# Shrinking Achievement Differences With Anchored Math Problems:

## *Challenges and Possibilities*

**Brian A. Bottge, Enrique Rueda, Ronald C. Serlin, Ya-Hui Hung,** and **Jung Min Kwon**
*University of Wisconsin–Madison*

Multiple measures administered in repeated waves within a nonequivalent dependent variables quasi-experimental design were used to test the effects of a reform-oriented instructional method called Enhanced Anchored Instruction (EAI) on the math achievement of 128 middle school students, including students with learning disabilities (LD). EAI problems are presented in multimedia and hands-on formats, a potential benefit for students with low skills in both reading and math. Overall, students of all ability levels benefited from EAI with effect sizes ($\eta^2$) ranging from .53 to .59. Results revealed that although students with LD scored lower on pretests, their learning trajectories matched those of students without LD. A maintenance test administered several weeks after instruction showed that students with LD retained what they had learned. Implications for instruction and suggestions for future research are provided.

Government and professional groups have urged educators to help *all* students acquire mathematical preparedness for post-secondary education and employment (e.g., *Standards for Technological Literacy: Content for the Study of Technology*, International Technology Education Association, 2000; Goals 2000: Educate America Act, U.S. Department of Education, 1994; *What Work Requires of Schools: A SCANS Report for America 2000*, U.S. Department of Labor, 1991). The most recent of these initiatives, the No Child Left Behind Act of 2001 (NCLB), outlines a national initiative for improving elementary and secondary education tied to high-stakes assessments. In response to these pressures, the National Council of Teachers of Mathematics (NCTM, 2000) has called for curricular reform that emphasizes more problem-based learning. According to the NCTM, these problems should develop the skills and concepts of middle school students in (a) working flexibly with whole numbers, fractions, and decimals; (b) constructing and interpreting scale drawings; (c) converting units of measure; and (d) interpreting tables and graphs.

Recent test scores show that these reforms may be paying off. Results from the National Assessment of Educational Progress (Perie, Grigg, & Dion, 2005) indicated that eighth graders scored higher in 2005 than in any previous year since the test was administered. However, this good news was accompanied by less positive findings showing more than one quarter of students without disabilities (28%) and more than two thirds of students with disabilities (69%) still scoring below *Basic* performance levels. *Basic* means students "should complete problems correctly with the help of structural prompts such as diagrams, charts, and graphs" and include "the appropriate use of strategies and technological tools to understand fundamental algebraic and informal geometric concepts in problem solving" (p. 20). Thus, the new standards call for a range of skills beyond procedural competency.

Higher expectations coupled with the sluggish math performance of students with disabilities have led some special educators (e.g., Jones, Wilson, & Bhojwani, 1997; Woodward, 2004; Woodward & Baxter, 1997; Woodward & Montague, 2002) to question whether traditional instructional methods for students with learning disabilities (LD) are appropriate and adequate. Central to this issue is how and to what extent the teaching practices used in special education—which have leaned toward behaviorist principles—can and should be modified to align with current reform-oriented practices in general education promoting more constructivist methods.

Shifting to a more constructivist approach will not be easy for several reasons. First, special educators (e.g., Vaughn, Klingner, & Hughes, 2000) have cautioned against embracing the reform-oriented methods for less capable students without an adequate research foundation because (a) students who have low math skills may not be able to solve more difficult types of math problems advocated by NCTM and (b) having such students spend time on problems they are not able to solve may limit the time available for basic skills. Various meta-analyses of special education research have shown the benefits of drill and repetition for many students with LD (Swanson,

*Address: Brian A. Bottge, Dept. of Rehabilitation Psychology and Special Education, University of Wisconsin–Madison, 432 N Murray St., Madison, WI 53706; e-mail: bbottge@education.wisc.edu*

2001; Swanson & Hoskyn, 1998). These methods carefully sequence content so task difficulty does not overload students' working memory (e.g., Carnine, 1998; Swanson & Deshler, 2003). Abandoning proven approaches at this time may be unwise, especially when general education is reevaluating their instruction of procedural skills (Star, 2005).

A second important question concerns how to individualize instruction for students with LD who are certain to face more complex content in general education math classrooms. In the past, most students with LD received their math instruction in small-group settings where teachers tailored specialized instruction to each student's learning needs. As more students with LD are included in general education classrooms, the question of whether and how they will get the support in regular education classes is still largely unanswered (Baker & Zigmond, 1990). Research suggests that pullout settings do not produce satisfactory results in school achievement over the long term (Fuchs & Fuchs, 1995; Rea, McLaughlin, Walther-Thomas, 2002), but few studies have identified practices that promote effective math instruction for students with LD in inclusive, reform-oriented settings.

The importance of these issues (i.e., curriculum, individualization) increases in complexity as students with LD transition to secondary school settings. Deshler et al. (2001) identified the struggles students encounter with higher and more focused academic expectations. Their graphic contrasting the upward learning trajectory of typical achievers and the learning plateau of students with LD is particularly informative. Although a variety of approaches has improved the academic performance of adolescents with LD in a number of important areas, such as algebra (Hutchinson, 1993; Maccini & Hughes, 2000; Witzel, Mercer, & Miller, 2003), equivalent fractions (Butler, Miller, Crehan, Babbitt, & Pierce, 2003), and problem solving (Jitendra, DiPipi, & Perron-Jones, 2002; Jitendra, Hoff, & Beck, 1999; Montague, 1997, 1998; Montague & Bos, 1986), more research is needed to show how these approaches can be used alongside of or embedded in general education curricula.

One instructional method that may have the potential for helping students with diverse abilities is called *Enhanced Anchored Instruction* (EAI; Bottge, 2001). Based on the concept of *anchored instruction* (AI; Cognition and Technology Group at Vanderbilt, 1990, 1997; Hickey, Moore, & Pellegrino, 2001), students first solve a problem in a multimedia format and then apply what they learn in related hands-on problems (e.g., building skateboard ramps, designing and manufacturing hovercrafts). A mix of experimental and quasi-experimental studies in inclusive and pull-out settings has yielded effect sizes ($r^2$) from .31 to .66 on problem-solving tests and from .14 to .38 on transfer tasks (Bottge, 1999; Bottge, Heinrichs, Chan, & Serlin, 2001; Bottge, Heinrichs, Mehta, & Hung, 2002). The focal point of these studies was on developing students' understanding of complex, multistep problems that required knowledge of fractions, linear measurement, and representational skills. One important advantage of EAI is its ability to directly immerse students in problem contexts, thus helping to eliminate the comprehension difficulties students with low skills in both math and reading often experience with complex text-based problems (Fuchs & Fuchs, 2002; Lesh & Kelly, 2000).

Although the EAI research has shown positive benefits for many students, the approach has its disadvantages. As Woodward (2004) pointed out, the difficulty level of anchored problems is high, which may at times overload the working memory of students with LD. In earlier studies with a limited number of scaffolds in place to reduce the cognitive load, some students with LD became frustrated by the complexity of EAI problems. A second issue relates to the lack of specificity in descriptions of the concepts and skills in the EAI problems. Prior reports of EAI have included overviews of learning objectives, but they did not provide specifics about what the assessments measured. Finally, the EAI research has highlighted the need to more adequately train teachers who use EAI, especially on how to organize and manage heterogeneous groups. In EAI settings, students with LD work with their classmates in small problem-solving groups, and teachers must be vigilant to ensure that students work cooperatively. In previous studies, group processes have at times broken down, and students with LD simply copied the work of more able students.

The goal of this study was to measure the effects of two EAI problems on the math performance of middle school students across a range of ability levels after improvements had been made in curricular design, instructional methods, and teacher training. Specifically, we assessed whether and to what extent the achievement and learning trajectories of students with and without LD widened, narrowed, or stayed the same as a result of working on two problems:

1. an EAI problem designed to teach students to calculate rates, interpret tables and construct graphs, and construct line of best fit for the purpose of making predictions based on relationships between variables
2. an EAI problem designed to teach students to interpret schematic plans, measure lengths in feet and inches, and compute combinations with whole numbers and fractions for the purpose of estimating the total cost of a building project.

# Method

## *Participants*

A total of 128 seventh-grade students in six intact middle school math classrooms participated in the 7-month study. The middle school is located in a small town in the upper Midwest. Two math teachers (MT1, MT2) each taught three 90-min blocks per day. Table 1 shows descriptive information for students in the six classes. MT1 taught an inclusive class, a pre-algebra class, and a class of students considered neither particularly high nor low achievers (i.e., typical). The inclusive class consisted of 13 students with disabilities, most with LD (hereafter referred to as LD), and a group of 13 students without disabilities who were considered to be typical achievers. The students with LD were placed in one class to allow the special education teacher or her assistant to provide extra help. The school had identified students for the pre-algebra class based on their math performance in sixth grade. Students had to have maintained an A average, completed 95% of their homework, attained at least 85% on a teacher-developed math test, and scored at or above the 90th percentile on the state math test. The students in MT1's other class and in all three of MT2's classes were considered by the school to be typical achievers because their skills were not high enough to be recommended for the pre-algebra class and not low enough to be referred for remedial help.

Table 2 shows demographic information of the students with LD in the inclusive class. Individualized Education Programs (IEPs) revealed that these students were receiving special education services an average of 625 min a week (range = 300–920). Twelve of the students were receiving special education services for LD, and the majority of them were receiving services in more than one content area. Two students with disabilities were considered at risk in math but had not yet been referred for LD services. According to the Wisconsin Department of Public Instruction (2006), students with LD have a significant discrepancy equal to or greater than 1.75 standard errors of the estimate below expected achievement based on individually administered, standardized achievement and ability tests. All of these students were male and Caucasian except one, who was an African American student.

Mean national percentile ranks (NP) on the *TerraNova Comprehensive Test of Basic Skills* (CTB/McGraw-Hill, 1997), which all students took the previous spring as part of statewide testing, showed that the students with disabilities in the inclusive class were achieving below average in math (NP = 16) *and* reading (NP = 19). The *Iowa Tests of Basic Skills* (ITBS; University of Iowa, 2001), which was administered immediately prior to the study, confirmed their low ability in computation (NP = 15) and in problem solving (NP = 14).

Based on the ITBS results, MT1's three classes differed in computation, $F(2, 68) = 67.03$, $p = .00$, $\eta^2 = .66$, and in problem solving, $F(2, 68) = 38.57$, $p = .00$, $\eta^2 = .53$. Pre-algebra students outscored students in the typical classes in computation, $t = 5.62$, $p = .00$, and in problem solving, $t = 3.87$, $p = .00$; and students in the typical class scored higher than did students in the inclusive class both in computation, $t = 5.45$, $p = .00$, and in problem solving, $t = 3.90$, $p = .00$. MT2's typical classes did not differ in computation, $F(2, 51) = 1.36$, $p = .26$, $\eta^2 = .05$, or in problem solving, $F(2, 51) = .85$, $p = .43$, $\eta^2 = .03$, nor were there differences between the four typical classes (one of MT1 and three of MT2) in computation, $F(3, 71) = 2.15$, $p = .10$, $\eta^2 = .08$, or in problem solving, $F(3, 73) = .76$, $p = .50$, $\eta^2 = .03$.

The two math teachers planned and were primarily responsible for instruction in the six classes. The special education teacher or the special education assistant was on hand

**TABLE 1.** Description of Students in the Inclusive, Pre-Algebra, and Typical Math Classes

| | Math Teacher 1 (MT1) | | | Math Teacher 2 (MT2) | | |
|---|---|---|---|---|---|---|
| | Inclusive | Pre-Algebra | Typical | Typical | Typical | Typical |
| Total | 26 | 27 | 21 | 16 | 17 | 21 |
| Gender | | | | | | |
| Boys | 22 | 8 | 8 | 6 | 5 | 11 |
| Girls | 4 | 19 | 13 | 10 | 12 | 10 |
| Special education IEP | 13 | 0 | 0 | 0 | 0 | 0 |
| **ITBS pretest scores** | | | | | | |
| Computation | | | | | | |
| *M* | 202 | 259 | 231 | 216 | 228 | 220 |
| *SD* | 12.55 | 17.37 | 17.63 | 19.22 | 26.22 | 12.19 |
| (NP) | (17) | (80) | (51) | (32) | (47) | (37) |
| Problem solving & data interpretation | | | | | | |
| *M* | 208 | 280 | 249 | 252 | 244 | 238 |
| *SD* | 37.03 | 18.41 | 27.65 | 28.36 | 31.68 | 34.60 |
| (NP) | (28) | (86) | (65) | (67) | (61) | (56) |

*Note*. IEP = Individualized Education Program; ITBS = *Iowa Tests of Basic Skills,* Level 12, *Form A* (University of Iowa, 2001); NP = national percentile.

**TABLE 2.** Description of Students With Disabilities in the Inclusive Class

| | | | | ITBS | | | | TerraNova CTBS | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Computation | | Problem solving & data interpretation | | Math | | Reading | |
| Student | Date of birth | Disability label(s) (Service areas) | Minutes per week in special education | SS | NP | SS | NP | SS | NP | SS | NP |
| Alex | 5/9/90 | LD/ADD (M,R,L) | 600 | 200 | 15 | 131 | 1 | 588 | 5 | 593 | 6 |
| Mark | 3/3/90 | LD/EBD (M,R,L) | 500 | 191 | 10 | 173 | 5 | 619 | 15 | 613 | 12 |
| John | 4/8/90 | SL/LD (L) | 450 | 200 | 15 | 213 | 33 | 641 | 30 | 635 | 26 |
| Ted | 3/31/90 | LD (M,L) | 750 | 182 | 5 | 204 | 24 | 613 | 13 | 554 | 2 |
| Rob | 6/14/90 | LD/ADHD (M,R,L) | 540 | 209 | 24 | 200 | 20 | 623 | 17 | 629 | 22 |
| Jim | 3/27/90 | LD/SL (L) | 665 | 191 | 10 | 185 | 10 | n.a. | n.a. | n.a. | n.a. |
| Les | 3/11/90 | LD (M,R,L) | 880 | 204 | 19 | 185 | 10 | 551 | 1 | 578 | 4 |
| Aaron | 9/4/89 | LD (M,R,L) | 800 | 191 | 10 | 173 | 5 | 551 | 1 | 529 | 1 |
| Brian | 11/29/89 | LD/EBD (M) | 300 | 196 | 12 | 200 | 20 | 592 | 6 | 543 | 1 |
| Rick | 11/25/89 | LD/ADD (M,R) | 480 | 217 | 33 | 161 | 1 | 570 | 2 | 602 | 8 |
| Willy | 10/17/89 | LD (M,R,L) | 920 | 221 | 38 | 185 | 10 | 604 | 9 | 597 | 7 |
| Ed | 5/26/89 | LD (M,R) | n.a. | 186 | 7 | 248 | 64 | 595 | 7 | 550 | 1 |
| Phil | 12/27/89 | ADD | n.a. | 191 | 10 | 233 | 51 | n.a. | n.a. | n.a. | n.a. |

*Note*. LD = learning disability; ADD = attention-deficit disorder; EBD = emotional and behavioral disorder; SL = speech and language disability; ADHD = attention-deficit/hyperactivity disorder; M = math; R = reading; L = written language; ITBS = *Iowa Tests of Basic Skills,* Level 12, *Form A* (University of Iowa, 2001); CTBS = *Comprehensive Test of Basic Skills* (CTB/McGraw-Hill, 1997); SS = scale score; NP = national percentile; n.a. = not applicable.

to help students with LD in the inclusive class, but they provided no direct, whole-group instruction. MT1 was chairperson of the math department, had a master's degree in middle-level education, was beginning his 13th year of teaching, and had incorporated EAI into his math classes the past three years. MT2 had a bachelor's degree, was in her second year of teaching, and had no previous experience teaching with EAI. The special education teacher was in her 15th year of teaching, had a master's degree in general education grades 1 through 8, and was licensed to teach special education in grades kindergarten through 12. Her assistant did not have a college degree, but she had several years of experience working with students with disabilities.

## *Instructional Methods and Materials*

Instruction involved two EAI problems that were integrated into the math curriculum during one academic school year. The teachers taught *Kim's Komet* for 13 days in October and *Fraction of the Cost* for 11 days in March. Between these times, the math teachers followed their prescribed curriculum, *Connected Mathematics* (Dale Seymour Publications, 2004).

MT1 and MT2 assigned two to four students to problem-solving groups based on their ability to work together. In the inclusive class, one student with LD worked with two or three students without disabilities. At the beginning of each class period, the teachers led students in a 10-min warm-up activity to review concepts they had worked on the previous day and to introduce new material. For the remainder of class, each group worked on a laptop computer, which contained the learning tools (i.e., scaffolds) students could use to help solve each of the subproblems. Teachers circulated from group to group, answering questions and posing new ones. After students solved the media-based anchor, they worked on the hands-on applications.

**Kim's Komet Instruction.** *Kim's Komet* is one episode in a series of video-based anchors called *The New Adventures of Jasper Woodbury* (Learning Technology Center at Vanderbilt University, 1997). As stated in the teacher manual, the purpose of *Kim's Komet* is to help students develop their informal understanding of pre-algebraic concepts, such as linear function, line of best fit, variables, rate of change (slope), and reliability and measurement error. Foundation skills needed to solve this problem include computation with whole numbers and decimals.

The video anchor involves two girls who compete in pentathlon events. The first challenge asks students to identify the three fastest qualifiers in three regional races, where the times and distances are known but the distances vary. Students learn, for example, whether a car that travels 15 ft in 0.9 s is faster or slower than a car that travels 20 ft in 1.3 s. The second challenge asks students to construct the "line of best fit" on their graph to predict the speed of cars at the end of a straightaway after they have been released from various heights on the ramp. The video gives students the opportunity to clock the speed of Kim's car from several heights. Eventually students realize that they should begin timing their cars on the straightaway, where the car's speed is relatively constant, rather than on the ramp, where the car is accelerating. The video helps students understand the concept of reliability of measurement, because students measure the speeds several times.

After students solved the problems in the video anchor, they participated in their own pentathlon competition by building and testing cars they made. The technology education teacher at the school built the ramp, straightaway track, and pentathlon stunts shown in the video anchor. At the beginning and end of the straightaway, an infrared detector measured the time it took for the students' cars to travel from one end of the straightaway to the other. Using these times, students made their own graphs showing the speeds of their cars for each release point on the ramp. Their graphs helped them predict where on the ramp to release their cars to achieve the speeds necessary to successfully navigate the stunts at the end of the straightaway.

**Standard Instruction.** After teaching *Kim's Komet* and before introducing *Fraction of the Cost*, the math teachers taught units on concepts related to geometry and proportional reasoning using their prescribed curriculum, *Connected Mathematics* (Dale Seymour Publications, 2004). The students completed several projects, such as figuring out the cost of an advertisement based on the proportion of the full page it covered. The teachers organized their instruction during this phase in ways that were similar to EAI. That is, they reviewed previous work, taught new concepts explicitly, and then provided time for students to solve related problems.

**Fraction of the Cost Instruction.** The authors developed the 8-min video-based anchor called *Fraction of the Cost*, which stars three students from a local middle school. Available in Spanish and English, the video was filmed at a local skateboarding store, a garage, and the backyard of a local home. The video opens in a skateboard store and rink, where the students are shown discussing how they can afford to buy materials for building a skateboard ramp. To solve the problem, students need to (a) calculate the percent of money in a savings account and sales tax on a purchase, (b) read a tape measure, (c) convert feet to inches, (d) decipher building plans, (e) construct a table of materials, (f) compute whole numbers and mixed fractions, (g) estimate and compute combinations, and (h) calculate total cost.

Instruction proceeded in much the same way as with *Kim's Komet*. Teachers showed the 8-min problem-solving anchor *Fraction of the Cost* without interruption the first day and asked students to describe the problems associated with it. Students then worked in small groups on their problem-solving packets. To figure out the most economical use of two-by-four (2" × 4") lumber in *Fraction of the Cost*, students first needed to identify the lengths of boards shown in the plan. Then they converted these dimensions from feet and inches to inches only and calculated several combinations of lengths to maximize their use of the wood available. Once they computed what additional wood was needed, they consulted the store advertisement to decide what to buy. Their final task was to make a materials list to show the total cost of the project.

To help solidify students' understanding, the math teachers often led students in brief practice sessions the first 10 min of class. For example, one day the teachers projected on the projector screen four narrow rectangles of various lengths that represented two-by-fours and asked students to respond to a series of questions about converting feet and inches to inches only and the most economical way of cutting the boards the students in the video found in the garage:

TEACHER: Okay, eyes up here. [Points to a board labeled 2 feet, 5½ inches.] What is this in inches?

CINDY: 2 feet, 5½ inches … 12 inches × 2 = 24 inches … 24 inches + 5½ inches = 29½ inches.

TEACHER: Good. [Points to a board labeled 5 feet, 11¼ inches.]

GREG: 5 feet, 11¼ inches … 5 × 12 inches =

60 inches … 60 inches + 11¼ inches = 71¼ inches.

TEACHER: Can I cut this piece (29¾") out of a board 2 feet, 5½ inches long?

ALL STUDENTS: No!

TEACHER: Can I cut 79 inches from a board that is 8 feet long? Why or why not?

BRIAN: 8 feet × 12 inches is 96 inches, and 96 inches is longer than 79 inches … so the answer is yes!

TEACHER: How much is left over?

ALL STUDENTS: 17 inches.

TEACHER: Are there any of these boards [pointing to schematic plan] that I can cut from what's left, 17 inches?

ALL STUDENTS: 16½ inches.

TEACHER: How much wood do I lose each time I cut?

ALL STUDENTS: ⅛ inch [referring to the kerf of the saw blade].

TEACHER: How many times did I cut?

ALL STUDENTS: Twice.

TEACHER: So how much wood did I lose from cutting the board twice?

ALL STUDENTS: 2/8 inches, or ¼ inch.

TEACHER: Let's take a vote. Should I cut the 8-foot board into 16½ inches and 79 inches?

ALL STUDENTS: Yes!

After students solved the problems posed in *Fraction of the Cost*, they worked on solving a related problem called the *Hovercraft Challenge*, in which students had to plan and construct out of PVC pipe a "rollover cage" for a hovercraft. The teacher divided the class into groups of three students, and each group planned how they could make the cage in the most economical way. Once the teacher approved the plans, students worked on measuring, cutting, and assembling. When the cages were complete, they lifted them onto a 4' × 4' plywood platform. A leaf blower inserted into a hole in the plywood powered the hovercraft. On the last day of the project, students rode on their hovercrafts in relay races up and down the halls of the school.

## *Instrumentation*

Two tests, the *Kim's Komet Challenge* (*KKC*) and the *Fraction of the Cost Challenge* (*FCC*), assessed minimally overlapping *sets of discrete math concepts* embedded in the two EAI problems. Constructed-response items measured knowledge of mathematics aligned to NCTM (2000) standards recommended for students in Grades 6–8 (i.e., Numbers and Operations, Measurement, Problem Solving, Communication, and Representation). Each test went through cycles of refinement based on student performances in previous research (Bottge, Heinrichs, Mehta, & Hung, 2002; Bottge et al., 2004; Bottge, Heinrichs, Chan, & Serlin, 2001; Hung, 2005) and on suggestions from math and assessment specialists (i.e., math teachers, math researchers, test consultants). Sets of problems were weighted according to their complexity and the contribution they were expected to make in solving the overarching problem. Within each set, items were awarded full or partial credit, which made it possible to analyze student work at both the item and concept levels (Lester & Kroll, 1990; Shafer & Romberg, 1999). Concurrent validity correlation coefficients based on pretest scores of the *Iowa Tests of Basic Skills* (ITBS) Problem Solving and Data Interpretation Subtest (University of Iowa, 2001) were .59 for the *FCC* and .52 for the *KKC*. Both correlations were significant and appear acceptable given that the range of mathematics concepts sampled by *FCC* and *KKC* was more restricted than that sampled on the ITBS. Internal consistency (Cronbach's coefficient alpha) of *KKC* and *FCC* were .92 and .80, respectively. The research staff calculated interrater reliabilities on 20% of the test protocols randomly selected from each of the scheduled test administrations and were computed by dividing the number of agreements by the total number of agreements and disagreements and multiplying by 100 (Sulzer-Azaroff & Mayer, 1977). Interrater reliability was 95% (range = 93–98%) for the *Kim's Komet* test and 94% (range = 90–97%) for the *Fraction of the Cost* test.

**Kim's Komet Challenge.** The 36-point *Kim's Komet Challenge* assessed students' ability to estimate and compute rates when times and distances are known. Students were assessed on their ability to interpret data in tables, to construct graphs, and to make predictions based on relationships between variables. These concepts and skills match closely those identified in the NCTM (2000) Algebra Standards for Grades 6–8.

Figure 1 shows three sets of test items (Items 5–7), which together account for 28 of the 36 total points. Student work is included to show correct responses. In Item 5, students use the information in the table to calculate the speeds of six cars, in feet per second, when distance and time are provided. The figure above the table is intended to depict the ramp and straightaway as they are shown in the instructional anchors (i.e., video-based and hands-on). After calculating each of the speeds, students indicate who in each town has the slowest and fastest cars and which car is fastest overall. Students can earn a total of 16 points if they calculate all speeds correctly or partial credit if they answer only some of the items correctly.

In Item 6, students calculate how fast cars are traveling at the end of the straightaway after they have been released from several heights on a ramp. This problem is difficult because students must first understand why they should *not* begin timing the cars at the release point on the ramp (i.e., heights of 44", 60", 108"). During instruction, students learn that cars accelerate on the down ramp and therefore accurate speeds cannot be computed. Instead, they must use the time

## Item 5

Two groups of high school students were racing their cars down a ramp. The ramp is show below. The car tracks were different lengths.

START

a. Calculate the speeds of the Lodi and Sun Prairie cars. Write the speeds in the tables below.

**LODI STUDENTS**

| STUDENT | LENGTH OF TRACK | TIME | SPEED (Your answer) |
|---|---|---|---|
| Ann | 10 feet | 1.6 seconds | 6.3 ft/sec |
| Mary | 10 feet | 2.0 seconds | 5 ft/sec |
| Rich | 10 feet | 2.8 seconds | 3.6 ft/sec |

**SUN PRAIRIE STUDENTS**

| STUDENT | LENGTH OF TRACK | TIME | SPEED |
|---|---|---|---|
| Tom | 15 feet | 2.5 seconds | 6 ft/sec |
| Jan | 15 feet | 2.0 seconds | 7.5 ft/sec |
| Dave | 15 feet | 3.2 seconds | 4.7 ft/sec |

b. Who has the fastest car in Lodi? Your answer: Ann's

c. Who has the fastest car in Sun Prairie? Your answer: Jan's

d. In Lodi and Sun Prairie, who has the fastest car? Your answer: Jan's

e. In Lodi and Sun Prairie, who has the slowest car? Your answer: Rich's

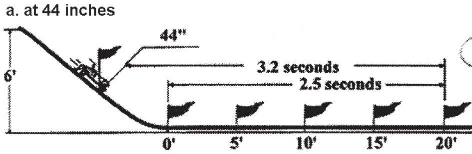
**Scoring Protocol**

| a. | Speed | Points |
|---|---|---|
| Ann | 6.25, 6.3, 6.2 | 2 ✓ |
| Mary | 5 | 2 ✓ |
| Rich | 3.6, 3.5, 3.57 | 2 ✓ |
| Tom | 6 | 2 ✓ |
| Jan | 7.5 | 2 ✓ |
| Dave | 4.7, 4.6, 4.68 | 2 ✓ |
| b. Ann | | 1 ✓ |
| c. Jan | | 1 ✓ |
| d. Jan | | 1 ✓ |
| e. Rich | | 1 ✓ |
| Total points | | 16/16 |

## Item 6

Calculate the speed of a car (feet per second) dropped from each height on the ramp.

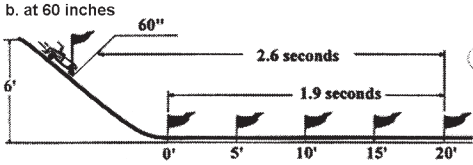
a. at 44 inches




Work Area: 20÷2.5=8

Your answer: 8 ft/sec

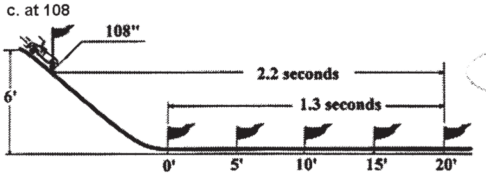
b. at 60 inches




Work Area: 20÷1.9=10.5 ft/sec.

Your answer: 10.5

c. at 108




Work Area: 20÷1.3= 15.4

Your answer: 15.4

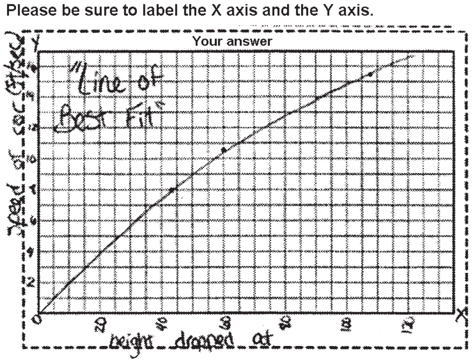
**Scoring Protocol**

| Work Area | | Partial | Full |
|---|---|---|---|
| @44" 20'÷ 2.5 sec. = | 8 ft/sec | 1 | 2 ✓ |
| @60" 20' ÷ 1.9 sec. = | 10.5, 10.52, 10.53 | 1 | 2 ✓ |
| @ 108 20' ÷ 1.3 sec = | 15.4, 15.3, 15.38 | 1 | 2 ✓ |
| Total points | | | 6 /6 |

## Item 7

Make a graph to show the speed of the car for each height on the ramp. Please be sure to label the X axis and the Y axis.

Your answer




**Scoring Protocol**

| | Partial | Full |
|---|---|---|
| X-Axis (correct label and correct scale) Speed or Rate inclusive of 16 OR Distance or Height inclusive of 108 | 1 | 2 ✓ |
| Y-Axis (correct label and correct scale) Speed or Rate inclusive of 16 OR Distance or Height inclusive of 108 | 1 | 2 ✓ |
| Line graph plotting 3 points: (44, 8) (60, 10.5) (108, 15.4) | 1 | 2 ✓ |
| Total points | | 6 /6 |

**FIGURE 1.** Representative items and student work on *Kim's Komet Challenge (KKC)*.

it takes for the cars to travel from the beginning to the end of the straightaway, where speeds are more constant. Students calculate the speeds of their cars by dividing the length of the straightaway by the time taken to travel the straightaway. Students can earn partial credit for showing correct procedures.

Item 7 asks students to graph the data from Item 6, label both axes, and draw a line of best fit. During instruction, students learn that this line shows the relationship between the height of the release point on the ramp and the speeds the cars will be traveling as they approach the end of the straightaway. The graph helps students pinpoint how fast a car will be traveling at the end of the straightaway for every release point (i.e., height) on the ramp. The scoring protocol shows that students can receive full credit for correctly graphing the data or partial credit if some of the information is correct. Students can earn partial or full credit in six major categories with scale weights based on more than 20 procedures.

**Fraction of the Cost Challenge.** The 37-point *Fraction of the Cost Challenge* assessed students' ability to interpret a three-dimensional schematic plan, measure lengths of building materials in feet and inches, estimate and compute combinations using whole numbers and fractions, interpret and record data in tables, and calculate total cost including sales tax. The test emphasized real-world knowledge such as reading a bank account statement, comparing treated and untreated lumber, selecting appropriate lengths and quantities of screws for specific applications, and developing a materials list with a cost estimate. Like the *KKC*, the procedures in the *FCC* were weighted according to difficulty and their contribution to the overall problem.

Figure 2 shows two sets of items, scoring protocols, and the information students use to answer each problem. In Item 3, students use the rulers to measure the lengths of boards that appear in the anchor problem. Students are to record the lengths to the nearest ⅛-inch in either feet and inches or in inches alone. Item 4 is more difficult because it requires students to show the most economical way to cut the wood shown in Item 3 for building the skateboard ramp. To solve this problem, students must first be able to "read" the schematic plan, identify the number of boards of each length, and then figure out the most economical way to use the available wood. Then they use these combinations to indicate where to saw the boards. Like the *KKC*, the *FCC* requires students to integrate and manage several problem sets.

## Research Design

Multiple measures administered in repeated waves within a nonequivalent dependent variables design assessed the effects of EAI on the achievement of students by ability and disability status. A unique feature of this design is that all students have access to the instructional interventions. At the most basic level, it involves a single group of students tested on two scales that are conceptually similar, but only one of which is expected to change because of the treatment. Adding multiple, repeated measures and predicting achievement patterns (i.e., *pattern matching*), as was done in this study, minimizes plausible threats to internal validity such as maturation, history, and testing (Cook & Campbell, 1979; Shadish, Cook, & Campbell, 2002). Figure 3 shows the schedule of instruction and testing represented in standard design notation.

Prior to instruction, students in all six classes were administered both problem-solving tests, *KKC* ($O_{1A}$) and *FCC* ($O_{1B}$). After teachers taught *Kim's Komet* ($X_A$), students took the *KKC* ($O_{2A}$), which served as a posttest for the concepts they were just taught, and the *FCC* ($O_{2B}$), which assessed math concepts not emphasized during *Kim's Komet*. For the next 14 weeks, the teachers taught with their usual math curriculum. In the third wave of testing, all students took the *KKC* ($O_{3A}$), which assessed students' retention of concepts they had last worked on several weeks before, and the *FCC* ($O_{3B}$), which served as the pretest for *Fraction of the Cost* instruction. In the last test wave, students took the *FCC* ($O_{4B}$), which measured what students learned from *Fraction of the Cost*.

A unique feature of this design is that the treatment is planned to affect only one of the measures at a time, and thus each outcome measure serves as a control for the other. Prior to conducting the study, we predicted that students would show overall improvement from test waves $O_{1A}$ to $O_{2A}$ while maintaining stable achievement from $O_{1B}$ to $O_{2B}$ and that, furthermore, the achievement pattern at $O_{1A}$ to $O_{2A}$ would be repeated at $O_{3B}$ to $O_{4B}$. We also predicted that this pattern would align closely with ability and disability status. That is, students in the pre-algebra class would score higher than students in the typical math classes; students in the typical math classes would score higher than students in the inclusive class; and, students with LD would score lower than all of the other groups.

## Fidelity of Implementation

Direct observation and video were used to monitor instruction in the math teachers' classrooms. Detailed daily lesson plans developed especially for use with the EAI problems and frequent checks of lesson implementation helped ensure that the instruction proceeded as planned. Classroom observers recorded whether teachers stated the objectives for the day, led students in warm-up review exercises, reviewed concepts from previous classes, introduced new concepts related to the EAI problems, allowed time for students in their small groups to discuss possible solutions, and attended to the needs of individual students. A primary and second observer took field notes 100% (24 days) and 67% (16 days) of the time, respectively, in the inclusive classroom. A primary and secondary observer also took field notes 71% (17 days) and 13% (3 days) of the time, respectively, in the MT2 classes.

In addition to the observational data, the teaching practices of MT1 and the special education staff in the inclusive classroom were chronicled with two cameras. One stationary camera was positioned in the back of the room facing the teacher. The wide-angle feature of the camera captured almost

## Item 3

What are the lengths of the wood in the garage?

Your answer

| Quantity | Description | Size in Feet | Size in Inches |
|---|---|---|---|
| 1 | Long Wood | 5' 10 5/8" | 70 5/8" |
| 1 | | 5' 1 1/4" | 61 1/4" |
| 1 | | 4' 11 3/4" | 59 3/4" |
| 1 | | 2' 4 1/2" | 28 1/2" |
| 1 | | 7' 8 1/2" | 92 1/2" |
| 1 | | | |
| 1 | | | |

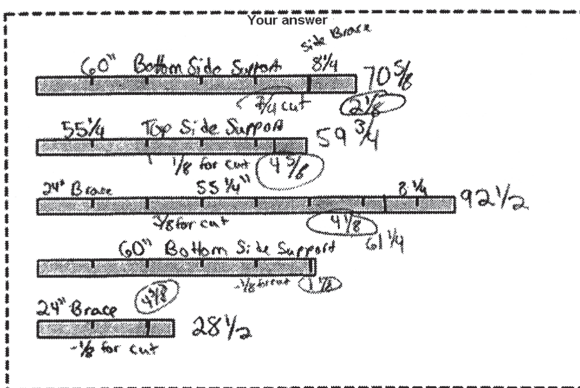
**Scoring Protocol**

| Feet | Inches | *Points | |
|---|---|---|---|
| 5' 10 $\frac{5}{8}$" | 70 $\frac{5}{8}$" | 0,1,2 correct= 0<br>3, 4 correct= 2<br>All correct = 4 | *Select one column (feet OR inches) to award points. |
| 4' 11 $\frac{3}{4}$" | 59 $\frac{3}{4}$" | | |
| 7' 8 $\frac{1}{2}$" | 92 $\frac{1}{2}$" | | |
| 5' 1 $\frac{1}{4}$" | 61 $\frac{1}{4}$" | | |
| 2' 4 $\frac{1}{2}$" | 28 $\frac{1}{2}$" | | |

Total Wood in Garage section = 4 /4

## Item 4

Label the wood that you found in the garage (drawn below) and show how you would use the wood in your skateboard ramp. Make sure that you show how much wood is left in each piece.

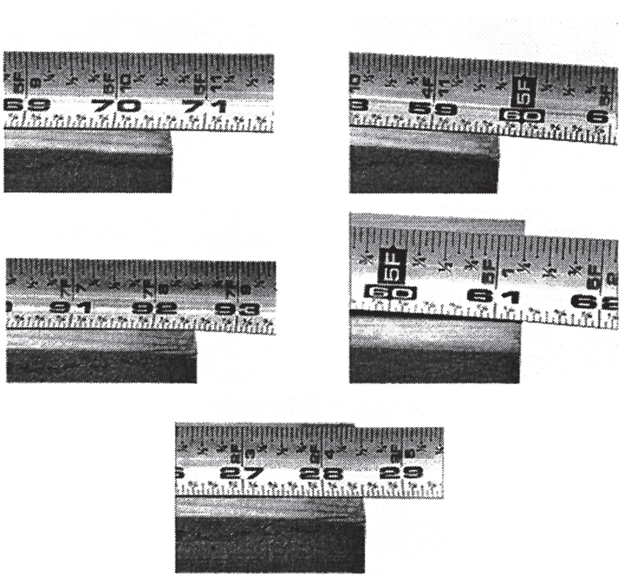



**Scoring Protocol**

| Feet | Inches | Label | Use |
|---|---|---|---|
| 5' 10 $\frac{5}{8}$" | 70 $\frac{5}{8}$" | Correctly labels 3 to 5 boards in feet or inches = 1 pt. | 1 pt. For each board cut in an economical way.<br>1 2 3 4 5 |
| 4' 11 $\frac{3}{4}$" | 59 $\frac{3}{4}$" | | |
| 7' 8 $\frac{1}{2}$" | 92 $\frac{1}{2}$" | | |
| 5' 1 $\frac{1}{4}$" | 61 $\frac{1}{4}$" | | |
| 2' 4 $\frac{1}{2}$" | 28 $\frac{1}{2}$" | | |

Total garage wood section = 6 /6

### Information from the Video



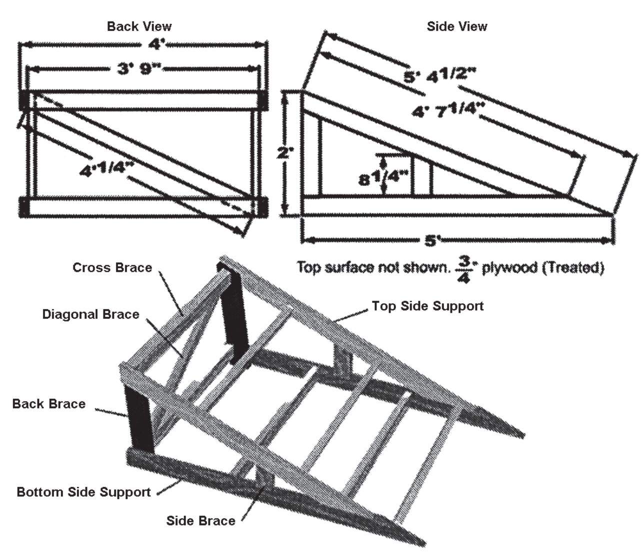



**FIGURE 2.** Representative items and student work on *Fraction of the Cost Challenge (FCC).*

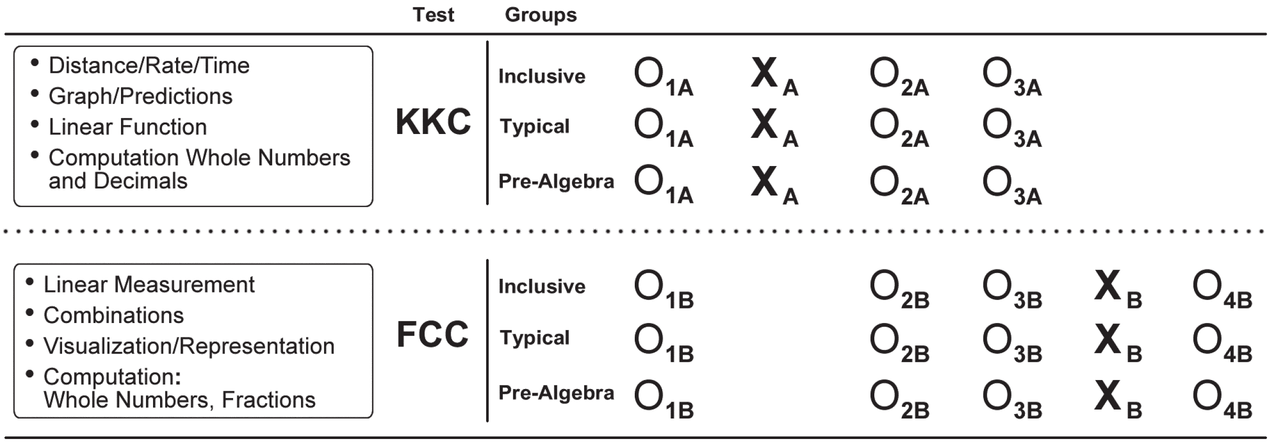



**FIGURE 3.** Nonequivalent dependent variables design with multiple measures in repeated waves.

the full range of classroom activity, which included the teachers' instructional behaviors and students in their small groups. A videographer with a mobile camera captured close-up views of students working and conversing. The two cameras produced more than 30 1-hr tapes, which were digitized and compressed in MPEG-1 encoding format. They were then transferred to video analysis software that made it possible to link the classroom lessons to time points of the observers' field notes. MT1 and MT2 were observed to have engaged in 95% and 97% out of the total number of activities, respectively. Interrater reliability was 99%.

# Results

## *Math Performance*

**Overall Comparisons.** We compared the effects of two EAI problems on the math skills of high-achieving students in a pre-algebra class, average-achieving students in typical math classes, and students with and without LD in an inclusive math class. Multiple comparison procedures tested differences on each measure by test wave and by disability status. An alpha level of .05 was chosen for the test of each of the predicted wave and achievement patterns, and Fisher's Least Significant Difference (LSD) method or Holm's sequentially rejective multiple comparison procedure was used (Holm, 1979; Seaman, Levin, & Serlin, 1991) to control Type I error rate among the between-group comparisons.

Table 3 shows the means and standard deviations of the students in the six math classes. We conducted two separate two-way split plot analysis of variances (ANOVAs) with repeated measures on students' scores, one for each measure. The within-student factor was instruction (*Kim's Komet* Test Wave 1, 2, 3 or *Fraction of the Cost* Test Wave 1, 2, 3, 4), and the between-students factor was class (inclusive, typical, pre-algebra). On the *KKC*, there was a main effect for instruction, $F$(2, 109= 157.88, $p < .001$, $\eta^2 = .59$; for class, $F(2, 109) = 41.67$, $p < .001$, $\eta^2 = .43$; and for instruction-by-class interaction, $F(4, 109) = 5.93$, $p < .001$, $\eta^2 = .10$. On the *FFC*, there

**TABLE 3.** Mean and Standard Deviations of Inclusive, Typical, and Pre-Algebra Classes by Test Wave

| | Measure | | | |
|---|---|---|---|---|
| | *Kim's Komet Challenge*[a] | | *Fraction of the Cost Challenge*[b] | |
| Achievement level | *M* | *SD* | *M* | *SD* |
| Test Wave 1 | | | | |
| Inclusive | 7.73 | 5.78 | 6.35 | 3.76 |
| Typical | 8.37 | 4.44 | 8.95 | 4.94 |
| Pre-Algebra | 16.42 | 8.02 | 13.61 | 6.88 |
| Overall | 10.10 | 6.64 | 9.35 | 5.58 |
| Test Wave 2 | | | | |
| Inclusive | 26.59 | 5.96 | 8.06 | 5.61 |
| Typical | 23.48 | 9.02 | 11.36 | 5.33 |
| Pre-Algebra | 32.88 | 2.57 | 17.28 | 3.83 |
| Overall | 26.25 | 8.33 | 11.87 | 5.83 |
| Test Wave 3 | | | | |
| Inclusive | 22.05 | 7.64 | 8.41 | 6.44 |
| Typical | 12.75 | 7.94 | 10.95 | 6.08 |
| Pre-Algebra | 26.77 | 9.36 | 17.03 | 4.58 |
| Overall | 17.79 | 10.18 | 11.62 | 6.45 |
| Test Wave 4 | | | | |
| Inclusive | — | — | 19.06 | 6.79 |
| Typical | — | — | 20.55 | 6.45 |
| Pre-Algebra | — | — | 26.94 | 6.33 |
| Overall | — | — | 21.45 | 9.35 |

*Note*. — = not tested.
[a]$n$ = Inclusive 22, Typical 65, Pre-Algebra 26. [b]$n$ = Inclusive 17, Typical 64, Pre-Algebra 18.

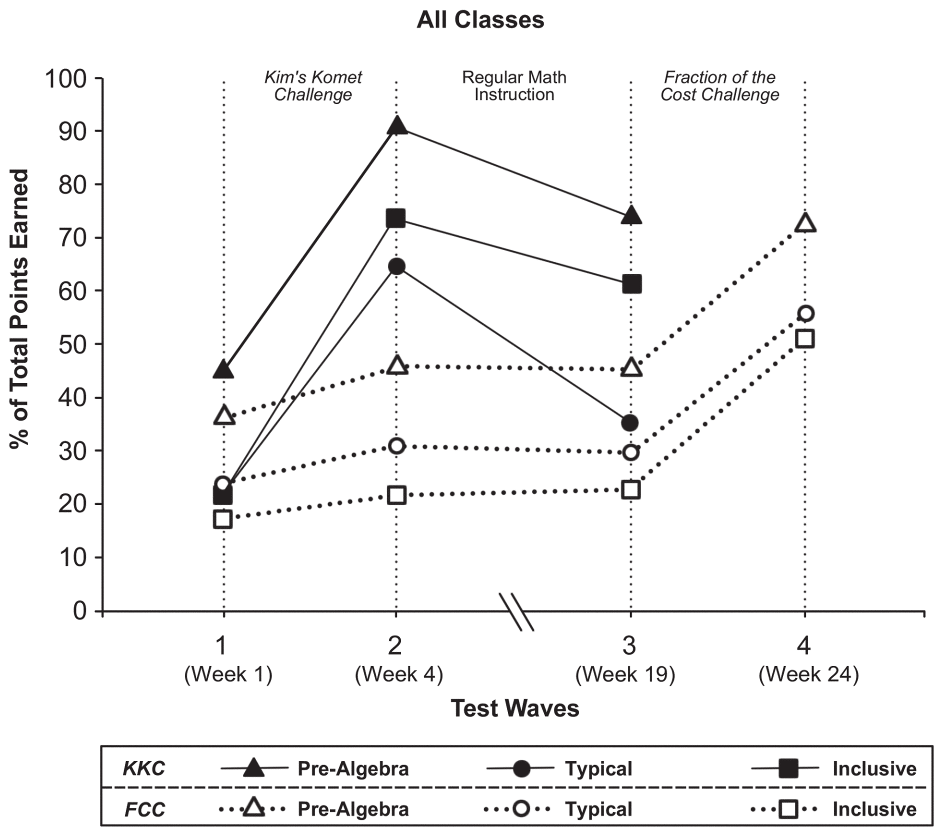



**FIGURE 4.** Performances of students in the pre-algebra, typical, and inclusive classes.

was a main effect for instruction $F(3, 93) = 107.90$, $p < .001$, $\eta^2 = .53$, for class, $F(2, 93) = 18.75$, $p < .001$, $\eta^2 = .28$, but not for instruction-by-class, $F(6, 93) = .38$, $p = .89$, $\eta^2 = .01$. Figure 4 shows the performance of students on both measures at each test wave.

Table 4 shows the main effects comparisons for the performances of students by test wave and by class. Overall, students scored higher on the *KKC* in Test Wave 2 (posttest) and in Test Wave 3 (maintenance) than in Test Wave 1, and students in the pre-algebra and inclusive classes outperformed students in the typical classes. On the *FCC*, students scored higher in Test Wave 4 (posttest) than in Test Wave 3 (pretest), but no differences were found between Test Waves 2 and 3 (double pretest). Students in the pre-algebra class scored higher than students did in the typical classes who, in turn, scored higher than students in the inclusive class did.

Table 5 shows the comparisons of improvement scores by class across assessment waves. These scores show whether the achievement gaps that may have existed at pretest levels stayed the same or increased at posttest. On the *KKC*, there were no differences between classes in the amount of improvement students made between Test Waves 1 and 2. However, students in both the pre-algebra class and the inclusive class showed greater improvement between Test Waves 1 and 3 than did students in the typical classes. On the *FFC*, no differences in improvement were found between classes over test waves.

**Comparisons of Students With and Without LD.** The procedures described above were also used to analyze the performance of students with and without LD in the inclusive class. Table 6 shows the means and standard deviations for each group of students at each test wave. We conducted two separate two-way split plot ANOVAs with repeated measures on students' scores, one for each measure. The within-student factor was instruction (Test Waves 1, 2, 3 for *Kim's Komet* and Test Waves 1, 2, 3, 4 for *Fraction of the Cost*), and the between-students factor was disability status. On the *KKC*, there was a main effect for instruction, $F(2, 20) = 76.36$, $p < .001$, $\eta^2 = .79$, and for instruction-by-disability interaction, $F(2, 20) = 4.01$, $p = .03$, $\eta^2 = .17$, but not for disability, $F(1, 20) = 1.26$, $p = .28$, $\eta^2 = .06$. On the *FFC*, there was a main effect for instruction, $F(3, 14) = 30.68$, $p < .001$, $\eta^2 = .67$, and for disability, $F(1, 14) = 6.52$, $p = .02$, $\eta^2 = .30$, but not for instruction-by-disability interaction, $F(3, 14) = 1.40$, $p = .26$, $\eta^2 = .09$. Figure 5 shows the performance of students with LD and students without LD on the *KKC* and *FCC* for each test wave.

**TABLE 4.** Comparisons on *Kim's Komet Challenge* (*KKC*) and *Fraction of the Cost Challenge* (*FCC*) by Test Wave and Class

| Measure and comparison | Contrast estimate | *SE* | Observed *t* | *p* value |
|---|---|---|---|---|
| *Kim's Komet Challenge* | | | | |
| Test wave comparisons | | | | |
| Wave 2 vs. Wave 1 | 16.15 | 0.76 | 21.35 | p < .001* |
| Wave 3 vs. Wave 1 | 7.69 | 0.89 | 8.64 | *p* < .001* |
| Achievement comparisons | | | | |
| Pre-algebra vs. Typical | 10.49 | 1.15 | 9.12 | *p* < .001* |
| Pre-algebra vs. Inclusive | 6.57 | 1.44 | 4.56 | *p* < .001* |
| Typical vs. Inclusive | −3.92 | 1.23 | −3.19 | *p* = .002* |
| *Fraction of the Cost Challenge* | | | | |
| Test wave comparisons | | | | |
| Wave 2 vs. Wave 1 | 2.52 | 0.53 | 4.76 | *p* < .001* |
| Wave 3 vs. Wave 1 | 2.27 | 0.61 | 3.74 | *p* = .001* |
| Wave 3 vs. Wave 2 | −0.25 | 0.50 | −0.49 | *p* = .624 |
| Wave 4 vs. Wave 3 | 9.83 | 0.68 | 14.55 | *p* < .001* |
| Achievement comparisons | | | | |
| Pre-algebra vs. Typical | 5.76 | 1.12 | 5.14 | *p* < .001* |
| Pre-algebra vs. Inclusive | 8.24 | 1.42 | 5.80 | *p* < .001* |
| Typical vs. Inclusive | 2.48 | 1.15 | 2.16 | *p* = .03* |

*Significant, α = .05 familywise, Fisher's LSD. ** Significant, α = .05 familywise, Holm.

**TABLE 5.** Comparisons of Improvement Scores for Inclusive, Typical, and Pre-Algebra Classes by Test Wave

| Measure and comparison | Contrast estimate | *SE* | Observed *t* | *p* value |
|---|---|---|---|---|
| *Kim's Komet Challenge* | | | | |
| Wave 2 and Wave 1 | | | | |
| Pre-algebra vs. Typical | 1.35 | 1.85 | 0.73 | *p* = .466 |
| Pre-algebra vs. Inclusive | −2.40 | 2.31 | −1.04 | *p* = .301 |
| Typical vs. Inclusive | −3.76 | 1.97 | −1.91 | *p* = .059 |
| Wave 3 and Wave 1 | | | | |
| Pre-algebra vs. Typical | 5.96 | 2.00 | 2.98 | *p* = .004** |
| Pre-algebra vs. Inclusive | −3.97 | 2.50 | −1.59 | *p* = .115 |
| Typical vs. Inclusive | −9.93 | 2.13 | −4.66 | *p* < .001** |
| *Fraction of the Cost Challenge* | | | | |
| Wave 1 and Wave 2 | | | | |
| Pre-algebra vs. Typical | 1.26 | 1.41 | 0.89 | *p* = .373 |
| Pre-algebra vs. Inclusive | 1.96 | 1.78 | 1.10 | *p* = .274 |
| Typical vs. Inclusive | 0.70 | 1.44 | 0.49 | *p* = .628 |
| Wave 1 and Wave 3 | | | | |
| Pre-algebra vs. Typical | 1.42 | 1.62 | 0.01 | *p* = .383 |
| Pre-algebra vs. Inclusive | 1.36 | 2.05 | 0.66 | *p* = .510 |
| Typical vs. Inclusive | −0.06 | 1.65 | −0.04 | *p* = .972 |
| Wave 3 and Wave 4 | | | | |
| Pre-algebra vs. Typical | 0.32 | 1.81 | 0.18 | *p* = .859 |
| Pre-algebra vs. Inclusive | −0.73 | 2.29 | −0.32 | *p* = .751 |
| Typical vs. Inclusive | −1.05 | 1.85 | −0.01 | *p* = .571 |

**Significant, α = .05 familywise, Holm.

**TABLE 6.** Means and Standard Deviations of Students in the Inclusive Class With LD and Without LD by Test Wave

| | Measure | | | |
|---|---|---|---|---|
| | *Kim's Komet Challenge* | | *Fraction of the Cost Challenge* | |
| Disability | *M* | *SD* | *M* | *SD* |
| Test Wave 1 | | | | |
| LD | 5.17 | 2.48 | 5.67 | 4.06 |
| Non-LD | 10.80 | 7.16 | 7.12 | 3.48 |
| Overall mean | 7.73 | 5.78 | 6.35 | 3.76 |
| Test Wave 2 | | | | |
| LD | 24.83 | 5.24 | 5.67 | 5.94 |
| Non-LD | 28.70 | 6.34 | 10.75 | 3.99 |
| Overall mean | 26.59 | 5.96 | 8.06 | 5.61 |
| Test Wave 3 | | | | |
| LD | 23.33 | 6.14 | 4.89 | 4.20 |
| Non-LD | 20.50 | 9.22 | 12.37 | 6.39 |
| Overall mean | 22.05 | 7.64 | 8.41 | 6.44 |
| Test Wave 4 | | | | |
| LD | — | — | 16.89 | 5.63 |
| non-LD | — | — | 21.50 | 7.50 |
| Overall mean | | | 19.06 | 6.79 |

*Note*. LD = learning disabled; non-LD = non–learning disabled; — = not tested.

Table 7 shows comparisons of achievement and improvement scores of students in the inclusive class by disability status across test waves. On the *KKC*, students scored higher in Test Wave 2 (posttest) and Test Wave 3 (maintenance) than in Test Wave 1 (pretest). Improvement from Test Wave 1 to 3 (pretest to maintenance) was larger for students with LD than it was for students without LD. On the *FCC*, students scored higher in Test Wave 4 (posttest) than in Test Wave 3 (pretest), but there were no differences in achievement across Test Waves 1, 2, and 3. Improvement scores between Test Waves 3 and 4 showed no difference between students with LD and students without LD.

**Item-/Concept-Level Analyses.** Tables 8 and 9 indicate the percentage of students who earned either full or partial credit on items within each concept assessed on *KKC* and *FCC*, respectively. On the *KKC*, the difference in the proportion of students earning credit before and after instruction was greatest for calculating rates, completing data tables, understanding concepts of constant speed and acceleration, and constructing a graph (see Figure 1, Items 5–7). Almost all of the students in the pre-algebra class (92%) were able to do Item 5 on the pretest, compared to only about half (55%) of the students with LD. However, post-instruction results showed that almost all of the students with LD (91%) received at least partial credit for their answers, and 68% of them earned a per-

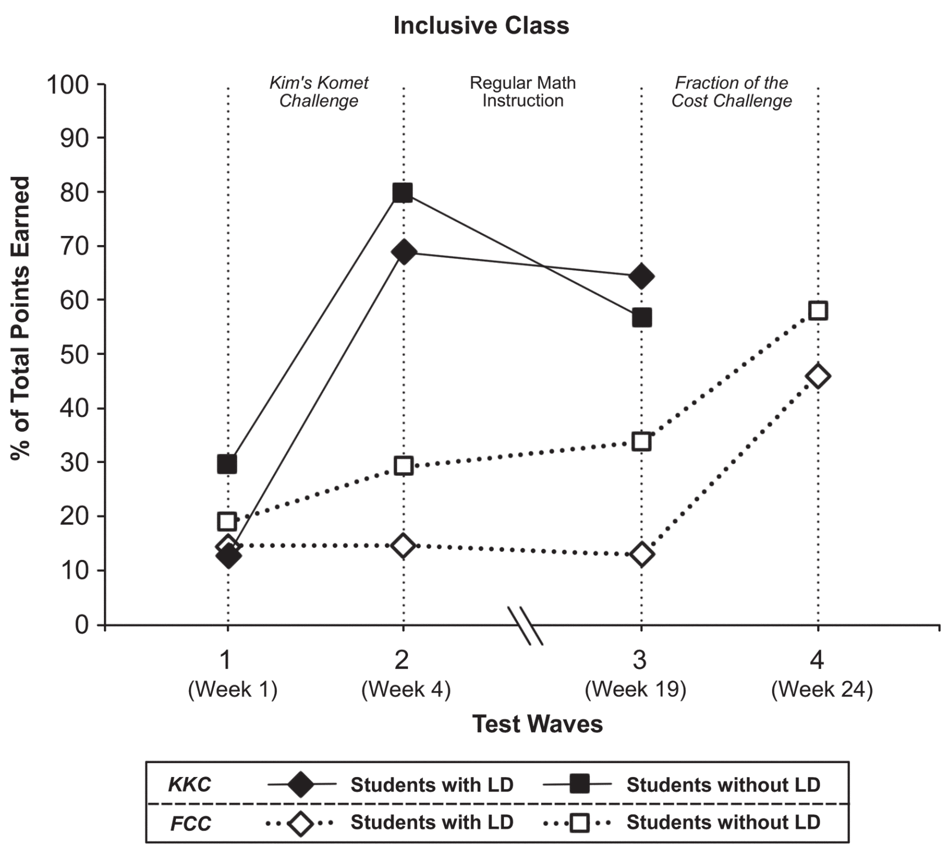



**FIGURE 5.** Performances of students with and without LD in the inclusive class.

**TABLE 7.** Achievement and Improvement Scores of Students in the Inclusive Class With LD and Without LD by Test Wave

| Measure and comparison | Contrast estimate | *SE* | Observed *t* | *p* value |
|---|---|---|---|---|
| *Kim's Komet Challenge* | | | | |
| Achievement | | | | |
| Wave 2 vs. Wave 1 | 18.86 | 1.23 | 15.36 | $p < .001$* |
| Wave 3 vs. Wave 1 | 14.32 | 1.90 | 7.55 | $p < .001$* |
| Improvement score non-LD vs. LD | | | | |
| Wave 2 vs. Wave 1 | −1.77 | 2.50 | −0.71 | $p = .487$ |
| Wave 3 vs. Wave 1 | −8.47 | 3.41 | −2.48 | $p = .022$* |
| *Fraction of the Cost Challenge* | | | | |
| Achievement | | | | |
| Wave 2 vs. Wave 1 | 1.71 | 1.07 | 1.60 | $p = .129$ |
| Wave 3 vs. Wave 1 | 2.06 | 1.26 | 1.64 | $p = .121$ |
| Wave 4 vs. Wave 3 | 10.65 | 1.49 | 7.17 | $p < .001$** |
| Improvement score non-LD vs. LD | | | | |
| Wave 2 vs. Wave 1 | 3.63 | 2.00 | 1.82 | $p = .09$ |
| Wave 3 vs. Wave 1 | 6.03 | 2.08 | 2.90 | $p = .011$** |
| Wave 4 vs. Wave 3 | −2.88 | 2.98 | −0.97 | $p = .350$ |

*Note*. Non-LD = non–learning disabled; LD = learning disabled.
*Significant, $\alpha = .05$ familywise, Fisher's LSD. **Significant, $\alpha = .05$ familywise, Holm.

fect score. On Item 6, none of the students with LD earned credit prior to instruction, whereas all of them received at least partial credit, and 94% earned full credit after instruction. In fact, a greater percentage of students with LD understood this item than did the students without LD who were in the same classroom. However, Item 7 showed that students with LD were much less successful in constructing a graph to represent their data.

On the *FCC* posttest (Test Wave 4), similar patterns emerged. All the students were at least partially successful in computing the amount of money each friend would contribute to buy materials for the skateboard ramp. Students with LD learned the most from instruction that focused on developing a materials list from a schematic plan, reading a tape measure to ⅛", converting feet and inches to inches alone, and comparing lengths when they are provided as mixed numbers (Item 3). However, most of the students in all ability groups had considerable difficulty when it came to showing the most economical way of cutting the lengths of wood (Item 4). If students were unable to figure out this item, they could not compute the cost of the project correctly.

# Discussion

The overall goal of this study was to measure the effectiveness of a reform-oriented instructional method called EAI on the achievement of middle school students. Specific objectives of the research were twofold. First, we tested the effects of EAI on the problem-solving performances of students without disabilities who were average and high achievers. Then, using their scores as performance benchmarks, we measured the treatment effects of EAI on students with LD.

## Summary of Findings

Results indicated that students in all six classes, including students with LD, benefited from both sets of EAI problems. On the *FCC*, which assessed linear measurement, computing fractions, and reading three-dimensional schematic plans, the obtained results followed predicted patterns. That is, students in the pre-algebra class scored higher than did students in both the typical and inclusive classes, and students in the typical classes scored higher than did the combination of students with and without LD in the inclusive class. However, there were no differences between groups in the size of improvement. As for the students with LD, they scored lower than their classmates without LD did, but there was no difference between the groups in improvement.

A somewhat different pattern of results was found for *KKC*. As on the *FCC*, the pre-algebra students outscored students in the typical and inclusive classes, but students in the inclusive class scored higher than students in the typical classes did. Furthermore, scores at follow-up 14 weeks after

**TABLE 8.** Percentage of Students Earning Partial or Full Credit on the *Kim's Komet Challenge (KKC)*

| | | Class/achievement | | | | |
|---|---|---|---|---|---|---|
| Concepts measured | Test wave | Pre-algebra | Typical | Inclusive non-LD | Inclusive LD | Total |
| Single/Two-step word problems | 1 | 100 | 83 | 100 | 82 | 91 |
| $d = r \times t$ | 2 | 100 | 94 | 100 | 73 | 92 |
| $r = d/t$ | | | | | | |
| (2 points) | Difference | 0 | +11 | 0 | -9 | +1 |
| Interpret line graphs, 1 and 2 data | 1 | 100 | 100 | 91 | 91 | 95 |
| lines, distance vs. time | 2 | 100 | 100 | 91 | 100 | 98 |
| (4 points) | Difference | 0 | 0 | 0 | +9 | +3 |
| Interpret data tables, compute | 1 | 92 | 74 | 73 | 55 | 73 |
| and compare rates (Item #5) | 2 | 100 | 97 | 100 | 91 | 97 |
| (16 points) | Difference | +8 | +23 | +27 | +36 | +24 |
| Interpret figures, compute rates, | 1 | 8 | 6 | 9 | 0 | 6 |
| understand concepts of constant | 2 | 100 | 62 | 91 | 100 | 88 |
| speed and acceleration (Item 6) | | | | | | |
| (6 points) | Difference | +92 | +56 | +82 | +100 | +82 |
| Construct graph, scale data, label | 1 | 65 | 36 | 36 | 9 | 37 |
| x and y axes, plot rates, draw line | 2 | 100 | 80 | 91 | 55 | 81 |
| of best fit (depends on previous item) | | | | | | |
| (Item #7) | | | | | | |
| (6 points) | Difference | +35 | +44 | +55 | +46 | +44 |
| Predict rates (answer depends on | 1 | 8 | 2 | 0 | 0 | 2 |
| previous item) | 2 | 54 | 41 | 45 | 9 | 37 |
| (2 points) | Difference | +46 | +39 | +45 | +9 | +35 |

*Note.* Non-LD = non–learning disabled; LD = learning disabled.

EAI showed that students in the inclusive class widened their performance margin and did not lose ground to the high-achieving students in the pre-algebra class. Comparing just the students in the inclusive class, students with LD scored slightly lower than did their classmates at both pretest and posttest, but their improvement from pretest to maintenance test was larger than that of their classmates without LD.

Item analyses also showed that students at all performance levels learned important skills and concepts, such as interpreting data tables and line graphs, calculating and comparing rates, interpreting schematic plans, and working with and converting linear measures. However, close inspection of these findings also revealed that many of the students, even those who were considered to be high achieving, had difficulty integrating several data sources into a descriptive representation of the overall problem solution. The *Fraction of the Cost* problem seemed especially difficult because the solution depended on understanding and working with fractions, a commonly reported area of misunderstanding for many students (Empson, 2003; Mack, 1995; Saxe, Taylor, McIntosh, & Gearhart, 2005).

## *Theoretical and Practical Implications*

Deshler et al. (2001) suggested that the achievement gap of students with LD in secondary inclusive classes could be reduced by (1) developing curricular materials that are appropriate for diverse groups of students and (2) training teachers to use it in ways students understand. On the surface, these goals appear to be common sense, but this study and previous EAI research have demonstrated that the path toward achieving these goals is neither easy nor straightforward. Although our findings suggest that EAI can address some of the learning needs of diverse groups of students, including students with LD, they also reveal the complexity involved in integrating cognitive supports in both the learning materials (e.g., multi-

**TABLE 9.** Percentage of Students Earning Partial or Full Credit on the *Fraction of the Cost Challenge (FCC)*

| | | Class/achievement | | | | |
|---|---|---|---|---|---|---|
| **Concepts measured** | **Test wave** | **Pre-algebra** | **Typical** | **Inclusive non-LD** | **Inclusive LD** | **Total** |
| Computing money each student contributed (Compute money from bank statement, compute percent) | 3 | 100 | 100 | 100 | 100 | 100 |
| | 4 | 100 | 100 | 100 | 100 | 100 |
| (3 points) | Difference | 0 | 0 | 0 | 0 | 0 |
| Interpret schematic plans (to 1/4 inch), complete materials list, convert feet and inches to inches. | 3 | 89 | 64 | 100 | 56 | 87 |
| | 4 | 100 | 97 | 88 | 89 | 92 |
| (12 points) | Difference | +11 | +33 | −12 | +33 | +5 |
| Read tape measure to 1/8 inch, complete data table (quantity, description, length in feet, length in inches) (Item #3) | 3 | 100 | 90 | 100 | 56 | 90 |
| | 4 | 95 | 97 | 88 | 89 | 95 |
| (4 points) | Difference | −5 | +7 | −12 | +33 | +5 |
| Compare and label lengths of wood (depends on constructing data table in previous item) | 3 | 72 | 28 | 50 | 44 | 49 |
| | 4 | 94 | 80 | 75 | 56 | 76 |
| (1 point) | Difference | +22 | +52 | +25 | +12 | +27 |
| Show most economical combinations, add and subtract mixed numbers, interpret figures (depends on answer in previous two items) | 3 | 22 | 0 | 13 | 0 | 9 |
| | 4 | 89 | 53 | 50 | 44 | 59 |
| (5 points) | Difference | +67 | +53 | +37 | +44 | +50 |
| Compute/Label combinations of wood, add and subtract mixed numbers, interpret tables and figures (depends on answers to previous items) (Item #4) | 3 | 17 | 3 | 0 | 0 | 5 |
| | 4 | 78 | 42 | 38 | 22 | 45 |
| (6 points) | Difference | +61 | +39 | +38 | +22 | +40 |
| Computing total cost, computing subtotals for materials (screws, dimension lumber), calculating sales tax (correct answer depends on correct answers to previous four items) | 3 | 0 | 0 | 0 | 0 | 0 |
| | 4 | 44 | 17 | 25 | 0 | 22 |
| (6 points) | Difference | +44 | +17 | +25 | 0 | +22 |

*Note*. Non-LD = non–learning disabled; LD = learning disabled.

media tools) and the instructional methods (e.g., managing problem-solving groups).

The need for such supports is based on cognitive load theory (CLT; Chandler & Sweller, 1991; Sweller, 1988), which, in its simplest form, suggests that all learners have limited working memory and that learning tasks should be structured in ways that do not overload it. One important consideration relates to what is called *element interactivity,* the extent to which relevant elements of content embedded in curricular materials interact. High interactivity materials cannot be fully understood until all the elements are processed simultaneously. Adding story contexts, visual representations, and multimedia applications to materials at this level of complexity can be extremely helpful in reducing cognitive overload (Mayer & Moreno, 2003; Mousavi, Low, & Sweller, 1995; Rittle-Johnson & Koedinger, 2005; Tabbers, Martens, & van Merrienboer, 2004).

The EAI problems in this study and those frequently encountered in life outside of school would most likely be situated high on the interactivity continuum. In some of our previous research, the sole use of the video-based anchor, in the absence of additional opportunities afforded by the learning tools and hands-on applications, did not meet the learning needs of some low-achieving students. Recent updates to *Fraction of the Cost* include several cognitive supports, among these a virtual tape measure, an interactive project in which students build and fly a hovercraft, and an explicit instruction module for helping students understand how to compute fractions. All of these were necessary additions to the EAI curriculum.

In addition to the multimedia tools, many students needed to "see" the problems (i.e., hovercraft, car pentathlon) in real life. In poststudy conversations, the math teachers thought the car pentathlon competition was the most effective tool of all. Students used their *smart tool*, or graph, to help them predict the correct height from which to release their cars to navigate the stunts at the end of the straightaway. Commenting on the performance of his students in his inclusive class, the math teacher pointed out, "the people who won [the pentathlon event] are normally people who are not successful in math. The people in my classes—first, second, and third—were people who never got the highest grades and had more difficulties, even the kids in my algebra class. The girl who won it normally comes on a daily basis and has a lot wrong on her homework. But she used her smart tools."

The teachers also had to be well prepared to deal with the complexity of the materials and the needs of their students. Recent studies have shown that content knowledge and the skills used to teach the math content are important contributors to student achievement in general education (e.g., Hill, Rowan, & Ball, 2005). The teachers had to anticipate the "bugs" in their students' problem-solving procedures to know when and how to redirect their efforts. Second, they had to know where to find the learning tools in the software and how to operate the other technology tools, such as the infrared timer used in the pentathlon. Finally, the math and special education teachers had to know how to facilitate the work of students in small groups so that all students had opportunities for giving their input, an especially important skill in the inclusive classroom.

The relative lack of training and experience of one of the math teachers may explain why the scores of typical achievers were lower on *KKC* than those of students in the inclusive class on the posttest and maintenance test waves. It was the first time this teacher had taught *Kim's Komet* and it is possible, if not probable, that she was unfamiliar with all the nuances involved in teaching and learning the EAI problems. Although observational data showed that she was an excellent teacher, unfamiliarity with the technologies and the subtleties of the problems may have led to her students' lower scores.

Finally, we comment on the complexity of the research endeavor itself as it relates to intervention research. As special educators have pointed out (e.g., Gersten, 2005; Graham, 2005; Palincsar, 2005), this type of research takes considerable time, from the theoretical justification for developing the new or modified intervention, to development and pilot work, to research design and implementation, and finally to the final analyses and reporting. As more students with disabilities are included in general education classrooms, developing and implementing curricula that meets the needs of students with and without disabilities will continue to present a formidable challenge, especially at the secondary level (Conderman & Katsiyannis, 2002; Fuchs & Fuchs, 1998).

## Limitations

These findings suggest that EAI can have important effects on learning, especially for students with LD, but we temper our results with the following limitations. First, the research design had advantages, as we noted earlier, but it also had several disadvantages. Probably most important, it did not assess the efficacy of the EAI approach versus other forms of anchored instruction or typical instruction, although comparisons have been made in previous work (e.g., Bottge, 1999; Bottge, Heinrichs, Chan, & Serlin, 2001). The design also required students to be tested several times on concepts they had not yet learned. Double pretests and repeated measurements helped minimize threats to internal validity (e.g., history), but teachers had to attend closely to student morale at all test waves.

Second, the outcome measures were aligned closely to situations taught during EAI, and thus we do not know to what extent students could generalize their understandings to new contexts. On the other hand, the pencil–paper format of the tests may have limited students' ability to show all they had learned. Judging from the scores of even the highest achievers, the tests were difficult. One reason for this may be a mismatch between learning and testing formats. Although the concepts tested were aligned with the instructional content, the paper tests did not offer students possibilities for searching problem contexts in the same way as during instruction. Creating interesting and meaningful problem contexts in paper format may have important implications for all students, especially for those who find it particularly difficult to picture text-based problems.

Third, although all of the seventh-grade students with LD were included in the study, the number was quite small and thus may have contributed to relatively low power. We frequently encounter this difficulty when conducting intervention studies in which there are few students with disabilities in each inclusive classroom. For this reason, we have tried to provide adequate detail of the interventions, measures, and results.

## Conclusion

Ceci and Roazzi (1994) made the insightful observation, "we cannot conclude that children lack certain cognitive abilities just because they do not exhibit them in a given context" (p. 93). We are cautiously optimistic about the possibilities technology-assisted, well-structured learning environments can have in reducing the differences between students with diverse skill levels. A decade ago, Darling-Hammond (1996) suggested educators should develop teaching methods that go far beyond dispensing information, focus on challenging curriculum goals for all students (i.e., be *learning centered*), and meet the needs and interests of individual learners (i.e., be *learner centered*). In this study, we have attempted to realize these goals by providing constructed contexts in which students at all achievement levels could succeed.

## AUTHORS' NOTE

The research reported in this article was supported by a grant from the U.S. Department of Education, Office of Special Education and Rehabilitative Services, Technology and Media Services for Individuals with Disabilities, Project No. H327A020030-04. Any opinions, findings, or conclusions are those of the authors and do not necessarily reflect the views of the supporting agency.

## REFERENCES

Baker, J. M., & Zigmond, N. (1990). Are regular education classes equipped to accommodate students with learning disabilities? *Exceptional Children, 56,* 515–526.

Bottge, B. A. (1999). Effects of contextualized math instruction on problem solving of average and below-average achieving students. *The Journal of Special Education, 33,* 81–92.

Bottge, B. A. (2001). Reconceptualizing math problem solving for low-achieving students. *Remedial and Special Education, 22,* 102–112.

Bottge, B. A., Heinrichs, M., Chan, S., & Serlin, R. (2001). Anchoring adolescents' understanding of math concepts in rich problem solving environments. *Remedial and Special Education, 22,* 299–314.

Bottge, B. A., Heinrichs, M., Mehta, Z., & Hung, Y. (2002). Weighing the benefits of anchored math instruction for students with disabilities in general education classes. *The Journal of Special Education, 35,* 186–200.

Bottge, B. A., Heinrichs, M., Mehta, Z. D., Rueda, E., Hung, Y., & Danneker, J. (2004). Teaching mathematical problem solving to middle school students in math, technology education, and special education classrooms. *Research in Middle Level Education Online, 27*(1), 1–17. (Available at http://www.nmsa.org/Publications/RMLEOnline/Articles/Vol27No1Article1/tabid/529/Default.aspx)

Butler, F. M., Miller, S. P., Crehan, K., Babbitt, B., & Pierce, T. (2003). Fraction instruction for students with mathematics disabilities: Comparing two teaching sequences. *Learning Disabilities Research & Practice, 18,* 99–111.

Carnine, D. (1998). Instructional design in mathematics for students with learning disabilities. In D. P. Rivera (Ed.), *Mathematics education for students with learning disabilities: Theory to practice* (pp. 119–138). Austin, TX: PRO-ED.

Ceci, S. J., & Roazzi, A. (1994). The effects of context on cognition: Postcards from Brazil. In R. J. Sternberg & R. K. Wagner (Eds.), *Mind in context: Interactionist perspectives on human intelligence* (pp. 74–101). New York: Cambridge University Press.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8,* 293–332.

Cognition and Technology Group at Vanderbilt University. (1990). Anchored instruction and its relationship to situated cognition. *Educational Researcher, 19*(3), 2–10.

Cognition and Technology Group at Vanderbilt University. (1997). *The Jasper Project: Lessons in curriculum, instruction, assessment, and professional development.* Mahwah, NJ: Erlbaum.

Conderman, G., & Katsiyannis, A. (2002). Instructional issues and practices in secondary special education. *Remedial and Special Education, 23,* 169–179.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings.* Boston: Houghton Mifflin.

CTB/McGraw-Hill. (1997). *TerraNova Comprehensive Test of Basic Skills.* Monterey, CA: Author.

Dale Seymour Publications. (2004). *Connected mathematics.* Upper Saddle River, NJ: Prentice Hall.

Darling-Hammond, L. (1996). The right to learn and the advancement of teaching: Research, policy, practice for democratic education. *Educational Researcher, 25*(6), 5–17.

Deshler, D. D., Schumaker, J. B., Lenz, B. K., Bulgren, J. A., Hock, M. F., Knight, J., et al. (2001). Insuring content-area learning by secondary students with learning disabilities. *Learning Disabilities Research & Practice, 16,* 96–108.

Empson, S. B. (2003). Low-performing students and teaching fractions for understanding: An interactional analysis. *Journal for Research in Mathematics Education, 34,* 305–343.

Fuchs, D., & Fuchs, L. S. (1995). What's special about special education? *Phi Delta Kappan, 76,* 522–530.

Fuchs, D., & Fuchs, L. S. (1998). Researchers and teachers working together to adapt instruction for diverse learners. *Learning Disabilities Research & Practice, 13,* 126–137.

Fuchs, L. S., & Fuchs, D. (2002). Mathematical problem-solving profiles of students with mathematics disabilities with and without comorbid reading difficulties. *Journal of Learning Disabilities, 35,* 563–573.

Gersten, R. (2005). Behind the scenes of an intervention research study. *Learning Disabilities Research & Practice, 20,* 200–212.

Graham, S. (2005). Further behind the scenes of intervention research. *Learning Disabilities Research & Practice, 20,* 221–224.

Hickey, D. T., Moore, A. L., & Pellegrino, J. W. (2001). The motivational and academic consequences of elementary mathematics environments: Do constructivist innovations and reforms make a difference? *American Educational Research Journal, 38,* 611–652.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42,* 371–406.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6,* 65–70.

Hung, Y-H. (2005). *The effects of prior knowledge on mathematics learning of adolescents with or without learning disabilities.* Unpublished doctoral dissertation. University of Wisconsin–Madison.

Hutchinson, N. L. (1993). Effects of cognitive strategy instruction on algebraic problem solving of adolescents with learning disabilities. *Learning Disability Quarterly, 16,* 34–63.

International Technology Education Association. (2000). *Standards for technological literacy: Content for the study of technology.* Reston, VA: Author.

Jitendra, A., DiPipi, C. M., & Perron-Jones, N. (2002). An exploratory study of schema-based word-problem-solving instruction for middle school students with learning disabilities: An emphasis on conceptual and procedural understanding. *The Journal of Special Education, 36,* 23–38.

Jitendra, A. K., Hoff, K., & Beck, M. M. (1999). Teaching middle school students with learning disabilities to solve word problems using a schema-based approach. *Remedial and Special Education, 20,* 50–64.

Jones, E. D., Wilson, R., & Bhojwani, S. (1997). Mathematics instruction for secondary students with learning disabilities. *Journal of Learning Disabilities, 30,* 151–163.

Learning Technology Center at Vanderbilt University. (1997). *The new adventures of Jasper Woodbury.* Mahwah, NJ: Erlbaum.

Lesh, R., & Kelly, A. (2000). Multitiered teaching experiments. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 197–230). Mahwah, NJ: Erlbaum.

Lester, F. K., Jr., & Kroll, D. L. (1990). Assessing student growth in mathematical problem solving. In G. Kulm (Ed.), *Assessing higher order thinking in mathematics* (pp. 53–70). Washington, DC: American Association for the Advancement of Science.

Maccini, P., & Hughes, C. A. (2000). Effects of a problem-solving strategy on the introductory algebra performance of secondary students with learning disabilities. *Learning Disabilities Research & Practice, 15,* 10–21.

Mack, N. K. (1995). Confounding whole-number and fraction concepts when building on informal knowledge. *Journal for Research in Mathematics Education, 26,* 422–441.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychology, 38*(1), 43–52.

Montague, M. (1997). Cognitive strategy instruction in mathematics for students with learning disabilities. *Journal of Learning Disabilities, 30,* 164–177.

Montague, M. (1998). Cognitive strategy instruction in mathematics for students with learning disabilities. In D. P. Rivera (Ed.), *Mathematics education for students with learning disabilities* (pp. 177–199). Austin, TX: PRO-ED.

Montague, M., & Bos, C. (1986). The effect of cognitive strategy training on verbal math problem solving performance of learning disabled adolescents. *Journal of Learning Disabilities, 19,* 26–33.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87,* 319–334.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

No Child Left Behind Act of 2001, 20 U.S.C. 70 § 6301 *et seq.* (2002)

Palincsar, A. S. (2005). Working theory into and out of design experiments. *Learning Disabilities Research & Practice, 20,* 218–220.

Perie, M., Grigg, W., & Dion, G. (2005). *The Nation's Report Card: Mathematics 2005* (NCES 2006-453). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Rea, P. J., McLaughlin, V. L., & Walther-Thomas, C. (2002). Outcomes for students with learning disabilities in inclusive and pullout programs. *Exceptional Children, 68,* 203–222.

Rittle-Johnson, B., & Koedinger, K. R. (2005). Designing knowledge scaffolds to support mathematical problem solving. *Cognition and Instruction, 23,* 313–349.

Saxe, G. B., Taylor, E. V., McIntosh, C., & Gearhart, M. (2005). Repesenting fractions with standard notation: A developmental analysis. *Journal for Research in Mathematics Education, 36,* 137–157.

Seaman, M. A., Levin, J. R., & Serlin, R. C. (1991). New developments in pairwise multiple comparisons: Some powerful and practicable procedures. *Psychological Bulletin, 110,* 577–586.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Shafer, M. C., & Romberg, T. A. (1999). Assessment in classrooms that promote understanding. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 159–184). Mahwah, NJ: Erlbaum.

Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education, 36,* 404–411.

Sulzer-Azaroff, B., & Mayer, G. R. (1977). *Applying behavior analysis procedures with children and youth.* New York: Holt, Rinehart & Winston.

Swanson, H. L. (2001). Research on intervention for adolescents with learning disabilities: A meta-analysis of outcomes related to high-order processing. *The Elementary School Journal, 101,* 331–348.

Swanson, H. L., & Deshler, D. (2003). Instructing adolescents with learning disabilities: Converting a meta-analysis to practice. *Journal of Learning Disabilities, 36,* 124–135.

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.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12,* 257–285.

Tabbers, H. K., Martens, R. L., & van Merrienboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *The British Journal of Educational Psychology, 74,* 71–81.

University of Iowa. (2001). *The Iowa tests of basic skills* (ITBS, *Form A).* Itasca, IL: Riverside.

U.S. Department of Education. (1994). Goals 2000: Educate America Act, 20 U.S.C. § 5801.

U.S. Department of Labor. (1991). *What work requires of schools: A SCANS report for America 2000.* Washington, DC: Author.

Vaughn, S., Klingner, J., & Hughes, M. (2000). Sustainability of research-based practices. *Exceptional Children, 66,* 163–171.

Wisconsin Department of Public Instruction. (2006). *Specific Learning Disabilities Eligibility Criteria Resources.* Retrieved October 23, 2006, from http://dpi.wi.gov/sped/eligild.html

Witzel, B. S., Mercer, C. D., & Miller, M. D. (2003). Teaching algebra to students with learning difficulties: An investigation of an explicit instruction model. *Learning Disabilities Research & Practice, 18,* 121–131.

Woodward, J. (2004). Mathematics education in the United States: Past to present. *Journal of Learning Disabilities, 37,* 16–31.

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

Woodward, J., & Montague, M. (2002). Meeting the challenge of mathematics reform for students with LD. *The Journal of Special Education, 36,* 89–101.