to the second, the first to the third, and the second to the third. However, this would violate assumptions related to independence of each effect size.

Conducting multiple comparisons this way provides redundant and potentially misleading information. Orthogonal contrasts, however, provide independent pieces of information (Keppel & Zedeck, 1989). They seemed the most appropriate and most elegant statistical approach to take. Thus, we conducted only orthogonal comparisons. These were determined by an in-depth reading of each study to determine what we viewed as the major research questions the author posed.

Effect size calculations for multiple dependent variables. To calculate effect sizes for the dependent variables, we followed standard procedures. If a study included more than one dependent measure of a similar construct (e.g., mathematics computation), the average of the measures was calculated and entered in the analysis.

We present results centered on each of the major categories. We approach each as a theme, in that, in each case, a series of researchers have addressed an instructionally important question. However, after careful consideration, we decided that one category was not relevant for a contemporary audience. After reviewing the computer-assisted instructional studies, we decided these studies addressed primarily efficacy of dated software (e.g., Bass, Ries, & Sharpe, 1986; Moore, 1988) and should be excluded. The types of software used 15 to 20 years ago are too dated in the context of contemporary instruction to be of much use. These studies are represented by the final category in the appendix. However, one study that used technology to provide precise feedback on student performance in mathematics (Clairiana & Smith, 1989) was still included under the category of provision of feedback. This study addressed questions that remain relevant for contemporary mathematics instruction.

Effect size calculations when the class is the unit of analysis. Two studies (Cardelle-Elawar, 1995; Fuchs et al., 1997) used the class or a subset of students in the class as the unit of analysis. Although this is a legitimate means of data analysis, the standard deviations presented in these studies are much smaller than studies where the student is the unit of analysis. Left uncorrected, this would tend to inflate effect sizes substantially. As a statistical correction, we multiplied the standard deviation of the unit of analysis presented by the square root of the class or unit size (Hopkins, 1982; Peckham, Glass, & Hopkins, 1969). This resulted in standard deviations that closely approximated what they would have been if students had been the unit of analysis.

Coding of dependent variables. We coded each dependent measure as either a computation measure or a general mathematics achievement measure. The purpose of computation measures was to determine how accurately—and in many cases, how quickly—students could add, subtract, multiply, and divide numbers (including numbers with decimals). These measures invariably were closely aligned with the focus of the instructional interventions. Some of the computation measures included the Math Operations Test (Fuchs et al., 1997), the Curriculum-Based Computation Rate (Ginsburg-Block & Fantuzzo, 1997), and the Math Skill Test (Schunk, 1982).

The assessments coded as general mathematics achievement measures included an array of mathematical topics. All included some computation, but they also included word problems and items assessing understanding of mathematical concepts such as equivalence of fractions. Published general math achievement measures included the ITBS total score in mathematics (Cardelle-Elawar, 1992; Woodward & Baxter, 1997b). General mathematics achievement measures developed by researchers connected to the target study included the Comprehensive Math Test (Fuchs et al., 1997). These measures were not completely aligned with the focus of the intervention. Rather, they assessed whether the intervention improved general mathematical competence.

We calculated separate effect sizes for computation and general achievement. Following standard procedures, however, if a study included two computation measures, we calculated the mean of the two measures in determining the computation effect size.

Results and Discussion

In this section we discuss the most important findings in each of the four major study categories. Our goal is to pinpoint what each set of studies says about means to improve the mathematical performance of low-achieving students. In particular, we look at how consistently the practice enhanced performance and the overall magnitude of the effect.

The number of studies meeting our initial criteria for inclusion in the synthesis was 17. With the exclusion of the two computer software studies by Bass et al. (1986) and Moore (1988), the total number of studies in the analysis was 15. These 15 studies were coded into four major categories and generated 39 independent effect sizes. These effect sizes ranged from −.59 to 1.49 (see App.).

In quantitative syntheses using effect sizes, it is common to present the overall mean effect size for all the intervention studies in the analysis. We decided not to do this because the mean effect size would have little meaning in the context of the range of questions being addressed in the set of studies. It would merely present the aggregate of an array of interventions ranging from curriculum redesign, to tutoring, or to improving the information teachers receive on student progress. Interpretation would be difficult.

Providing Data or Recommendations to Teachers and Students

In four studies, students and/or their teachers were provided with specific data on student performance. In some of these studies, the computer also generated recommendations about what types of problems to work or how many problems to work on a given topic. The comparison group in these four studies either was provided with no performance feedback or with such limited feedback that a relevant contrast between the experimental and comparison group was meaningful.

Six comparisons were conducted across these four studies. All of these comparisons were orthogonal—that is, they contributed independent sources of information (Keppel & Zedeck, 1989) (see Table 2). In five comparisons (involving all four studies) students received information on their effort or performance in solving mathematics problems or received recommendations from the teacher or computer regarding the number of problems they should work in a given time. In many cases, a computer provided this feedback to teachers as well as students. The overall effect size for these five comparisons was .57 (unweighted, .71), and the confidence interval indicated that the overall effect was significantly different from zero. This is a moderate effect and the second largest mean effect size we found in this synthesis.

One study in this category was conducted by Fuchs, Fuchs, Hamlett, Phillips, and Bentz (1994), and we describe it in more depth than the others. It is, in our view, the most complex study in the set but also the richest. Fuchs et al. included two experimental groups and one comparison group. On a computer, students in the two experimental groups took weekly tests on items that reflected state content standards. The software created individualized graphs depicting each student's performance over time. Performance graphs were given to both teachers and students. Teachers also received a performance summary of all students in the class. In the comparison condition, teachers used their own techniques for monitoring student progress. Thus, the major difference between the experimental groups and the comparison group was providing teachers and students with weekly information on student progress in mathematics.

One of the unique features of this study was the difference between the two experimental groups. In the more complex experimental group, teachers also received computer-generated recommendations on what content to teach the full class in upcoming lessons based on the aggregate performance data. Recommendations regarding which students to group together for small-group instruction were provided (based on their individual Curriculum-Based Measurement math tests). Teachers also received a listing of computer lessons to use with individual students and suggestions on how to use peer tutoring (Greenwood, Delquadri, & Hall, 1989) to provide students with practice and reinforcement on concepts and skills with which they were struggling.

We conducted two orthogonal contrasts from this study. The first investigated the effect of providing weekly progress information to students and their teachers. Here, we contrasted the average performance of the two experimental groups with the performance of the comparison group. This effect size was .29, indicating a small effect on mathematics achievement when teachers systematically monitored the progress of their low-achieving students and graphed this information for themselves and their students.

In the second contrast we investigated the effect of providing teachers with instructional recommendations based on the progress-monitoring data. For this question we examined the difference between the two experimental groups. The effect size for this comparison was .51, which is considered moderate. The two effect sizes suggest that perhaps merely providing teachers with data on student performance may not be as beneficial as the combination of providing data and then making specific instructional recommendations to address problem areas identified in current student performance.

In summary, the small number of studies and comparisons supported the practice of providing feedback to students and recommendations to them on what problems to work. Advantages and limitations of providing teachers with feedback versus providing them with feedback accompanied by instructional recommendations are speculative and need further investigation. Computers seem to be valuable in generating mathematics problems for students to work, being able to target areas where more practice is needed, and providing performance feedback to students and teachers along with specific recommendations.

Peer-Assisted Learning

A major focus of research on mathematics instruction for low-achieving students has been on development and evaluation of strategies and structures that enable students to provide each other with feedback and support. Six studies addressed this topic. The overall aggregate effect size for this category appears in Table 2.

We believe there are several reasons for the relatively heavy emphasis on studying how students can work with each other to learn mathematics. While working on mathematics problems independently, students who are uncertain about problem solving often want to ask questions about what to do. A teacher cannot be available to each individual student to address questions and uncertainty. Oftentimes peers can provide the answer, or (if taught to do so) provide suggestions that help students solve the problem themselves. Related to the idea of peers working together is the finding that success in mathematics requires considerable task persistence (Kolligian & Sternberg, 1987). Researchers investigating peer tutors have stressed that peer tutoring is likely to encourage low achievers to persist in their work.

Peer-assisted learning interventions invariably led to positive effects on student achievement. The average effect size was .62, with a median of .51. Effect sizes ranged from .34 to 1.26. With the exception of the one outlier (Heller & Fantuzzo, 1993), which had the weakest comparison group, the effects were reasonably consistent and in the low to moderate range. The majority of the studies (five of the six) examined effects on computation; only two studies included a measure of general math achievement. The magnitude of effect sizes was greater on computation than general math ability. The average effect size on computation problems was .62 (weighted), which was significantly greater than zero. On general math achievement, the two effect sizes were .06 and .40, producing a weighted mean of .29 that was not significantly different from 0. It is safest to conclude that the peer-assisted learning approaches demonstrated a consistent, moderately strong positive effect on the computation abilities of low achievers. To date, it is unclear how helpful peer tutoring might be in other areas of mathematics.

In discussing the individual studies, it is helpful to disentangle the two major streams of research that have been con-

TABLE 2. Summary of Aggregate Effect Sizes for Four Categories of Studies Reviewed

Category	No. of Comparisons	*N*	Effect Size		95% Confidence Interval	
			Average	Weighted	Lower	Upper
Providing data or recommendations to teachers and students:						
Providing student with data/information[a]	5	200	.71	.57	.27	.87
Providing instructional recommendations to teachers	1	20	.51	.51	−.38	1.40
Peer-assisted learning[a]	6	302	.62	.66	.42	.89
Explicit teacher-led and contextualized teacher-facilitated approaches:						
Explicit instruction[a]	4	485	.65	.58	.40	.77
Teacher-facilitated instruction and practice	4	203	−.04	.01	−.26	.29
Concrete feedback to parents	2	81	.43	.42	−.02	.87

[a]Effect size is significantly greater than zero.

ducted on this topic. Lynn Fuchs and her colleagues have conducted the first. Fantuzzo and colleagues conducted the second. There are many similarities between the two approaches as well as some subtle differences. Both rely on students working in pairs (dyads) as opposed to the groups of four to six that typify cooperative learning. In each approach, curriculum-based assessment data are used to pair students and to determine the content of the tutoring sessions.

Both approaches are reciprocal in nature. In other words, students alternate between the role of tutor and tutee. (The one exception is the study by Fuchs, Fuchs, Phillips, Hamlett, and Karns [1995], which used a more traditional tutor-tutee model.) Role reciprocity became a critical innovation as Fuchs et al.'s peer tutoring approach evolved and may be a central feature in its success. Both Fuchs et al. and Fantuzzo et al. use a tightly structured format for the tutoring sessions. Students are carefully trained in tutoring procedures. Both approaches employ prompt cards and procedures for tutors to use as they help fellow students work through a series of problems.

There are two major differences between the approaches. The research of Fantuzzo and colleagues focused only on computation, whereas the work of Fuchs and colleagues included a broader range of mathematics topics and involved a more complex feedback system for the peer tutors to use. In the Fuchs et al. system, each tutor gives the partner feedback on each step in the math problem attempted. Although tutors are provided with a specific step-by-step strategy for approaching each type of problem, when the tutees encounter difficulty, tutors are encouraged to construct explanations in their own words.

Once students master the basics of the peer tutoring procedures (typically by the third week of the program), they are taught several teaching strategies to use as they work with their partners. These strategies were adapted from the research of Hiebert and Wearne (1993). The goal of the tutoring strategies is to help students contextualize problems. Tutors are encouraged to represent abstract mathematical quantities with visuals or manipulatives. Tutors are also encouraged to discuss solution strategies with their partner.

In summary, the use of peers to provide feedback and support is consistently supported by research as a means to improve computational abilities and is a promising means to enhance problem-solving abilities.

Explicit Teacher-Led and Contextualized Teacher-Facilitated Approaches

Seven studies investigated the effects of instructional practices on math achievement of low achievers. Effect sizes for these studies are presented in Table 2. Although this is the largest set of studies, it is still an extremely small number of research studies on the broad topic of effective mathematics instruction for low-achieving students. This paucity reflects the general lack of experimental field research in mathematics (Kilpatrick et al., 2001).

The seven studies fell into two general categories: those involving explicit instruction in mathematics, and those that stressed contextualized approaches. Three studies, each contributing one comparison to the overall effect size, investigated an approach involving explicit instruction in mathematics. In these studies, the manner in which concepts and problem solving were taught to students was far more explicit than is typical.

Three studies investigated the effects of math instruction that emphasized the context of the mathematics problems in which teaching occurs and stressed conceptual understanding over procedural compliance and accuracy. Following this approach, teaching emphasized real-world applications of mathematical principles. One study (Woodward, Baxter, & Robinson, 1999) had both an explicit instruction group and a contextualized instruction group and consequently was included in both contrasts.

Explicit instruction. Two of the interventions followed principles articulated in Engelmann and Carnine (1982), which are often referred to as direct instruction. Direct instruction involves teaching rules, concepts, principles, and problem-solving strategies in an explicit fashion. This includes providing a wide range of examples of the principle or concept and providing extensive review and discriminative practice. Two other studies in this category, both by Cardelle-Elawar (1992, 1995), used an approach that was also explicit but focused on teaching generic problem-solving strategies using a specific set of problems. The method of problem solving used was derived from the cognitive research of Mayer (1987).

The aggregate weighted effect size in the explicit instruction category was .58 (unweighted = .65). Individual effect sizes ranged from .32 (Cardelle-Elawar, 1995) to 1.1 (Moore & Carnine, 1989). The 95% confidence interval was .40 to .77, indicating that the effect was statistically significant. This result indicates that, overall, the approaches that used explicit instruction had a positive, moderately strong effect on the mathematics achievement of at-risk students.

Cardelle-Elawar (1992, 1995) investigated the effects of the Mayer problem-solving approach on a general measure of mathematics achievement. The overall effect for the problem-solving approach was .55 (weighted), which was statistically significant. Woodward et al. (1999) examined effects of the intervention on a measure of computation involving decimals. Moore and Carnine (1989) assessed students' proficiency and understanding of elementary problems involving decimals and proportions (e.g., What is 25% of 22?). Both the Woodward et al. and Moore and Carnine studies involved curricula designed according to Engelmann and Carnine's (1982) principles of direct instruction. The weighted effect size was .80, also statistically significant.

Cardelle-Elawar's instructional approach focused on strategy instruction. Specifically, teachers extensively modeled how students should ask themselves a series of questions when faced with mathematics problems. Instruction emphasized story problems. Deciphering the vocabulary used in problems was stressed. In part, this was because the students in these studies were primarily English language learners. After working to understand the vocabulary used, students learned to determine if the necessary information was available to solve the problem, and if so, how to organize the problem. Then they proceeded step by step through the calculations phase to arrive at the correct answer.

After extensive teacher modeling of all components, students worked individually on similar problems, under close supervision and monitoring by the teacher. Teachers provided feedback to students that closely followed the question-asking strategy students learned from the teacher modeling. At the end of each lesson, students were required to formulate in their own words what they learned that day as a means of processing the key principles and strategies presented in the lesson.

In the explicit instruction approaches in the Moore and Carnine (1989) and Woodward et al. (1999) studies, concepts and operations involving ratios and proportions were taught following the principles of instructional design articulated by Engelmann and Carnine (1982). The curriculum stressed presenting a wide range of examples to demonstrate each concept, extensive practice, and cumulative review of previously taught material. The curriculum used in the Moore and Carnine (1989) study explicitly taught students strategies for discerning relevant from irrelevant material in the problems. Students also worked on problems independently, errors were corrected quickly by the teacher, and students reviewed the strategies taught or the relevant mathematical principle. There was virtually no emphasis on students being able to verbalize their problem-solving strategies in this approach, unlike the approach Cardelle-Elawar used. In Woodward et al. (1999), students in one of the two groups were taught concepts involving decimals using the identical theoretical framework as the Moore and Carnine study.

Contextualized instruction and practice. Four recent studies (all conducted in the 1990s) investigated what we have labeled contextualized teaching. A characteristic of this approach was for some or all of the instruction in the experimental group to stress real-world applications and, at least to some extent, to focus on understanding underlying concepts of authentic problems. The treatments in studies in this category sought to teach students about mathematical thinking, arguing that a more vigorous emphasis on concept development was critical to mathematics success and would lead not only to a deeper understanding of math but to computational proficiency as well. These studies were all influenced by the framework of mathematics instruction developed by the NCTM in 1989 and recently revised drastically (NCTM, 2000). One study (Woodward et al., 1999) investigated explicit instruction versus the NCTM framework and thus fell into both the explicit instruction and contextualized instruction categories.

The overall effect size of studies in contextualized instruction was .01, essentially zero. In other words, students in the comparison groups scored as well as students in the experimental groups. In two of the four studies, the overall effect size favored the experimental group, and in two of the four it favored the comparison group. The effect size in the study by Henderson and Landesman (1992) was .18, indicating that the effect of contextualized instruction was small. Henderson and Landesman administered both a general achievement measure (concepts and applications) and a computation measure. Both measures produced small, positive effects (.22 and .14, respectively).

The study by Bottge and Hasselbring (1993) produced the largest positive effect for this group of studies (effect size = .48). This was one of the most interesting and thoughtful studies in the synthesis because of the creative nature of the design and instruction and the way the authors tried to assess the effect of the intervention on a range of dependent measures.

Bottge and Hasselbring did two things that are worth discussing. First, during a 5-day baseline phase they taught both experimental and comparison students mathematics skills that would help the researchers better understand the subsequent effects of the intervention. The intervention compared math learning via contextualized instruction—the presentation of an authentic problem delivered via videodisc—versus learning from instruction delivered through a more traditional focus on word problems.

The second noteworthy feature was that dependent measures were carefully constructed to assess learning and transfer. Two dependent measures were closely aligned to each of the teaching approaches—that is, one measure was aligned to instruction the experimental group experienced, and one measure was closely aligned to instruction the comparison group experienced. Two measures assessed transfer, one of the most troublesome areas of special education research and practice. Students who received contextualized instruction scored higher on a contextualized word problem test and on transfer measures than students in the control group.

The two studies by Woodward and his colleagues (Woodward & Baxter, 1997a; Woodward et al., 1999) were the closest to Bottge and Hasselbring in trying to investigate how contextualized instruction affects mathematics performance. The overall effect sizes for both of these studies were negative. Woodward and Baxter (1997a) compared students receiving instruction based on the Everyday Mathematics (Bell, Bell, & Hartfield, 1993) program developed at the University of Chicago to students receiving traditional instruction following a basal mathematics program. This program de-emphasized computation and included many problems with relevance to students' everyday experiences. Problems were selected so that students could not capitalize on "key words" as shortcuts to understanding the concepts. "Students are encouraged to use or develop a variety of number models which display relevant quantities (e.g., total and parts; start-change-end; quantity-quantity-difference) to be manipulated in solving these problems" (Woodward & Baxter, 1997a, p. 376). Estimation is actively encouraged. An array of mathematical games is an integral part of the curriculum. Unlike the approach taken in the Moore and Carnine (1989) study, where relatively brief problems were taught first and students gradually built to complex problems, students in the Woodward and Baxter (1997a) study were confronted early on with complex multi-step problems. The goal was to develop in-depth conceptual understanding.

The overall effect size, based on performance on a general measure of math achievement (total mathematics score on the ITBS), was −.24, meaning that students in the basal group (the comparison condition) performed better than students in the NCTM group (the experimental condition). A strength of the Woodward and Baxter (1997a) study was the use of a standardized measure of mathematics achievement. The study also included a student interview to assess conceptual understanding, but this was only administered to a small subsample, precluding statistical analyses.

In the subsequent study in 1999, Woodward and his colleagues used a stronger comparison group—explicit instruction—and investigated the effects of instruction using two experimenter-developed measures of computation involving decimals. Students solved computation problems with and without the use of a calculator. The effect sizes were nearly identical, −.59 and −.58, indicating that students in the comparison group did better than students in the experimental group, and the effect in both cases was moderate. This finding is not too surprising given that the computational focus could be expected to favor the explicit instruction group.

Woodward et al. (1999) also presented the results of an individual student interview scored quantitatively, which significantly favored students in the NCTM group. The authors indicated that this measure showed how NCTM instruction benefits students' conceptual knowledge of mathematics compared to explicit instruction. We did not include this interview in our synthesis, however, because it was not clear how the measure was scored, and its reliability was uncertain. Also, it was not clear whether the interviewers were blind as to which group students were in when they were interviewed.

Instruction for the conceptual, or contextualized, condition in the Woodward et al. (1999) study emphasized the development of conceptual understanding through the use of visual representations and physical manipulatives (e.g., wood block rectangles). Lessons from *Mathematics in the Mind's Eye* (Bennett, Maier, & Nelson, 1988) served as a basis for daily instruction. Students were taught to develop visual representations of problems using pie chart diagrams and wood block rectangles, squares, or cubes. According to Woodward et al. (1999, pp. 17–18), "Instruction was intended to provide much greater depth in initial decimal concepts than these students had received in the past." Links between fractions and decimals were stressed.

In summary, the studies involving contextualized mathematics instruction present a complex puzzle of findings, open to multiple interpretations. These studies have furthered the understanding of how instruction focusing on concept development compares to other approaches, such as explicit instruction in fundamentals and problem solving. Further studies in this area should be conducted, using, as Woodward et al. (1999) did, techniques such as high-quality control groups, and employing, as Bottge and Hasselbring (1993) did, a range of dependent measures of learning and transfer. Researchers should also attempt to combine the best aspects of contextualized instruction with other approaches such as explicit instruction with the same group of students.

Regarding explicit instruction in mathematics, effects were consistently positive. Approaches ranged from very explicit instruction in mathematics operations, with extensive and carefully crafted practice, to approaches that focused on the explicit teaching of strategies students needed in order to understand the content of story problems.

Providing Parents with Information about Student Successes

Fantuzzo and colleagues (Fantuzzo et al., 1995; Heller & Fantuzzo, 1993) conducted studies in this category. In both cases, providing information to parents was assessed as an "add on" to a reciprocal peer-tutoring program. Both studies focused on improving computational skills and used measures of math computation. The studies produced identical effect sizes of .42. Although on the surface this seems like a moderate effect, it is not statistically significant. The "low cost" of the intervention is still impressive, however, and warrants a closer look at the studies and further empirical investigations.

The parent support intervention was "designed to enhance the parent's role as supporter and motivator of students' academic effort and success" (Fantuzzo et al., 1995, p. 274). It involved regular home-school contacts (by note or telephone) that described examples of students' efforts and successes in mathematics. All examples were positive, focusing on what the student had learned or accomplished or on activities on which the student had worked particularly hard. Messages focused on instances of students showing academic initiative and task persistence. The purpose of these contacts was to encourage parental celebrations of students' successes in mathematics. It is important to emphasize that in these interventions the parent's role was not that of a math teacher but rather of a knowledgeable supporter to her/his children in their efforts to work hard on learning mathematics.

These studies suggest that providing parents with information on their children's mathematics accomplishments and encouraging parents to celebrate those accomplishments with their children can result in improved student achievement. Moreover, this parent-support technique is remarkably easy to implement, and the positive effect has been replicated. The approach has, at times, worked in conjunction with other instructional efforts such as peer tutoring.

Summary and Conclusions

The set of 15 studies provides some ideas about ways to improve the performance of low-achieving students in mathematics. Although this is not a large body of research, four findings are consistent enough to be considered components of best practice. Other findings are more tentative, based on only a few studies.

One consistent finding is that providing teachers and students with specific information on how each student is performing seems to enhance mathematics achievement consistently. This practice has been recommended for many years, yet the extent to which it is implemented is unclear. The effect of such practice is substantial, raising scores, on average, by .68 SD units.

A second consistent finding represents an important strand in contemporary research. Using peers as tutors or guides enhances achievement. Research shows that the use of peers to provide feedback and support improves low achievers' computational abilities and holds promise as a means to enhance problem-solving abilities. If nothing else, having a partner available to provide immediate feedback is likely to be of great benefit to a low achiever struggling with a problem. A crucial feature of this approach is that the topics being covered are ones on which curriculum-based measurement data suggest areas where a student needs extra practice and support.

Third, providing clear, specific feedback to parents of low achievers on their children's successes in mathematics seems to have the potential to enhance achievement, although perhaps only modestly. More research needs to be conducted before firm conclusions are drawn. Advantages of this approach are that it is relatively easy to implement and can lead to other long-range benefits in school-home communication. The two relevant studies suggest that the feedback should (*a*) be specific, objective, and honest and (*b*) detail successes (or relative successes) as opposed to failures or difficulties.

Fourth, in terms of curricula, a small body of research suggests that principles of direct or explicit instruction can be useful in teaching mathematical concepts and procedures. This includes both the use of strategies derived from cognitive psychology to develop generic problem-solving strategies and more classic direct instruction approaches where students are taught one way to solve a problem and are provided with extensive practice. With the latter approach, concepts involving fractions, ratios, or decimals are presented using a wide range of examples.

There is less clarity about the benefits of contextualized approaches, where the teacher serves primarily as a facilitator as students work through real-world examples of mathematical concepts and discuss alternative solutions to problems with their peers or teachers. A small positive effect was found when students worked out complex, real-world problems only after they had been clearly and explicitly taught the underlying foundational mathematical concepts (e.g., Bottge & Hasselbring, 1993). Thus, low achievers seem not to do well at authentic problem solving and discussion of mathematical concepts without solid preparation in the underlying mathematical foundations.

At this point, however, it is premature to draw strong inferences on the effectiveness of this recently developed approach. For one thing, the mean effect size of the set of studies is essentially zero, indicating that there is no clear trend in the findings. In addition, the concept of contextualized instruction is an emerging one, and we did not find strong coherence in the approaches used in the four studies.

There are other plausible explanations for the overall ineffectiveness of the four studies that used a contextualized instruction approach. All four were quasi-experiments, so differences between the experimental and comparison groups unrelated to the intervention may have influenced the outcome. Also, three of the four studies involved older students (grades 7–9), which was a far larger percentage than we found in the other categories. In addition, only one of the four studies included a measure of implementation fidelity. Not only is this a lower percentage of measured fidelity than in the other categories, but because contextualized instruction is not yet a well-defined approach, lack of assessment of what actually occurred during lessons may have led to erratic implementation and thus to fewer effects on student learning than would have been obtained with expert implementation.

Examined as a group, the instructional studies seem to support the position taken by the National Research Council (Kilpatrick et al., 2001), which argues for a mix of explicit instruction in procedures and ample opportunity to apply procedures to open-ended problems with real-world relevance. Manzo (2001, p. 1) commented that the report emphasized that, "while both computational skills and a deep understanding of math concepts are essential parts of a complete mathematics education . . . other elements, including problem-solving and reasoning abilities, as well as an awareness of the relevance of math in everyday life, are also necessary for mathematical proficiency." Furthermore, the panel also suggested that algebraic principles should be built into the curriculum beginning in the early grades. Earlier research (e.g., Gersten & Carnine, 1984) suggests that this is a wise course of action for at-risk students.

It is unfortunate that such a limited number of controlled research studies address means for improving the mathematics knowledge of students who are considered low achievers. By limiting our synthesis to studies that were well controlled, we did not address some of the subtle, intricate, and profound issues in the teaching of mathematics raised, for example, by Ball (1990, 1993, 1995), Griffin, Case, and Siegler (1994), and Hiebert and Wearne (1993). Yet we find in these studies a burgeoning sense of the concept of mathematical proficiency, that is, "the integrated attainment of conceptual understanding, procedural fluency, strategic competence, adaptive reasoning, and productive disposition" (Kilpatrick et al., 2001, p. 313).

Each of the researchers cited in the instructional/curricula strand grappled with this issue (Bottge & Hasselbring, 1993; Cardelle-Elawar, 1992, 1995; Henderson & Landesman, 1995; Woodward & Baxter, 1997a), albeit using different language and coming at it from different traditions. There are too few studies on curriculum and instruction to allow inferences to be drawn; there is a pressing need for well-designed research on this topic. Even the National Research Council report (Kilpatrick et al., 2001) is elusive, at best, as to how to design and orchestrate instruction for students with chronic problems in mathematics. But the report's admonition that "sound research can help guide the design of effective mathematics instruction" (p. 24) reinforces the important point that strong research studies are not only the best way to answer questions about the effects of specific approaches, but they should also play a substantive role in shaping better methods to teach mathematics.

The National Research Council (Kilpatrick et al., 2001, p. 26) also suggests that "high-quality research should play a central role in any effort to improve mathematics learning. That research can never provide prescriptions, but it can be used to help guide skilled teachers in crafting methods that will work in their particular circumstances." Designing instruction so that even students with chronic problems in mathematics can succeed and develop a solid conceptual understanding is a formidable challenge. The authors of the National Research Council report note "it can be challenging to draw scientifically sound conclusions from a selected set of observations. In contrast, experimental methods . . . establish stronger bases for drawing conclusions, although even these conclusions have important limitations and qualifications" (Kilpatrick et al., 2001, p. 25).

Some of the qualifications cited in the National Research Council report seem valid. Others do not apply to the set of studies we reviewed. For example, the report suggests that effects found in controlled studies may not apply to the real world of classrooms. Yet most of the studies in our synthesis were conducted in school classrooms or tutoring settings that exist in Title I programs. The report's authors note that "most published studies in education confirm the predictions made by the investigators" (Kilpatrick et al., 2001, p. 26). This was not always the case in the studies we reviewed. We believe our synthesis provides suggestions that can serve as initial steps in the improvement, and perhaps the ultimate transformation, of the teaching of mathematics for low achievers.

Appendix

Studies Included in Meta-Analysis, by Category

Category 1: How Effective Is Providing Data and Feedback to Teachers and Students?

Effectiveness of Providing Students with Information and/or Data

Clairiana, R. B., & Smith, L. J. (1989). *Progress reports improve students' course completion rate and achievement in math computer-assisted instruction.* (ERIC Document Reproduction Service No. ED 317 170) (Effect size = .39)

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., & Bentz, J. (1994). Classwide curriculum-based measurement: Helping general educators meet the challenge of student diversity. *Exceptional Children,* **60,** 518–537. (Effect size = .29)

Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, C. L., Katzaroff, M., & Dutka, S. (1997). Effects of task-focused goals on low-achieving students with and without learning disabilities. *American Educational Research Journal,* **34,** 513–543. (Effect size = .25)

Schunk, D. H. (1982). *Efficacy and skill development through social comparison and goal setting.* (ERIC Document Reproduction Service No. ED 222 279) (Effect size = 1.31)

Effectiveness of Providing Instructional Recommendations to Teachers

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., & Bentz, J. (1994). Classwide curriculum-based measurement: Helping general educators meet the challenge of student diversity. *Exceptional Children,* **60,** 518–537. (Effect size = .51)

Category 2: How Effective Are Peer-Assisted Learning Formats?

Fantuzzo, J. W., Davis, G. Y., & Ginsburg, M. D. (1995). Effects of parent involvement in isolation or in combination with peer tutoring on student self-concept and mathematics achievement. *Journal of Educational Psychology,* **87**(2), 272–281. (Effect size = .47)

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., & Bentz, J. (1994). Classwide curriculum-based measurement: Helping general educators meet the challenge of student diversity. *Exceptional Children,* **60,** 518–537. (Effect size = .51)

Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, C. L., Katzaroff, M., & Dutka, S. (1997). Effects of task-focused goals on low-achieving students with and without learning disabilities. *American Educational Research Journal,* **34,** 513–543. (Effect size = .40)

Fuchs, L. S., Fuchs, D., Phillips, N. B., Hamlett, C. L., & Karns, K. (1995). Acquisition and transfer effects of classwide peer-assisted learning strategies in mathematics for students with varying learning histories. *School Psychology Review,* **24,** 604–620. (Effect size = .34)

Ginsburg-Block, M., & Fantuzzo, J. W. (1997). Reciprocal peer tutoring: An analysis of "teacher" and "student" interactions as a function of training and experience. *School Psychology Quarterly,* **12**(2), 134–149. (Effect size = .69)

Heller, L. R., & Fantuzzo, J. W. (1993). Reciprocal peer tutoring and parent partnership: Does parent involvement make a difference? *School Psychology Review,* **22,** 517–534. (Effect size = 1.49)

Category 3: How Effective Are Explicit Teacher-Led and Contextualized Teacher-Facilitated Approaches?

Effectiveness of Explicit Teacher-Led Instruction

Cardelle-Elawar, M. (1992). Effects of teaching metacognitive skills to students with low mathematics ability. *Teaching and Teacher Education,* **8**(2), 109–121. (Effect size = .32)

Cardelle-Elawar, M. (1995). Effects of metacognitive instruction on low achievers in mathematics problems. *Teaching and Teacher Education,* **11,** 81–95. (Effect size = .61)

Moore, L. J., & Carnine, D. W. (1989). A comparison of two approaches to teaching ratio and proportions to remedial and learning disabled students: Active teaching with either basal or empirically validated curriculum design material. *Remedial and Special Education,* **10**(4), 28–37. (Effect size = 1.1)

Woodward, J., Baxter, J., & Robinson, R. (1999). Rules and reasons: Decimal instruction for academically low achieving students. *Learning Disabilities Research and Practice,* **14,** 15–24. (Effect size = .59)

Effectiveness of Contextualized Teacher-Facilitated Approach

Bottge, B., & Hasselbring, T. S. (1993). A comparison of two approaches for teaching complex, authentic mathematics problems to adolescents in remedial math classes. *Exceptional Children,* **59,** 556–566. (Effect size = .48)

Henderson, R. W., & Landesman, E. M. (1995). Effect of thematically integrated mathematics instruction on students of Mexican descent. *Journal of Educational Research,* **88,** 290–300. (Effect size = .18)

Woodward, J., & Baxter, J. (1997a). The effects of an innovative approach to mathematics on academically low-achieving students in mainstreamed settings. *Exceptional Children,* **63**(3), 373–388. (Effect size = −.24)

Woodward, J., Baxter, J., & Robinson, R. (1999). Rules and reasons: Decimal instruction for academically low-achieving students. *Learning Disabilities Research and Practice,* **14,** 15–24. (Effect size = −.59)

Category 4: Can Parents Be Used to Enhance the Math Achievement of Their Children?

Fantuzzo, J. W., Davis, G. Y., & Ginsburg, M. D. (1995). Effects of parent involvement in isolation or in combination with peer tutoring on student self-concept and mathematics achievement. *Journal of Educational Psychology,* **87**(2), 272–281. (Effect size = .44)

Heller, L. R., & Fantuzzo, J. W. (1993). Reciprocal peer tutoring and parent partnership: Does parent involvement make a difference? *School Psychology Review,* **22,** 517–534. (Effect size = .42)

Category 5: How Effective Is Computer Instruction Using Software from the 1980s?

Bass, G., Ries, R., & Sharpe, W. (1986). Teaching basic skills through microcomputer assisted instruction. *Journal of Educational Computing Research,* **2**(2), 207–219. (Effect size = .15)

Moore, B. M. (1988). Achievement in basic math skills for low-performing students: A study of teachers' affect and CAI. *Journal of Experimental Education,* **57**(1), 38–44. (Effect size = .21)

Note

This research was supported in part by a grant from the Texas Education Agency and in part from a contract from the Research to Practice Division of the Office of Special Education Programs, U.S. Department of Education, awarded to American Institutes for Research. We wish to acknowledge the assistance in data analysis and coding provided by Janet Otterstedt, Jonathan Flojo, and Joyce Smith-Johnson. Jennifer Palmer and Janet Otterstedt provided valuable assistance in preparing the manuscript. Finally, we would like to thank Doug Carnine and Bob Slavin for their helpful feedback on an earlier draft of this article.

References

Ball, D. L. (1990). Reflections and deflections of policy: The case of Carol Turner. *Educational Evaluation and Policy Analysis,* **12,** 247–259.

Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *Elementary School Journal,* **93,** 373–397.

Ball, D. L. (1995). Blurring the boundaries of research and practice. *Remedial and Special Education,* **16,** 354–363.

Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 433–456). Washington, DC: American Educational Research Association.

Bass, G., Ries, R., & Sharpe, W. (1986). Teaching basic skills through microcomputer assisted instruction. *Journal of Educational Computing Research,* **2**(2), 207–219.

Bell, M., Bell, J., & Hartfield, R. (1993). *Everyday mathematics.* Evanston, Il: Everyday Learning Corporation.

Bennett, A., Maier, E., & Nelson, T. (1988). *Math in the mind's eye: Modeling rationals.* Salem, OR: Math Learning Center.

Bottge, B. A., & Hasselbring, T. S. (1993). A comparison of two approaches for teaching complex, authentic mathematics problems to adolescents in remedial math classes. *Exceptional Children,* **59,** 556–566.

Cardelle-Elawar, M. (1992). Effects of teaching metacognitive skills to students with low mathematics ability. *Teaching and Teacher Education,* **8**(2), 109–121.

Cardelle-Elawar, M. (1995). Effects of metacognitive instruction on low achievers in mathematics problems. *Teaching and Teacher Education,* **11,** 81–95.

Clairiana, R. B., & Smith, L. J. (1989). *Progress reports improve students' course completion rate and achievement in math computer-assisted instruction.* (ERIC Document Reproduction Service No. ED 317 170)

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2d ed.). Hillsdale, NJ: Erlbaum.

Cooper, H., & Hedges, L. V. (Eds.). (1994). *Handbook of research synthesis.* New York: Russell Sage.

Engelmann, S., & Carnine, D. (1982). *Theory of instruction.* New York: Irvington.

Fantuzzo, J. W., Davis, G. Y., & Ginsburg, M. D. (1995). Effects of parent involvement in isolation or in combination with peer tutoring on student self-concept and mathematics achievement. *Journal of Educational Psychology,* **87,** 272–281.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., & Bentz, J. (1994). Classwide curriculum-based measurement: Helping general educators meet the challenge of student diversity. *Exceptional Children*, **60**, 518–537.

Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, C. L., Katzaroff, M., & Dutka, S. (1997). Effects of task-focused goals on low-achieving students with and without learning disabilities. *American Educational Research Journal*, **34**, 513–543.

Fuchs, L. S., Fuchs, D., Phillips, N. B., Hamlett, C. L., & Karns, K. (1995). Acquisition and transfer effects of classwide peer-assisted learning strategies in mathematics for students with varying learning histories. *School Psychology Review*, **24**, 604–620.

Gersten, R., & Carnine, D. (1984). Direct instruction mathematics: A longitudinal evaluation of low-income elementary school students. *Elementary School Journal*, **84**, 395–407.

Ginsburg-Block, M., & Fantuzzo, J. W. (1997). Reciprocal peer tutoring: An analysis of "teacher" and "student" interactions as a function of training and experience. *School Psychology Quarterly*, **12**, 134–149.

Greenwood, C. R., Delquadri, J. C., & Hall, R. V. (1989). Longitudinal effects of classwide peer tutoring. *Journal of Educational Psychology*, **81**, 371–383.

Griffin, S. A., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual prerequisites for first formal learning of arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 24–49). Cambridge, MA: MIT Press.

Heller, L. R., & Fantuzzo, J. W. (1993). Reciprocal peer tutoring and parent partnership: Does parent involvement make a difference? *School Psychology Review*, **22**, 517–534.

Henderson, R. W., & Landesman, E. M. (1992). *Mathematics and middle school students of Mexican descent: The effects of thematically integrated instruction* (Research Report 5). Santa Cruz: University of California, National Center for Research on Cultural Diversity and Second Language Learning.

Henderson, R. W., & Landesman, E. M. (1995). Effect of thematically integrated mathematics instruction on students of Mexican descent. *Journal of Educational Research*, **88**, 290–300.

Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal*, **30**, 393–425.

Hopkins, K. D. (1982). The unit of analysis: Group means versus individual observations. *American Educational Research Journal*, **19**, 5–18.

Keppel, G., & Zedeck, S. (1989). *Data analysis for research designs: Analysis of variance and multiple regression/correlation approaches*. New York: Freeman.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Kolligian, J., & Sternberg, R. J. (1987). Intelligence, information processing, and specific learning disabilities: A triarchic synthesis. *Journal of Learning Disabilities*, **20**, 8–17.

Lee, D. (2000). *A meta-analysis of mathematics interventions reported for 1971–1998 on the mathematics achievement of students identified with learning disabilities and students identified as low achieving*. Unpublished doctoral dissertation, University of Oregon, Eugene.

Lee, D., Kame'enui, E., & Gersten, R. (2000). *A synthesis of the instructional research in mathematics for students with learning disabilities and low achieving students: An extension of the Swanson-Hoskyn instructional components analysis approach*. Eugene: University of Oregon.

Maccini, P., & Hughes, C. A. (1997). Mathematics interventions for adolescents with learning disabilities. *Learning Disabilities Research and Practice*, **12**, 168–176.

Manzo, K. K. (2001, January 31). Forget math feud, take broader view, NRC Panel urges. *Education Week*, pp. 1, 12.

Mastropieri, M. A., Scruggs, T. E., & Shiah, S. (1991). Mathematics instruction for learning disabled students: A review of research. *Learning Disabilities Research and Practice*, **6**, 89–98.

Mayer, E. R. (1987). *Educational psychology*. Boston: Little, Brown.

Miller, S. P., Butler, F. M., & Lee, K. (1998). Validated practices for teaching mathematics to students with learning disabilities: A review of literature. *Focus on Exceptional Children*, **31**(1), 1–24.

Moore, B. M. (1988). Achievement in basic math skills for low-performing students: A study of teachers' affect and CAI. *Journal of Experimental Education*, **57**(1), 38–44.

Moore, L. J., & Carnine, D. W. (1989). A comparison of two approaches to teaching ratio and proportions to remedial and learning disabled students: Active teaching with either basal or empirically validated curriculum design material. *Remedial and Special Education*, **10**(4), 28–37.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.

National Council of Teachers of Mathematics. (2000). *Standards 2000 project*. [On-line]. Reston, VA: NCTM. Available: http://standards.nctm.org/document/index.htm.

National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Washington, DC: National Institute of Child Health and Human Development.

Peckham, P. D., Glass, G. V., & Hopkins, H. D. (1969). The experimental unit in statistical analysis. *Journal of Special Education,* **3,** 337–349.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research,* **64,** 479–530.

Schunk, D. H. (1982). *Efficacy and skill development through social comparison and goal setting.* (ERIC Document Reproduction Service No. ED 222 279)

Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper & L. V. Hedges (Eds.), *Handbook of research synthesis* (pp. 261–321). New York: Russell Sage.

Swanson, H. L., & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research,* **68,** 277–321.

Swanson, H. L., Hoskyn, M., & Lee, C. (1999). *Interventions for students with learning disabilities: A meta-analysis of treatment outcomes.* New York: Guilford.

Swanson, H. L., McMahon, C. M., Sachse-Lee, C., Hoskyn, M., Restori, A., Ruffner, J., McLaughlin-Cheng, E., Carson, C., Kraemer, R., O'Shaughnessy, T., & Murray, B. (1996). *Synthesis of intervention research for students with learning disabilities* (Preliminary Report for the U.S. Department of Education, grant H023E40014). Riverside: University of California—Riverside.

Woodward, J., & Baxter, J. (1997a). The effects of an innovative approach to mathematics on academically low-achieving students in mainstreamed settings. *Exceptional Children,* **63,** 373–388.

Woodward, J., & Baxter, J. (1997b, April). *Making the task meaningful: Technology, literacy, and the instructional needs of students with learning disabilities.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Woodward, J., Baxter, J., & Robinson, R. (1999). Rules and reasons: Decimal instruction for academically low-achieving students. *Learning Disabilities Research and Practice,* **14,** 15–24.

Wortman, P. M., & Bryant, F. B. (1985). School desegregation and black achievement: An integrative review. *Sociological Methods and Research,* **13,** 289–324.

http://www.jstor.org

LINKED CITATIONS

You have printed the following article:

A Synthesis of Empirical Research on Teaching Mathematics to Low-Achieving Students
Scott Baker; Russell Gersten; Dae-Sik Lee
The Elementary School Journal, Vol. 103, No. 1. (Sep., 2002), pp. 51-73.
Stable URL:
http://links.jstor.org/sici?sici=0013-5984%28200209%29103%3A1%3C51%3AASOERO%3E2.0.CO%3B2-8

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

References

Reflections and Deflections of Policy: The Case of Carol Turner
Deborah Loewenberg Ball
Educational Evaluation and Policy Analysis, Vol. 12, No. 3. (Autumn, 1990), pp. 247-259.
Stable URL:
http://links.jstor.org/sici?sici=0162-3737%28199023%2912%3A3%3C247%3ARADOPT%3E2.0.CO%3B2-L

With an Eye on the Mathematical Horizon: Dilemmas of Teaching Elementary School Mathematics
Deborah Loewenberg Ball
The Elementary School Journal, Vol. 93, No. 4. (Mar., 1993), pp. 373-397.
Stable URL:
http://links.jstor.org/sici?sici=0013-5984%28199303%2993%3A4%3C373%3AWAEOTM%3E2.0.CO%3B2-F

Effects of Task-Focused Goals on Low-Achieving Students with and without Learning Disabilities
Lynn S. Fuchs; Douglas Fuchs; Kathy Karns; Carol L. Hamlett; Michelle Katzaroff; Suzanne Dutka
American Educational Research Journal, Vol. 34, No. 3. (Autumn, 1997), pp. 513-543.
Stable URL:
http://links.jstor.org/sici?sici=0002-8312%28199723%2934%3A3%3C513%3AEOTGOL%3E2.0.CO%3B2-F

http://www.jstor.org

Direct Instruction Mathematics: A Longitudinal Evaluation of Low-Income Elementary School Students
Russell Gersten; Doug Carnine
The Elementary School Journal, Vol. 84, No. 4. (Mar., 1984), pp. 395-407.
Stable URL:
http://links.jstor.org/sici?sici=0013-5984%28198403%2984%3A4%3C395%3ADIMALE%3E2.0.CO%3B2-V

Instructional Tasks, Classroom Discourse, and Students' Learning in Second-Grade Arithmetic
James Hiebert; Diana Wearne
American Educational Research Journal, Vol. 30, No. 2. (Summer, 1993), pp. 393-425.
Stable URL:
http://links.jstor.org/sici?sici=0002-8312%28199322%2930%3A2%3C393%3AITCDAS%3E2.0.CO%3B2-D

The Unit of Analysis: Group Means versus Individual Observations
Kenneth D. Hopkins
American Educational Research Journal, Vol. 19, No. 1. (Spring, 1982), pp. 5-18.
Stable URL:
http://links.jstor.org/sici?sici=0002-8312%28198221%2919%3A1%3C5%3ATUOAGM%3E2.0.CO%3B2-4

Reciprocal Teaching: A Review of the Research
Barak Rosenshine; Carla Meister
Review of Educational Research, Vol. 64, No. 4. (Winter, 1994), pp. 479-530.
Stable URL:
http://links.jstor.org/sici?sici=0034-6543%28199424%2964%3A4%3C479%3ARTAROT%3E2.0.CO%3B2-I

Experimental Intervention Research on Students with Learning Disabilities: A Meta-Analysis of Treatment Outcomes
H. Lee Swanson; Maureen Hoskyn
Review of Educational Research, Vol. 68, No. 3. (Autumn, 1998), pp. 277-321.
Stable URL:
http://links.jstor.org/sici?sici=0034-6543%28199823%2968%3A3%3C277%3AEIROSW%3E2.0.CO%3B2-Z