COGNITION AND INSTRUCTION, *23*(3), 313–349




# Designing Knowledge Scaffolds to Support Mathematical Problem Solving

Bethany Rittle-Johnson
*Vanderbilt University*

Kenneth R. Koedinger
*Carnegie Mellon University*

We present a methodology for designing better learning environments. In Phase 1, 6th-grade students' ($n = 223$) prior knowledge was assessed using a difficulty factors assessment (DFA). The assessment revealed that scaffolds designed to elicit contextual, conceptual, or procedural knowledge each improved students' ability to add and subtract fractions. Analyses of errors and strategies along with cognitive modeling suggested potential mechanisms underlying these effects. In Phase 2, we designed an intervention based on scaffolding this prior knowledge and implemented the computer-based lessons in mathematics classes. In Phase 3, we used the DFA and supporting analyses to assess student learning from the intervention. The posttest results suggest that scaffolding conceptual, contextual, and procedural knowledge are promising tools for improving student learning.

Well-structured, organized knowledge allows people to solve novel problems and to remember more information than do memorized facts or procedures (see Bransford, Brown, & Cocking, 2001, for a review). Such well-structured knowledge requires that people integrate their contextual, conceptual and procedural knowledge in a domain. Unfortunately, U.S. students rarely have such integrated and robust knowledge in mathematics or science (Beaton et al., 1996; Reese, Miller, Mazzeo, & Dossey, 1997). Designing learning environments that support integrated knowledge is a key challenge for the field, especially given the low number of established tools for guiding this design process (Lesh, Lovitts, & Kelly, 2000). In this article, we use our design of a lesson within a larger design ex-

Requests for reprints should be sent to Bethany Rittle-Johnson, Vanderbilt University, 230 Appleton Place, Peabody #512, Nashville, TN 37203. E-mail: bethany.rittle-johnson@vanderbilt.edu

periment to illustrate the combination of a variety of methodologies drawn from cognitive science and education to guide and evaluate our design.

In the introduction, we first provide an overview of our methods for designing and evaluating a learning environment. Next, we review previous research on three critical types of knowledge that we tried to elicit in our learning environment: contextual, conceptual, and procedural knowledge. Finally, we describe this study on adding and subtracting fractions with sixth graders, including the scaffolds used to elicit each type of knowledge.



## DESIGNING LEARNING ENVIRONMENTS

Appropriate methods for designing and evaluating learning environments are still emerging (Barab & Squire, 2004; Lesh et al., 2000). In this study, we illustrate the potential benefits of combining four key methods drawn from cognitive science and educational research for informing the design, evaluation, and revision of learning environments. First, we used difficulty factors assessment (DFA) and analyses of errors and strategy use to identify students' prior knowledge.

DFA can be used to identify what problem features (i.e., factors) facilitate problem solving. Target factors are systematically varied and crossed with specific problem contexts and numerical values so that the factors are not confounded with these other problem features, leading to multiple versions of the assessment (Koedinger & Nathan, 2004). For example, to evaluate the impact of having a story context, one version of the assessment might ask children to add $\frac{1}{2} + \frac{2}{3}$ in the context of the story problem and to add $\frac{3}{4} + \frac{1}{7}$ without a story context. The format of the two questions would vary only in the presence of the story context. A second version of the assessment would do the opposite: add $\frac{3}{4} + \frac{1}{7}$ in the context of a story problem and $\frac{1}{2} + \frac{2}{3}$ without a story context. Collapsing data across the two versions provides estimates of student accuracy when a problem does and does not have a story context, and these estimates are not confounded by potential differences in the difficulty of particular numbers used. In this study, we used DFA to evaluate the effects of three different knowledge scaffolds on problem-solving accuracy.

Analyses of students' errors and strategy use on the DFA provide further insights into how each problem factor impacts problems solving. Students' errors can be used to infer students' incorrect problem-solving strategies (e.g., Siegler, 1976). Additional information on students' strategies can be elicited via requests for verbal self-reports or through analysis of overt behavior, including written work (e.g., Koedinger & Nathan, 2004; Rittle-Johnson & Siegler, 1999).

Evidence regarding the factors that facilitate problem solving and clues regarding how these factors influence the problem-solving process can be used to inform a cognitive task analysis of the domain. The goal of a cognitive task analysis is to elucidate the likely decision processes, knowledge, and procedural skills needed to complete a task (Lovett, 1998; Schraagen, Chipman, & Shalin, 2000). Such an analysis can improve the design of instruction and assessment (e.g., Klahr & Carver, 1988). To increase its power, a cognitive task analysis can be instantiated in a cognitive model designed to simulate the proposed problem-solving process. This model precisely articulates potential mechanisms and provides an avenue for making predictions that are empirically falsifiable (Anderson, 1993; Koedinger & MacLaren, 2002). We use a form of cognitive modeling we call *knowledge-component modeling,* which is a middle ground between a paper-and-pencil task analysis and a completely implemented production system model. Knowledge-component modeling yields many of the benefits of a production system simulation but with lower resource costs. Our knowledge-component modeling of student problem solving yielded a detailed exploration of potential mechanisms underlying the impact of each scaffold on problem-solving difficulty.

Thus, results from the DFA, analyses of students' errors and strategies, and cognitive modeling provide a detailed description of students' current problem-solving processes. In turn, this description provides a foundation for the design of instruction to improve children's knowledge. For example, it identifies prior knowledge possessed by many children that may be useful for bridging to (i.e., scaffolding) more complete knowledge of the domain.

This information is particularly valuable in combination with a design experiment. Design experiments involve iterative redesign over the course of the project based on observations and assessments gained while the study is being conducted (e.g., Cobb, McClain, & Gravemeijer, 2003). During this iterative redesign, it can be difficult to disentangle which features of the design experiment are driving its effectiveness and thus, which features should be kept and which should be altered. Identifying these key features is crucial for developing a pedagogical domain theory, for extending the findings to new domains, and for contributing to theory on how people learn (A. L. Brown, 1992). The combination of using DFA, strategy and error analyses, and cognitive modeling at multiple phases of the design experiment can provide such data. In this study, a DFA and supporting analyses were conducted at the beginning of one unit within a design experiment and at the end of the first iteration of the unit to help identify critical design features and to provide suggestions for revisions.

## TARGET KNOWLEDGE SOURCES

Our initial design was guided by prior research in cognitive science and mathematics education on the importance of at least three types of knowledge for problem solving: contextual, conceptual, and procedural knowledge. We consider the role of each type of knowledge following.

## Contextual Knowledge

First, consider the role of contextual knowledge in problem solving. This is our knowledge of how things work in specific, real-world situations, which develops from our everyday, informal interactions with the world (J. S. Brown, Collins, & Duguid, 1989; Leinhardt, 1988; Mack, 1990; Saxe, 1988). Students' contextual knowledge can be elicited by situating problems in story contexts. In contrast, problems can be presented symbolically using only mathematical symbols such as numerals, operators, and variables.

There are contradictory claims for whether presenting problems in story contexts facilitates or hampers problem solving. A commonly held belief among the general public, textbook authors, teachers, mathematics education researchers, and learning science researchers is that story problems are harder than symbolic problems (Nathan & Koedinger, 2000; Nathan, Long, & Alibali, 2002). This belief is supported in part by national assessment data indicating that elementary-school children in the United States generally do worse on story problems compared to similar symbolic computation problems (Carpenter, Corbitt, Kepner, & Reys, 1980; Kouba, Carpenter, & Swafford, 1989). Systematic studies of U.S. elementary-school children solving single-digit arithmetic problems also indicate that children often do better on symbolic problems than comparable story problems (e.g., Cummins, Kintsch, Reusser, & Weimer, 1988; Riley, Greeno, & Heller, 1983). Linguistic difficulties, rather than insufficient mathematical knowledge, seem to account for young children's overall poorer performance on arithmetic story problems compared to symbolic problems (Briars & Larkin, 1984; Cummins et al., 1988; de Corte, Verschaffel, & de Win, 1985; Hudson, 1983; Kintsch & Greeno, 1985; Riley et al., 1983).

In other instances, story problems can be easier to solve than symbolic problems. In particular, in multidigit arithmetic and early algebra, children are often more successful at solving problems in familiar story contexts than comparable problems presented symbolically (Baranes, Perry, & Stigler, 1989; Carraher, Carraher, & Schliemann, 1985, 1987; Koedinger & Nathan, 2004; Saxe, 1988). The benefits of story contexts arise because they elicit alternative, informal solution strategies and/or improved problem comprehension. These findings converge with learning theories that have emphasized the role of contextual knowledge in supporting the development of symbolic knowledge (e.g., Greeno, Collins, & Resnick, 1996; Vygotsky, 1978).

How can we resolve these seemingly contradictory empirical results and theoretical perspectives on the relative difficulty of story and symbolic problems? The explanation for differences across domains (e.g., single-digit addition vs. algebra) may involve at least two key factors: (a) differences in children's comprehension of the target words and symbols and (b) differences in their facility with symbolic strategies. These differences are largely based on children's prior exposure to the target language and symbol systems. First, young children have pervasive exposure to single-digit numerals, but some words and syntactic forms are still unknown or unfamiliar. In comparison, older children have less exposure to large, multidigit numerals and algebraic symbols and have much better reading and comprehension skills. Second, by early elementary school, many children have mastered multiple strategies for adding single-digit numerals (Siegler, 1987; Siegler & Jenkins, 1989). In comparison, older children frequently make errors when implementing symbolic multidigit arithmetic or equation-solving strategies (Fuson, 1990; Hiebert & Wearne, 1986; Sleeman, 1985)

In this study, we evaluated the predictive value of these two factors for a new age group and new domain—sixth-grade students learning to add and subtract fractions. Students' familiarity with fraction symbols (e.g., $\frac{2}{3}$) is midway between that of single-digit numerals and algebra symbols, and sixth graders are fairly proficient readers. By sixth grade, students can invent strategies for solving fraction problems presented in real-world contexts (Mack, 1990, 1993) and have typically learned but not mastered symbolic strategies for adding and subtracting fractions (Kouba et al., 1989). We predicted that well-designed story problems would elicit sixth graders' contextual knowledge and be easier to solve than symbolic fraction problems. Our claim is that although not all story contexts support learning and problem solving, some do.

## Conceptual Knowledge

Students also need to develop conceptual knowledge in a domain, which is integrated knowledge of important principles (e.g., knowledge of number magnitudes) that can be flexibly applied to new tasks. Conceptual knowledge can be used to guide comprehension of problems and to generate new problem-solving strategies or to adapt existing strategies to solve novel problems (Hiebert, 1986). Indeed, improving children's conceptual knowledge can lead them to use better problem-solving strategies (Perry, 1991; Rittle-Johnson & Alibali, 1999).

Although there is widespread agreement on the importance of conceptual knowledge, how to tap this knowledge and encourage students to integrate it with their contextual and procedural knowledge is less clear (Kilpatrick, Swafford, & Findell, 2001). Visual representations of problems (such as pictures and diagrams) are one potential scaffold for eliciting conceptual knowledge and facilitating integration. Within the domain of fractions, visual representations, such as fraction circles and fraction bars, help to illustrate key concepts (Cramer, Post, & del Mas, 2002; Fuson & Kalchman, 2002). In turn, past research in a variety of domains has indicated that visual representations can aid problem solving (e.g., Griffin, Case, & Siegler, 1994; Koedinger & Anderson, 1990; Larkin & Simon, 1987; Novick, 2001) and improve learning and transfer (Mayer, Mautone, & Prothero, 2002). For example, Fuson and Kalchman found that using a single visual representation of fractions facilitated students' understanding of fractions and allowed them to invent and refine strategies for multiplying fractions. In this study, we predicted that a visual representation of fractions (fraction bars) would improve learning and problem solving.

### Procedural Knowledge

Finally, consider a direct source of problem-solving knowledge—knowledge of subcomponents of a correct procedure. Procedures are a type of strategy that involve step-by-step actions for solving problems (Bisanz & LeFevre, 1990), and most procedures require integration of multiple skills. For example, the conventional procedure for adding fractions with unlike denominators requires knowing how to find a common denominator, how to find equivalent fractions, and how to add fractions with like denominators. Learning all of these subcomponents simultaneously may overwhelm students (Anderson, Corbett, Koedinger, & Pelletier, 1995; Sweller, 1988). A widely studied approach for avoiding this problem is to model expert performance on the task either directly by a teacher or tutor (Chiu, Chou, & Liu, 2002; Knapp & Winsor, 1998; Schoenfeld, 1985), by written cue cards (Mayer et al., 2002; Scardamalia & Bereiter, 1985), or by worked examples of problem solutions (Sweller, 1988). We predicted that providing a component of the conventional procedure would improve learning and problem solving.



## THIS STUDY

We evaluated the role of each of the three types of knowledge—conceptual, contextual, and procedural—for sixth-grade students' solving fraction addition and subtraction problems. Students' knowledge was evaluated before and after they completed computer-based classroom instruction that scaffolded the three types of knowledge. The unit was part of a larger design experiment on creating a sixth-grade Cognitive Tutor course (Koedinger, 2002).

National and international assessments indicate that although fourth-grade students have basic knowledge of fraction quantities, they have limited knowledge of more difficult concepts, such as equivalent fractions, and of correct procedures for computing with fractions (Kouba et al., 1989; Kouba, Zawojewski, & Strutchens, 1997). For example, on the mathematics section of the fourth National Assessment of Educational Progress, only 53% of seventh-grade students correctly solved the problem $3\frac{1}{2} - 3\frac{1}{3}$ (Kouba et al., 1989). In contrast to the relatively poor performance on written, symbolic problems, interviews with students have revealed students' informal knowledge of fractions, especially in everyday contexts (Empson, 1999; Leinhardt, 1988; Mack, 1990, 1993; Mix, Levine, & Huttenlocher, 1999; Pothier & Sawada, 1983). Eliciting students' prior, albeit fragmented, knowledge during problem solving should improve performance. To accomplish this, we designed a scaffold to elicit a representative component of contextual, conceptual, and procedural knowledge. The contextual scaffold involved situating the computation in a real-world, candy bar context. Past research on fraction learning indicates that food contexts are particularly meaningful contexts for students (Mack, 1990, 1993). During earlier lessons on comparing and finding equivalent fractions, discussing fractions as parts of a candy bar emerged as a meaningful way for our students to think about fractional quantities. Although a candy bar context may not be the ideal context for thinking about fraction computations, it did provide a familiar context for our students that had been developed in previous lessons.[1] A candy bar context also integrated with our visual representation of fractions—fraction bars. Fraction bars are a particularly useful visual representations for facilitating thinking about rational number concepts (Cramer et al., 2002; Fuson & Kalchman, 2002). For example, they illustrate the meaning of fractions as parts of a whole, the importance of equal-size parts for combining fractions, and the impossibility of adding two fractions and getting a fraction smaller than either addend. In prior lessons on fraction concepts, fraction bars were the visual representation that our students found particularly useful for linking to part–whole concepts, and thus, fraction bars may elicit their conceptual knowledge.[2] Finally, the procedural scaffold was to provide a common denominator. Students' most common error is adding the numerators and denominators rather than using a common denominator. Thus, providing a common denominator should signal to students that they should not add the denominators and should help them to implement a correct procedure.

To evaluate whether each scaffold facilitated addition and subtraction of fractions, we used DFA. We expected all three scaffolds to improve accuracy, reduce common errors, and improve use of effective strategies at pretest. After participating in an intervention containing all three scaffolds, we expected the scaffolds to have less impact on problem solving if the intervention had successfully facilitated knowledge integration, thus reducing the need for explicit scaffolding.

[1]We acknowledge that our candy bar context does not encompass the range of situations in which students would need to add and subtract fractions, is somewhat contrived (e.g., most candy bars cannot be broken into 20 pieces), and may be ambiguous to interpret at times (e.g., Is the whole one or two candy bars?). We were available to answer student questions and our students seemed to relate to the problems (partially because the context had been developed in previous lessons). However, in extending this work to other student populations, careful attention should be paid to the optimal contexts to use.

[2]It is possible that fraction bars did not tap children's conceptual knowledge. We return to this issue in the Discussion.

# METHOD

## Participants

Participants were 223 sixth-grade students drawn from two different populations to increase the generalizability of the findings. Students participated as part of their regular mathematics course, and thus, all students in the given classrooms participated. The first population was 137 students (69 female) from eight mathematics classrooms at an urban, public middle school. At the school, 78% of students were considered economically disadvantaged; approximately 50% were White; and 49% were African American. Average daily attendance in the school district was 89%. Students were using the National Council of Teachers of Mathematics standards-based curriculum *Connected Mathematics* (Lappan, Fey, Fitzgerald, Friel, & Phillips, 2002) 4 days a week in the classroom and using the intelligent tutoring system component of our sixth-grade Cognitive Tutor course 1 day a week.

The second population was 86 students (41 female) from four mathematics classrooms in two suburban public schools (one elementary and one middle school). At these schools, the percentage of economically disadvantaged students was 12% and 20%, respectively, and over 95% of the students were White. Average daily attendance in each of these school districts was 94%. At one school, the mathematics teacher was helping us to design the sixth-grade Cognitive Tutor curriculum, and he was simultaneously using the curriculum in his classroom. The course incorporated both a problem-based paper curriculum used in the classroom 3 days a week and an intelligent tutoring system used in the computer laboratory 2 days a week (Koedinger, Anderson, Hadley, & Mark, 1997). At the second school, the teacher was piloting the new curriculum. Thus, the student body, the classroom curriculum, and the teaching styles and attitudes varied across the schools.

## Procedure and Materials

The study was conducted in three phases. Phase 1 was to identify students' prior knowledge of fraction arithmetic using a pretest that incorporated a DFA. Analyses of student errors and strategy use and cognitive modeling of performance on the assessment was used to better understand students' prior knowledge. Phase 2 was to design and implement an intervention that built off of students' prior knowledge in the domain. The intervention was implemented via an intelligent tutoring system. Phase 3 was to administer a posttest that was identical to the pretest to assess learning from the intervention. Students participated in the three phases over approximately three, 40-min classroom sessions. Students went to their schools' computer laboratory for these sessions once or twice a week as part of their normal math class.

*Phase 1: Identifying prior knowledge.* To evaluate students' prior knowledge, students first solved eight problems on adding and subtracting fractions. Half of the problems were created using a DFA methodology. These four problems involved adding or subtracting two symbolic fractions with unlike denominators. Children completed one instance of each scaffold: none, contextual, conceptual, or procedural. A critical feature of a DFA is that factors of interest are not confounded with the particular calculations in the problem. Thus, four versions of the assessment were generated based on fully crossing the four scaffold types with the four required calculations as illustrated in Table 1. Figure 1 presents each of the scaffolds for the calculation $\frac{2}{3} - \frac{1}{9}$ and illustrates how the four problems varied only in the type of scaffold.

To strengthen the impact of the conceptual scaffold, we included fraction bars depicting common incorrect answers to help students recognize the conceptual violations produced by the errors. To do so, we made all the questions multiple choice, and each problem had four answer choices (see Figure 1): (a) correct, (b) combine-both error (combine both numerator and denominator), (c) fail-to-convert error (fail to convert numerators after finding a common denominator), and (d) other error.

In addition to the four problems created using the DFA methodology, there were four other problems designed to get a broader index of prior knowledge. They were adding two fractions with the same denominator ($\frac{3}{7} + \frac{2}{7}$), adding three fractions ($\frac{9}{10} + \frac{1}{5} + \frac{1}{2}$), subtracting mixed numbers ($13\frac{1}{2} - 3\frac{1}{8}$), and identifying a verbal description of the conventional procedure. These four problems were the same across the four versions of the assessment.

The four versions of the assessment were randomly distributed to the students at the beginning of the study, with 54 to 57 students completing a particular version. Students spent approximately 10 min completing the assessment.

Students' responses on the assessment were coded as one of the four response categories outlined previously. If students wrote "Don't know" or left the answer blank, the response was coded as an other error. Students' strategy use was coded from their written work, using the codes described in Table 2. Unfortunately, students' written work was not available for two out of the four classes at the suburban

TABLE 1
Difficulty Factors Assessment Design: Distribution of Scaffolds Across the Four Fraction Problems With Unlike Denominators

| *Form* | $\frac{3}{4} + \frac{2}{12}$ | $\frac{1}{4} + \frac{1}{5}$ | $\frac{2}{3} + \frac{1}{9}$ | $\frac{1}{2} + \frac{2}{7}$ |
|---|---|---|---|---|
| 1 | Contextual | None | Procedural | Conceptual |
| 2 | Conceptual | Contextual | None | Procedural |
| 3 | Procedural | Conceptual | Contextual | None |
| 4 | None | Procedural | Conceptual | Contextual |

a) No scaffold

**Subtract:** $\mathbf{\frac{2}{3} - \frac{1}{9} =}$

a. $\frac{1}{6}$ b. $\frac{5}{9}$ c. $\frac{1}{9}$ d. Other: ______

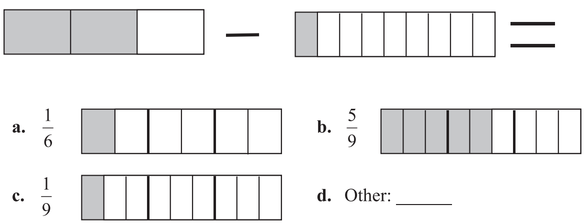
b) Contextual scaffold

Mrs. Jules bought each of her children a chocolate bar. Jarren ate $\mathbf{\frac{2}{3}}$ of a chocolate bar and Alicia ate $\mathbf{\frac{1}{9}}$ of a chocolate bar. How much **more** of a chocolate bar did Jarren eat than Alicia**?**

**a.** $\frac{1}{6}$ **b.** $\frac{5}{9}$ **c.** $\frac{1}{9}$ **d.** Other: ______

c) Conceptual scaffold:

$\mathbf{\frac{2}{3}}$ - $\mathbf{\frac{1}{9}}$ =




d) Procedural scaffold:

Hint: Convert $\frac{2}{3}$ and $\frac{1}{9}$ so they both have a denominator of 9.

$\mathbf{\frac{2}{3} - \frac{1}{9}} = \frac{\square}{\mathbf{9}} + \frac{\square}{\mathbf{9}} =$

**a.** $\frac{1}{6}$ **b.** $\frac{5}{9}$ **c.** $\frac{1}{9}$ **d.** Other: ______

FIGURE 1 Example of each scaffold type used on the pretest and posttest for the problem $\frac{2}{3} - \frac{1}{9}$.

schools (one from each school). Students' accuracy on the assessment did not differ across the four suburban classrooms, suggesting that the written work available was representative of student problem solving.

*Phase 2: Instructional intervention.* All students participated in the same instructional intervention, which was designed to build on their prior knowledge of fractions identified in Phase 1. During the intervention, students solved fraction addition and subtraction problems on a computer-based intelligent tutoring system. The conceptual and contextual scaffolds were present for all problems. Simi-

TABLE 2

Strategy Use: Percentage of Problems on Which Students Used Each Strategy, by Scaffold Type and Assessment Time

| | | *Pretest: Scaffold Type*[a] | | | | *Posttest: Scaffold Type*[b] | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Strategy* | *Written Work* | *None* | *Concept* | *Context* | *Procedure* | *None* | *Concept* | *Context* | *Procedure* |
| None | No work shown | 78 | 72 | 81 | 22 | 65 | 66 | 59 | 22 |
| Common denominator | Wrote a common denominator and equivalent fractions | 15 | 16 | 15 | 35 | 35 | 30 | 38 | 60 |
| Error in equivalent fractions | Wrote a common denominator but not equivalent fractions | 2 | 2 | 0 | 36 | 0 | 2 | 1 | 17 |
| Visual estimation | Drew picture or marked given fraction bars | 0 | 2 | 5 | 0 | 0 | 1 | 1 | 0 |
| Potential numeric estimation[c] | Only written work was a fraction not used as answer | 3 | 7 | 0 | 2 | 0 | 1 | 0 | 0 |
| Other | | 2 | 2 | 0 | 5 | 0 | 1 | 2 | 1 |

[a]$N = 179$. [b]$N = 168$. [c]Only 5 out of the 22 instances were within one tenth of the correct answer, suggesting most were not numeric estimates.

lar to the conceptual scaffold on the DFA, fraction bars displayed the value of each given fraction as well as the value of the students' answers. Similar to the contextual scaffold on the DFA, all problems were presented in the context of two students combining or finding the remaining amount of a candy bar (e.g., "Jared had ⅔ of a candy bar and Latesha ate 2⁄4 of the candy bar. How much of the candy bar remained?"). To increase the relevance of the context to the students, real names of sixth graders at each school were randomly chosen to be included in each of the story problems (Cordova & Lepper, 1996). Anecdotal evidence suggested that students often reasoned about the fraction bars as if they represented candy bars and were motivated by having classmates' names included in the problems.

The presence of the procedural scaffold (i.e., providing a common denominator) varied across three blocks of problems. In the first block of 12 problems, the fractions had the same denominators, so the procedural scaffold was not needed. On the second block of 24 problems, the fractions usually had unlike denominators, and the procedural scaffold was present—students were given a common denominator. A "conversion scratch pad" was provided for students to find equivalent fractions with a common denominator. Figure 2 contains a screen shot of a completed problem from the second block of problems and illustrates the implementation of the conceptual, contextual, and procedural scaffold. On the third

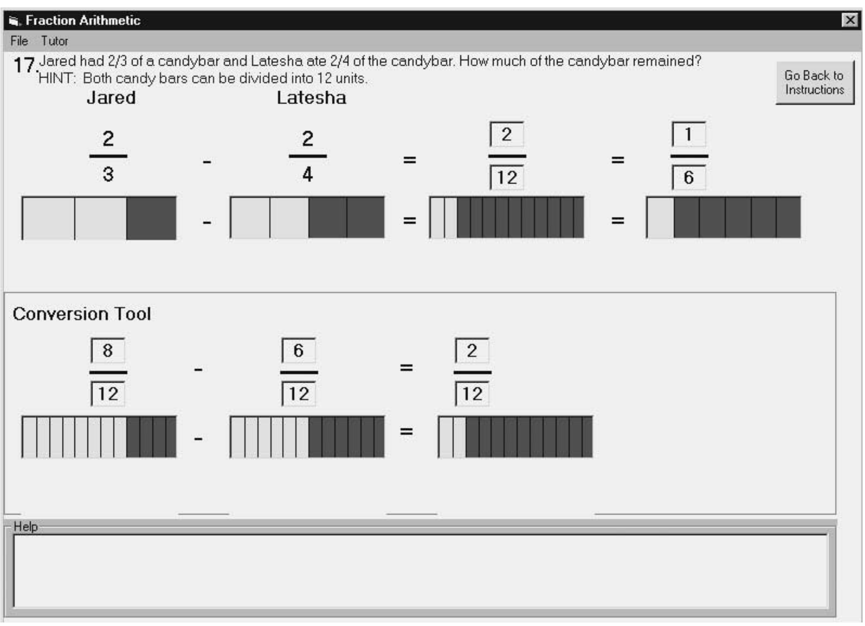



FIGURE 2 Screen shot of a completed problem from Block 2 on the intelligent tutoring system used during the intervention.

TABLE 3
Characteristics of Blocks of Problems Presented During the Intervention

| *Block* | *Fraction Type* | *Number of Problems* | *Scaffolds Present* |
|---|---|---|---|
| 1 | Same denominator (e.g., $\frac{3}{9} + \frac{2}{9}$) | 12 | Concept and context |
| 2 | Unlike denominators (e.g., $\frac{1}{4} + \frac{2}{3}$) | 24 | Concept, context, and procedure |
| 3 | Unlike denominators (e.g., $\frac{3}{4} + \frac{3}{7}$) | 25 | Concept and context |

block of 25 problems, the procedural scaffold was removed (fading of scaffolding). However, the conversion scratch pad was still available. Table 3 outlines the three blocks of problems used during the intervention including the presence of each scaffold.

When students made errors, a feedback message automatically appeared on the screen in the help window. The feedback message noted what the error was (if it was a common one) or suggested a first step for solving the problem. The feedback message combined the three scaffold types. For example, if students entered $\frac{2}{9}$ as the answer to $\frac{1}{4} + \frac{1}{5}$, a feedback message appeared saying, "You cannot just add the top and bottom numbers together. Look at the pictures of the candy bars. Together, Brittany and Mike have more than 2-out-of-9 pieces of a candy bar." The intervention program was written in Microsoft Visual Basic (Version 5.0) and was presented on computers in the schools' computer laboratories.

Students received a brief introduction to the tasks and computer interface by their classroom teacher or Bethany Rittle-Johnson or another member of the project team and then worked individually at their own pace for two to three class periods on the intervention (approximately 1½ hr spent working on the intervention problems). Before beginning the second block of problems, students studied a worked example of adding two fractions with unlike denominators including fraction bars and a candy bar context to justify each step (presented on the computer). The normal classroom teacher and sometimes a member of the project team circulated among the students and helped the students who were having difficulty. The availability of the teacher to help individual students who are having difficulty while other students remain engaged in the activity is one critical feature of Cognitive Tutors.

*Phase 3: Posttest.* The posttest was identical to the pretest. Five of the items on the assessment were similar to those presented during the intervention (i.e., adding fractions with the same denominator and the four DFA items on adding or subtracting fractions with unlike denominators). The other three items had not been seen during the intervention (i.e., adding three fractions, subtracting mixed numbers, and identifying a verbal description of the conventional procedure) and thus served as an index of transfer.

Students completed the posttest when they finished the intervention or at the end of the 3rd day (even if they had not completed the intervention). They were randomly given one of the four versions of the assessment (between 43–60 students completed each version). Unfortunately, there were not opportunities for students who missed class to make up missed work on the computer, and thus, some students solved substantially fewer intervention problems. Nevertheless, these students were included in the sample to evaluate the effectiveness of the intervention under typical classroom conditions.

## RESULTS

We report the results in three sections. We begin with the results from Phase 1 on students' prior knowledge. This section includes findings from the DFA, the supporting error and strategy analyses, and the knowledge-component modeling of the findings. Next, we report the posttest results from Phase 3 to give an overview of learning from the intervention. Finally, we report descriptive data from Phase 2 to provide a qualitative sense of learning during the intervention.



### Phase 1: Identifying Prior Knowledge

Our sixth-grade students had some prior knowledge of fraction addition and subtraction. Average accuracy across the eight pretest items was 45%. Considering the individual items, over 80% of students correctly added two fractions with the same denominator. Adding and subtracting fractions with unlike denominators was much more difficult; accuracy was 40% averaged across the four items of this type. Similarly, accuracy ranged from 37% to 42% correct on the other three items (i.e., adding three fractions, subtracting mixed numbers, identifying a verbal description). Students at the suburban schools had higher accuracy scores than students at the urban schools at pretest (*M*s = 62% vs. 35% correct), $F(1, 221) = 51.9$, $p < .0001$. Accuracy scores did not differ for male students and female students (both 45% correct; there were no gender differences in any analyses). Our primary interest in Phase 1 was identifying whether and how the conceptual, contextual, and procedural scaffolds aided problem solving.

*Accuracy on the DFA items at pretest.* Students solved four DFA problems on the pretest, all involving adding or subtracting fractions with unlike denominators. Each problem had one of four scaffold types; to avoid confounding scaffold type with particular calculations, four different assessment forms were used (see Table 1). As expected, overall performance on the four forms was similar ($p = .26$).

As shown in Table 4, students had the least success on the no-scaffold problem and the most success on the procedural scaffold problem, and the effects of the scaffolds were similar across the two school locations (even though students at the suburban schools outperformed the students at the urban schools (*M*s = 52% vs.

TABLE 4
DFA Performance at Pretest: Proportion of Children Making Each Response Type Across the Four Scaffold Types Overall and by School Setting

| | *Scaffold Type* | | | |
|---|---|---|---|---|
| *Response Type* | *None* | *Conceptual* | *Contextual* | *Procedural* |
| Overall[a] | | | | |
| Correct | .30 | .42 | .41 | .49 |
| Combine both | .38 | .28 | .30 | .12 |
| Fail to convert | .21 | .12 | .19 | .27 |
| Urban school[b] | | | | |
| Correct | .20 | .37 | .32 | .44 |
| Combine both | .50 | .36 | .39 | .15 |
| Fail to convert | .18 | .08 | .18 | .28 |
| Suburban schools[c] | | | | |
| Correct | .44 | .50 | .55 | .57 |
| Combine both | .19 | .14 | .15 | .06 |
| Fail to convert | .27 | .19 | .20 | .27 |

*Note.* DFA = difficulty factors assessment.
[a]$N = 233$. [b]$N = 137$. [c]$N = 86$.

33% correct), $F(1, 221) = 13.93$, $p = .0002$. A student's response for each scaffold type was categorical, so pairwise sign tests were used to compare performance on problems with different scaffolds collapsed across the two schools. Children were more likely to choose the correct solution on each of the scaffolded problems compared to the no-scaffold problem ($p$s < .001). Their success on the procedural-scaffold problem was marginally better than on the contextual- or conceptual-scaffold problems ($p = .07$ and .08, respectively).

*Errors on the DFA items.* To understand how the scaffolds improved accuracy, we examined the kinds of errors students made on the four DFA problems. The scaffolds reduced problem difficulty largely because they reduced the common error of adding or subtracting both the numerators and denominators (combine-both error; see Table 4). Students were less likely to make a combine-both error on each of the scaffolded problems compared to the no-scaffold problem (all $p$s = .05). Although the procedural scaffold did not eliminate this error (as might be predicted), it did reduce this error more than the other two scaffolds ($p$s < .0001).

The conceptual scaffold also reduced the error of using a common denominator but not converting the numerator (fail-to-convert error) compared to the no-scaffold problem ($p = .004$). The contextual and procedural scaffolds did not reliably impact this error type ($p = .58$ and .10, respectively). Only the conceptual scaffold provided direct support for noticing that the fail-to-convert error leads to an impossible answer.

The scaffolds did not impact whether students attempted to solve the problems, largely due to a ceiling effect. Students only left the answer blank or responded "don't know" on 5% of trials, with little difference between problem types.

*Strategy use on the DFA items.* Students' written work provided some additional insight into how each scaffold impacted problem solving. As shown in Table 2, the procedural scaffold greatly increased how often students showed their work; one third of students showed evidence of using a correct common-denominator procedure, and another third of the students used a common denominator but had difficulty finding equivalent fractions on this problem. On the other problem types, the most frequent work shown was the correct common-denominator procedure. Students showed very limited evidence of using alternative strategies, such as visual or numeric estimation, on any of the problems. Finally, there was no evidence of students using multiplication or division on any of the problems, and none of the incorrect answers were a result of multiplying or dividing the given fractions. Overall, most written work indicated correct or incorrect use of the conventional common-denominator procedure.

In summary, the conceptual, contextual, and procedural scaffolds each helped students to add and subtract fractions correctly. Each of the scaffolds reduced combine-both errors, but only the conceptual scaffold consistently reduced fail-to-convert errors. Students' written work suggests that correct answers were found using the conventional common-denominator procedure across the four problem types.

*Knowledge-component modeling of pretest DFA results.* To better understand how the scaffolds impacted problem solving, we created and compared three potential cognitive models of pretest performance. First, we generated models based on hypothesized knowledge components needed to solve the problems, and then we fit the models to the accuracy and error data reported previously. Inspecting where the models converged and deviated from the observed data provided insights into the strengths and weaknesses of the proposed models.

All models consisted of independent variables capturing the difficulty of hypothesized knowledge components and16 equations using the independent variables to predict students' responses (correct, combine-both error, fail-to-convert error, and other error) on the four problem types (conceptual, contextual, procedural, or no scaffold). Although quantitative in nature, the models are meant primarily as a tool for thinking qualitatively about experimental data and comparing alternative theoretical models for explaining the data (Aleven & Koedinger, 2002).

First, we generated four base equations for predicting the frequency of the four response types on the unscaffolded problem. Students' written work suggested that they primarily used the conventional procedure of finding a common denominator to add or subtract fractions. This procedure involved three core knowledge components—(a) find a common denominator (*cd*), (b) find equivalent fractions (*ef*), and

(c) combine (add or subtract) numerators when the denominators are the same (*cn*). For our purposes, these knowledge components represented both the selection and the implementation of the knowledge. The probability of finding the correct answer was the product of the probability of using each of these knowledge components (*cd* × *ef* × *cn;* additional descriptions of the equations are in Appendix A). The probability of making a combine-both error could simply be the default approach if students did not find a common denominator (and our modeling work supported this inference; *not cd*). The fail-to-convert error reflected the probability that students found a common denominator, did not find equivalent fractions, but still combined numerators (*cd* × *not ef* × *cn*). All remaining errors were in the "other" category.

Next, we adapted these base equations to capture the effect of each scaffold. For the procedural scaffold, we added a knowledge component, use given common denominator (*u*), which replaced find common denominator in the four equations for this scaffold (see Appendix A). For example, to add $\frac{1}{4} + \frac{1}{5}$ when given the common denominator 20, students did not need to find a common denominator. However, they did need to notice and use the given common denominator. Thus, the probability of finding a correct answer on this problem would be the product of the probability of *u* (e.g., 20), *ef* (e.g., $\frac{5}{20}$ and $\frac{4}{20}$), and *cn* (e.g., to get $\frac{9}{20}$; *u* × *ef* × *cn*).

For the conceptual and contextual scaffolds, we explored three potential mechanisms underlying their effects on performance. All three mechanisms were based on prior research and theory that has suggested that conceptual and contextual knowledge support magnitude based, part–whole representations of fractions (Cramer et al., 2002; Leinhardt, 1988; Mack, 1990, 1993). The three mechanisms were (a) alternative common denominator (eliciting a magnitude-based alternative to finding a common denominator), (b) estimation (eliciting an alternative, estimation-based solution strategy), and (c) reject implausible answers (prompting students to reject implausible solutions). Modeling each of these mechanisms required adding one new knowledge component to the base model, and the equations used in each model are listed in Appendixes A, B, and C. To model this data, we generated separate models for each mechanism and within each model used the same equations for the conceptual- and contextual-scaffold problems. This approach was a practical rather than theoretical decision. There was simply not enough variability in observed responses between the different scaffolded problems to include multiple mechanisms in the same model or to model the effects of conceptual and contextual scaffolds separately. Future research using multiple instances of each problem type should yield data that makes a more fine-grained approach feasible.

First, consider the alternative-common-denominator mechanism. The conceptual or contextual scaffold may have elicited a magnitude-based representation of fractions that allowed students to reason through the need for equal-size pieces (i.e., a common denominator) and use their number sense to find a common denominator. For example, to add $\frac{1}{4} + \frac{1}{5}$ in the candy bar context or with fraction bars, a student could either find a common denominator symbolically (e.g., by multiplying $4 \times 5$) or she or he could use the context or pictures to reason about the problem, recognizing that quarters and fifths cannot be combined and that both can be further divided into 20ths. To model this alternative approach, the find-common-denominator knowledge component was supplemented with a knowledge component for finding the common denominator by reasoning about magnitudes in each base equation (see Appendix A).

Second, consider the estimation mechanism, which provides an alternative to the common-denominator procedure. Students could estimate the answers visually (e.g., drawing or using the given fraction bars) or use benchmarks such as $\frac{1}{4}$, $\frac{1}{2}$, and 1. Because the assessment was multiple choice, estimation was a sufficient procedure for selecting a correct answer. For example, to add $\frac{1}{4} + \frac{1}{5}$, a child might estimate that $\frac{1}{5}$ is close to one fourth, know that combining two fourths yields one half, and thus find the answer closest to $\frac{1}{2}$. Alternatively, a child might look at or draw fraction bars for $\frac{1}{4}$ and $\frac{1}{5}$ and mentally combine the amounts, noting that the answer is close to $\frac{1}{2}$. To model this estimation procedure, the four base equations were adjusted to included the probability of finding the correct answer through estimation as an alternative to using the conventional common-denominator procedure (see Appendix B).

Finally, consider the reject-implausible-answers mechanism in which students continue to use incorrect procedures to find answers, but their conceptual or contextual knowledge of fractions prompts them to reject implausible answers. Both the combine-both and fail-to-convert errors lead to answers that are smaller than either addend and thus to answers that violate a principle of addition. For example, after making a combine both error when adding $\frac{1}{4} + \frac{1}{5}$, a student might recognize that $\frac{2}{9}$ is smaller than $\frac{1}{4}$ and thus cannot be the correct answer. This should motivate him or her to solve the problem again (although this knowledge component does not provide information on how to solve it differently). To model this rejection of implausible answers, the base equations for the combine-both error and the fail-to-convert error included a knowledge component representing the probability of rejecting these responses (see Appendix C).

To evaluate the three models (one for each potential mechanism underlying the conceptual/contextual scaffold), we used the Generalized Reduced Gradient method offered by the Microsoft® Excel Solver to find the values of the knowledge components in each model that best predicted the frequency of the four response types on the four problem types at pretest (using maximum likelihood estimates). This yielded 16 predicted response frequencies for each problem type as well as probability estimates of students using each knowledge component. Comparing the predicted and observed response frequencies offered insights into how each potential mechanism would impact problem solving. In Figure 3, for each model, the predicted frequency of each response is displayed as a line graph superimposed over the observed frequency of each response displayed as a bar graph. The ability

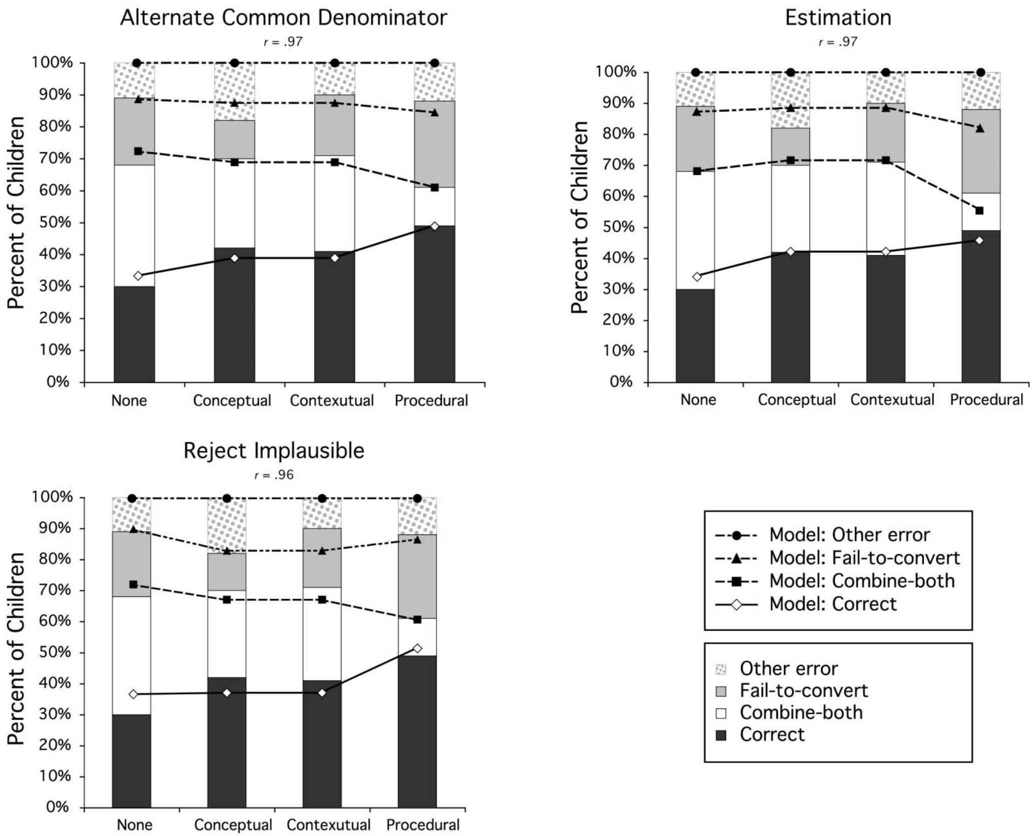



FIGURE 3 Predicted frequency of response choices produced by each knowledge-component model at pretest. The predicted frequencies are displayed as line graphs superimposed over the observed frequency of each response at pretest (displayed as bar graphs).

of each model to capture the seven key empirical findings of the DFA at pretest is outlined in Table 5.

Both the alternate-common-denominator and the estimation models captured each key empirical finding except the decrease in fail-to-convert errors for the conceptual scaffold. In contrast, the reject-implausible-answers model predicted the decreases in fail-to-convert errors but did not predict the observed increases in correct answers. This suggests that the reject-implausible-answers mechanism may be used in conjunction with one of the other two mechanisms. Overall, all three models predicted the observed frequency of responses quite well, supporting the plausibility of our cognitive models and the potential role of each mechanism.

The models also provided estimates of the probability of successfully implementing each knowledge component as shown in Table 6. Comparing the probability estimates provided an estimate of the relative difficulty and use of each knowledge component. Finding a common denominator and finding equivalent fractions were equally difficult; the probability of doing either correctly was in the mid .60s across the three models. This suggested that we should consider scaffolding finding equivalent fractions as well. By comparison, combining numerators once the denominators were the same and using a given common denominator were easier skills (probabilities in the .80s). Finally, the alternate-common-denominator, estimation, and reject-implausible-knowledge components were used infrequently (probabilities of using each correctly were below .25 across the three models).

Overall, the conceptual- and contextual-knowledge scaffolds seemed to provide backup approaches for solving the problems such as finding a common de-



TABLE 5
Summary of Each Knowledge-Component Models' Ability to Capture the Key Findings of the Pretest Difficulty Factors Assessment

| *DFA Finding* | *Alternate Common Denominator* | *Estimation* | *Reject Implausible Answers* |
|---|---|---|---|
| Frequency of correct responses | | | |
| Procedure > none | √ | √ | √ |
| Concept and context > none | √ | √ | x |
| Procedure > concept and context | √ | √ | √ |
| Frequency of combine-both-errors | | | |
| Procedure < none | √ | √ | √ |
| Concept and context < none | √ | √ | √ |
| Procedure < concept and context | √ | √ | √ |
| Frequency of fail-to-convert errors | | | |
| Concept < none | | | √[a] |

*Note.* DFA = difficulty factors assessment. "Procedure > none" for correct responses indicates that correct responses were more likely on the procedural scaffolded item than on the no-scaffold item.

[a]Incorrectly predicts reduction in fail-to-convert errors on context problems.

TABLE 6
Probability Estimates for Each Knowledge Component in the Knowledge-Component Models at Pretest and Posttest

| Knowledge Component | Pretest | | | Posttest | | |
|---|---|---|---|---|---|---|
| | *Alternate Common Denominator* | *Estimate* | *Reject Implausible Answers* | *Alternate Common Denominator* | *Estimate* | *Reject Implausible Answers* |
| Find common denominator | .62 | .65 | .66 | .82 | .84 | .82 |
| Find equivalent fractions | .67 | .64 | .66 | .79 | .79 | .78 |
| Combine numerators | .82 | .81 | .87 | .90 | .90 | .93 |
| Use given common denominator | .88 | .88 | .88 | .89 | .89 | .89 |
| Magnitude-based component[a] | .24 | .12 | .10 | .16 | .00 | .15 |

[a]Varied by model.

nominator through reasoning about fraction quantities, estimating the solution, or rejecting implausible answers. Scaffolding procedural knowledge seemed to replace a relatively difficult skill of finding a common denominator with an easier skill of using a common denominator.

*Summary and implications.* The sixth-grade students had a fair amount of prior knowledge of fractions. For example, most could add fractions with the same denominator. In contrast, students had great difficulty adding or subtracting fractions with unlike denominators, and the most frequent error was to combine the numerators and denominators. Providing a procedural scaffold improved accuracy, suggesting that some students had procedural knowledge for finding equivalent fractions and adding like fractions but that this knowledge was masked by not using a common denominator. The conceptual and contextual knowledge scaffolds improved accuracy because they seemed to elicit prior knowledge that allowed them to implement backup approaches for solving the problems. Instruction using these three scaffolds should improve students' ability to add and subtract fractions.

## Phase 3: Posttest Results

In this section, we explore the impact of the intervention, which incorporated all three scaffolds, on learning from pretest to posttest. Of the total number of students, 33 students did not complete the posttest because they were absent from the classroom on the 3rd day. These students scored lower on the pretest than did students who completed the posttest ($M = 32\%$ vs. 48% correct), $F(1, 221) = 6.96$, $p < .009$. Such attrition is common in classroom research, and it is not surprising that students who tended to miss school more frequently had lower initial knowledge. The following results may slightly overestimate learning from our intervention in this population, but the existing sample is still more representative of the general population than is typical in educational research.

*Changes in accuracy at posttest.* Students solved more problems correctly at posttest compared to pretest ($M = 45\%$ to 61% correct), $t(189) = 7.50$, $p < .0001$. Amount of gain was nearly identical for students in the urban and suburban locations and for male and female students.

The assessment included both learning and transfer problems. Five problems were structurally similar to those presented during the intervention (adding and subtracting two fractions with like and unlike denominators), and students improved in performance on these learning problems ($M$s = 51% to 66% correct), $t(189) = 7.19$. Three of the problems were novel (e.g., subtracting mixed numbers); students also improved in accuracy on these transfer problems ($M$s = 43% to 53% correct), $t(189) = 4.52$. Indeed, students' accuracy improved significantly on seven out of the eight individual items as shown in Table 7.

To assess individual change from pretest to posttest, we calculated a percent gain score for each student: (posttest percentage correct – pretest percentage correct) / (100 – pretest percentage correct). This yielded a score from 0 to 100, reflecting the percent gain for each child. Of the students, 7% ($n = 16$) were at ceiling at pretest and thus were omitted from this analysis. Students' average gain from



TABLE 7
Pretest-to-Posttest Changes: Proportion of Children Making Each Response Type on Individual Assessment Problems

| | | *Correct Answer* | | *Combine-Both Error* | | *Fail-to-Convert Error* | |
|---|---|---|---|---|---|---|---|
| *Problem* | *Type* | *Pretest* | *Posttest* | *Pretest* | *Posttest* | *Pretest* | *Posttest* |
| Unlike denominators | | | | | | | |
| No scaffold | Learn | .32 | .52* | .33 | .17* | .23 | .18 |
| Conceptual | Learn | .43 | .58* | .26 | .16* | .12 | .17 |
| Contextual | Learn | .44 | .57* | .27 | .18* | .19 | .16 |
| Procedural | Learn | .49 | .66* | .10 | .04* | .28 | .17* |
| Same denominator | Learn | .85 | .95* | .13 | .04* | na | na |
| Adding 3 | Transfer | .42 | .55* | .26 | .16* | .18 | .16 |
| Subtracting mixed | Transfer | .44 | .58* | .25 | .22 | na | na |
| Verbal description | Transfer | .42 | .47 | .16 | .05* | na | na |

*Note.* $N = 190$. na = not applicable. Students who did not complete the posttest are not included in the pretest data.

*$p < .05$, change from pretest to posttest based on pairwise comparisons.

pretest to posttest was 37%, and we categorized students as making no gain, modest gain (less than 50% gain), or high gain (at least 50%). Approximately a third of the sample (35%) made high gains, another third (28%) made modest gains, and the remaining third (38%) made no gain.

A critical factor in predicting percent gain was how many intervention problems a student solved. The two were correlated, $r(172) = .42$, $p < .0001$, and students in the three categories differed in the number of intervention problems solved, $F(2, 171) = 18.96$, $p < .0001$. In particular, students in the no gain and modest gain groups were much less likely to have started the final section of the intervention on which the procedural scaffold was no longer given than students in the high gain group (66% and 63% vs. 95% of students started the final section, respectively; $\chi^2(2, N = 173) = 19.252$, $p < .0001$.

Although there was no effect of school location on overall amount of gain, the distribution of students in the three gain categories varied by school location, $\chi^2(2, N = 173) = 16.82$, $p = .0002$. Students at the suburban schools were twice as likely to be in the high-gain group as students at the urban school (51% vs. 24% of students), whereas students in the urban schools were more likely to be in the modest gain group (13% vs. 37% of students). Again, differences in percent of gain were linked to number of intervention problems solved, with students at the suburban school completing more intervention problems in the allotted time (*M*s = 93% vs. 72% of intervention problems), $F(1, 172) = 30.14$, $p < .0001$. Students in the suburban schools were more likely to be in class and to be engaged in instructional activities. In summary, many students seemed to learn from the intervention, but some students did not, especially if they did not have sufficient time to complete the no-procedural scaffold section of the intervention.

*Changes in errors at posttest.* Students' improvement at posttest was partially due to a reduction in the number of combine-both errors across the problems from pretest to posttest (22% to 13% of problems), $t(189) = 5.93$, $p < .0001$. A reduction in combine-both errors occurred on every problem but one (see Table 7). There was not a decrease in fail-to-convert errors across the problems ($p = .10$), but fail-to-convert errors decreased for the procedural scaffold problem (*M*s = 28% to 17% of problems), $p = .009$.

*Effects of scaffolds (DFA) at posttest.* Given students' improved knowledge of strategies for adding and subtracting fractions, what was the relative impact of the different scaffolds at posttest? As shown in column 4 of Table 7, the contextual and conceptual scaffolds had less impact on accuracy at posttest, and accuracy on the problems was no longer reliably different from accuracy on the no-scaffold problem ($p = .22$ and .11, respectively). In contrast, the procedural scaffold continued to improve accuracy at posttest compared to the no-scaffold problem ($p = .0004$). The procedural scaffold also led to better performance than the contextual or conceptual scaffolds ($p$s < .05). The superior performance on the procedure-scaffold item was partially due to a reduction in combine-both errors compared to each of the other problems ($p$s < .0001; see column 6 of Table 7). None of the scaffolds reduced fail-to-convert errors at posttest.

Inspection of students' written work further supports these findings (see Table 2). Compared to pretest, students were much more likely to show written evidence of using a correct common-denominator procedure and somewhat less likely to show use of other strategies at posttest. The largest change was on the procedural scaffold item; over 60% of students showed evidence of using the correct common-denominator procedure, and students were much less likely to make errors in finding equivalent fractions.

*Knowledge-component modeling at posttest.* The same three models used to capture performance on the DFA items at pretest were refit to students' performance on the DFA items at posttest. Our key interest was in changes in the probability estimate for each knowledge component from pretest to posttest (all three models had an extremely good fit to the posttest data, $r$s = .99). As shown in Table 6, the probability of correctly finding a common denominator and correctly finding equivalent fractions both improved substantially from pretest to posttest, regardless of which model is considered. There was also a smaller improvement in combining numerators correctly. Interestingly, changes in the probability estimates for the magnitude-based knowledge component, which varied by model, were not consistent across the three models. The probability of using the alternate-common-denominator approach or using estimation both decreased (and such a decrease is also supported by their written work; see Table 2). In contrast, the probability of rejecting implausible answers increased. Although such differences must be interpreted with caution, they suggest that instruction and experience may impact each mechanism differently.

*Summary.* The posttest results suggest that the intervention, which included scaffolds for all three types of prior knowledge, facilitated learning to add and subtract fractions. At posttest compared to pretest, children were more accurate across a range of problems, made many fewer common errors such as adding the numerator and denominator, had less need for the scaffolds, and seemed more likely to correctly use the conventional procedure. The knowledge-component models suggested that children had improved in their ability to carry out each component of the conventional procedure. Changes in their use of alternative, backup approaches seemed to vary based on the proposed mechanism.

### Learning During the Intervention

Descriptive data from the intervention sessions provided clues for how the intervention supported learning and why some students did not learn much. Students' answers during the intervention were recorded through the computer software. Due to technical difficulties, data at the urban school were not reliably recorded during the intervention. Thus, data in this section are from the two suburban schools and may not be representative of learning pathways at the urban school.

The answers provided by two students illustrate a successful and an unsuccessful learning pathway. First, consider Brittany, who figured out the common-denominator procedure during the first half of the intervention and quickly generalized the procedure to problems without the procedural scaffold. At pretest, Brittany solved two problems correctly and made fail-to-convert errors on the remaining problems. On the first block of intervention problems (in which fractions had the same denominator), she solved all 12 problems correctly on the first try. However, when she started the second block of problems (in which fractions had unlike denominators), she had great difficulty finding equivalent fractions. On the 1st problem, she initially made a fail-to-convert error and then a combine-both error. Next, she started using the conversion scratch pad (because she was prompted to do so) but seemed to randomly guess what the new numerator should be (e.g., converting $1/4$ to 12ths, she tried: $4/12$, $5/12$, $6/12$, $1/12$, $2/12$). She persisted with random guessing for the new, converted numerators for 5 more problems. On the 7th problem, she made a single error finding equivalent fractions but solved the problem correctly on the second attempt. She solved 14 out of the remaining 18 problems in this block correctly on the first attempt (76%).

When a common denominator was no longer given, Brittany continued to solve a majority of the problems correctly (84%). She always found a common denominator and equivalent fractions and only made one error when doing so. However, she did not find the least common denominator but rather multiplied the two denominators together, even when one denominator was a multiple of the other (e.g., used 20 sevenths for $8/9 + 1/3$). This led to arithmetic errors when she needed to reduce the fraction for the final answer. Thus, she had learned a correct but cumbersome procedure for adding and subtracting fractions. Brittany maintained her use of a correct procedure on the posttest, solving all of the learning problems and two out of the three transfer problems correctly.

Contrast this with Katie, who did not learn from the intervention. Katie also solved 2 problems correctly at pretest and typically made fail-to-convert errors on the other problems. On the first block of problems during intervention, she solved 75% of problems correctly. During the second block of problems (in which fractions no longer had the same denominator), she only solved 38% of the problems correctly. Her initial attempts typically included a fail-to-convert error, followed by seemingly random guessing (e.g., for $1/3 + 1/6$: answers $2/6$, then $2/3$, then $2/12$, and then attempts to use the conversion tool by converting $1/6$ to $2/6$, $3/6$, and $6/6$). She solved problems correctly when the denominators were “friendly” (e.g., $1/2 - 1/4$; $1/5 + 1/10$) and toward the end of Block 2 solved some harder problems correctly (e.g., $2/10 - 1/15$), but she continued to have difficulty converting to equivalent fractions on many problems. She began the third block of problems but only had time to solve

10 of the 25 problems. On these problems, she always identified a common denominator but continued to make errors in finding equivalent fractions. Katie showed no improvement on the posttest.

Brittany's pattern of figuring out a correct procedure when a common denominator was given and then quickly adapting this procedure to find a common denominator on her own (as indexed by performance above the median in Block 2 to Block 3) was evident in 43% of students. Of students in this high performance group during the intervention, 66% were in the high pretest-to-posttest gain category. Katie's pattern of struggling throughout the problems with unlike denominators (as indexed by performance below the median on both Blocks 2 and 3) was evident in 30% of students. Of the students in this low-performance group during the intervention, 50% showed no gain from pretest to posttest.

Brittany and Katie also illustrate a common pattern across the students: Finding equivalent fractions was very challenging for students whether or not a common denominator was given. This arose even though students had recently completed a unit on fraction concepts (and were 80% correct on equivalent fraction problems on the unit posttest). Students who learned from the intervention improved their ability to find equivalent fractions within the context of adding and subtracting fractions, whereas for other students, the intervention was insufficient for supporting this learning.

## DISCUSSION

In this study, we illustrated the use of a multimethod, multiphase approach to designing learning environments. Using this approach, we identified promising knowledge scaffolds, explored potential mechanisms underlying their effectiveness, and designed and evaluated a learning environment incorporating these scaffolds. In the discussion, we first consider the methodology used in this study, then discuss implications for potential mechanisms underlying each knowledge scaffold, and finally discuss implications for designing better learning environments.

### Benefits of Integrated Methodology

In Phase 1, DFA was combined with error and strategy use analyses and knowledge-component modeling to identify what types of scaffolds facilitated problem solving and to provide clues to the strategies and mechanisms underlying these effects. In Phase 2, we designed a learning environment based on the findings of Phase 1 and implemented the intervention with a wide range of students as part of their regular mathematics curriculum. In Phase 3, we used the same methodology as in Phase 1 to evaluate the effectiveness of the intervention and to inform the iterative redesign of the intervention.

DFA is an efficient and effective method for gaining unconfounded, causal evidence for what problem factors make problems easier or harder to solve. DFA involves systematically varying target factors and creating multiple versions of the assessment so that these factors are not confounded with other problem features (Koedinger & Nathan, 2004). Comparisons of problem-solving accuracy provide direct evidence for the effects of the target factors on problem-solving performance (e.g., that a story context made fraction problems easier to solve). Thus, DFA provides direct evidence of students' prior knowledge.

DFA can also be used in pretest–posttest designs to identify changes in which factors facilitate or hamper problem solving after instruction or experience. For example, in this study, the conceptual and contextual scaffolds facilitated problem solving at pretest but not at posttest. This suggested that these knowledge scaffolds had been integrated with children's problem-solving knowledge by posttest and no longer needed to be elicited directly. In contrast, the procedural scaffold continued to improve performance at posttest, suggesting that students needed additional experience to fully integrate knowledge of finding a common denominator with their problem-solving strategy.

In addition to evaluating students' accuracy on a DFA, analyses of errors and strategy use provide insights into potential mechanisms underlying the effects of different problem factors. Consider two examples from this study. First, the procedural scaffold did not eliminate the combine-both error, indicating that some students' difficulty went beyond failure to know how to find a common denominator and included not knowing how to use a common denominator and find equivalent fractions. Second, students typically showed evidence of using the conventional common-denominator procedure rather than alternative strategies such as estimation across the problem types. At the same time, only the conceptual scaffold reliably reduced both combine-both and fail-to-convert errors, suggesting that some alternative strategy must have been supported by this scaffold.

The use of knowledge-component modeling allowed us to explore how alternative mechanisms could account for these outcomes. Perhaps the most useful and least transparent benefit of cognitive modeling is helping researchers clarify and specify their thinking. The mechanisms explored in this study had been proposed by others and were supported by our data, but the process of modeling these mechanisms forced us to hypothesize how different mechanisms might be implemented (e.g., an alternate common-denominator process likely does not replace finding a common denominator symbolically but rather serves as a backup approach) and predict error patterns (e.g., a reject-implausible-answers mechanism reduces the frequency of common, implausible errors without improving accuracy). Such insights guide future iterations of the design process and contribute to domain theories of how children learn specific content knowledge. A second benefit of knowledge-component modeling is the ability to compare probability estimates at a given time and over time. These comparisons provide clues to which knowledge components are most difficult or most commonly used at a given time, which knowledge components are strengthened by an intervention, and which knowledge components still need additional instructional focus. Overall, knowledge-component modeling provides a research tool that is less time consuming and more accessible to the average researcher than implementing a running computer model while retaining the need for greater precision and evaluation than a paper-and-pencil task analysis or verbal model.

This set of methods is an ideal addition to design experiments. Design experiments have emerged as a predominant methodology for studying and improving student learning within the complexity of functioning classrooms (Barab & Squire, 2004). Nevertheless, it is often difficult to disentangle necessary, sufficient, and peripheral features of the design. Integrating DFA, analyses of strategies and errors, and knowledge-component modeling with design experiments offers a powerful methodology for evaluating what instructional features facilitate or harm problem solving and for exploring how the features impact learning and performance.

## Mechanisms Underlying the Impact of Each Scaffold

This methodology allowed us to specify and compare potential mechanisms underlying the effects of our conceptual, contextual, and procedural scaffolds. First, consider why the conceptual and contextual knowledge scaffolds may have facilitated accurate problem solving. This research provides direct evidence to support the hypothesis put forth in the National Research Council's synthesis of the mathematics education literature:

> More specifically, say these researchers, instruction should build on students' *intuitive understanding* [italics added] of fractions and *use objects or contexts* [italics added] that help students make sense of the operations. The rationale for that approach is that students need to understand the key ideas in order to have something to connect with procedural rules. For example, students need to understand why the sum of two fractions can be expressed as a single number only when the parts are of the same size. That understanding can lead them to see the need for constructing common denominators. (Kilpatrick et al., 2001, pp. 240–241)

Situating a problem in a real-world context or providing visual models of fraction concepts seemed to do just that. Students were much less likely to make a conceptual error such as adding the numerator and denominator and were more likely to find the correct answer when either scaffold was present. We hypothesize that scaffolding conceptual and contextual knowledge improves students' problem solving via improved problem representation. In particular, both types of knowledge may help students to represent the magnitude of fractions as parts of a whole rather than simply the symbol of one number over another (e.g., Cramer et al., 2002; Hiebert & Wearne, 1996; Hiebert, Wearne, & Taber, 1991; Mack, 1990, 1993; Rittle-Johnson, Siegler, & Alibali, 2001). We explored three potential mechanisms that rely on magnitude-based representations: (a) an alternative, magnitude-based, backup approach to finding a common denominator; (b) an estimation-based solution strategy (Mack, 1990, 1993); and (c) rejection of implausible answers (Hiebert & LeFevre, 1986). The first was the mechanism most compatible with our error and strategy use analyses; students most often showed evidence of using the common-denominator procedure and rarely showed evidence of using estimation, whereas the reject-implausible mechanism did not provide a route to improved accuracy. Rejecting implausible answers may be an additional mechanism elicited by conceptual knowledge, as they were the only mechanism and scaffold, respectively, to reduce fail-to-convert errors. Future research using finer grain assessments and think-aloud protocols should be used to further evaluate the viability of each of these mechanisms both alone and in conjunction with one another and how and when the mechanisms underlying the influences of contextual and conceptual knowledge differ.

Compared to the conceptual and contextual knowledge scaffolds, the procedural knowledge scaffold seemed to influence problem solving through a different mechanism. The procedural scaffold (i.e., providing a common denominator) seemed to facilitate problem solving by replacing the difficult skill of finding a common denominator with the easier skill of using a common denominator. This might avoid overwhelming students' working memory resources and free up resources to focus on identifying relevant information and successfully completing other components of the task.

## Implications for Designing Learning Environments

An understanding of students' prior knowledge and problem-solving processes guided our design of a computer-based intervention incorporating the three knowledge scaffolds. This intervention seemed to support learning and transfer for a majority of students. The sixth-grade students' ability to add or subtract two fractions with unlike denominators (without scaffolding) was below the national average for seventh graders at pretest (32% vs. 53%) but matched this average at posttest (52%; Kouba et al., 1989). Similarly, on the transfer item of adding three fractions, students were below the international average for eighth graders at pretest (42% vs. 49%) and above this average at posttest (55%; Harmon et al., 1997). The absence of a control group in this study prevents us from making causal conclusions on the source of learning, but the findings suggest that scaffolding conceptual, contextual, and procedural knowledge during instruction is a promising tool for improving learning.

Three general design suggestions emerged from integrating these findings with past research: (a) story contexts may be useful scaffolds for introducing new tasks or problem types, (b) visual representations may facilitate problem solving, and (c) scaffolding intermediate procedural steps and then fading the scaffolding may support learning and problem solving. We are not claiming that these principles are always true but rather that they can be helpful when designing instruction. We consider each suggestion in turn.

First, in many domains, carefully chosen context problems can provide a useful way to introduce symbolic problems (cf. Carraher et al., 1987). Contrary to conventional wisdom, adding and subtracting fractions in a story context was easier than the symbolic version of the problem. This adds to prior evidence that students are more successful on context problems than symbolic problems in early algebra (Koedinger & Nathan, 2004) and sometimes in multidigit arithmetic (Baranes et al., 1989; Carraher et al., 1987; Saxe, 1988). The general lesson is that students must comprehend mathematical symbols just as they must comprehend English sentences.

Unless students already have extensive exposure to the target symbol system and effective strategies for working with the symbols (e.g., single-digit addition for middle-class children), we predict that students will do better on story problems and benefit from initial instruction being embedded in familiar story contexts. This prediction is in line with constructivist ideas that people's everyday experiences provide funds of knowledge on which they can build formal, symbolic knowledge (e.g., Greeno et al., 1996; Vygotsky, 1978).

Second, visual representations can support learning and problem solving (e.g., Griffin et al., 1994; Koedinger & Anderson, 1990; Larkin & Simon, 1987; Novick, 2001). In this study, fraction bars were used to illustrate key fraction concepts (e.g., $3/5$ as 3-out-of-5 pieces), and their inclusion led to improved performance at pretest and may have facilitated learning during the intervention. The visual representation was the only scaffold that reliably reduced both types of common errors at pretest—combining numerators and denominators and failing to convert the numerators after finding a common denominator. Both of these errors violate fraction concepts (e.g., adding two fractions cannot lead to a smaller quantity). The visual representation may have allowed students to reason through how to find common denominators based on magnitude and to reject answers that were not conceptually possible.

Other intervention projects on rational numbers and algebra have also used a single, unifying visual representation (e.g., number line) to elicit conceptual knowledge and improve problem solving (Fuson & Kalchman, 2002; Kalchman, Moss, & Case, 2001; Moss & Case, 1999). Nevertheless, the visual representation used in this study may have facilitated learning and performance by more surface-level processes than tapping conceptual knowledge. For example, students may have used the fraction bars to estimate the solutions by visually combining, or subtracting, the two quantities without linking the quantities to fraction concepts. This option is less feasible during the intervention when students needed to enter a specific fraction as the answer before a visual representation of the answer was provided. Overall, we hypothesize that visual representations of key concepts can elicit conceptual knowledge and facilitate learning and problem solving.

Third, procedural scaffolds may support learning and problem solving by allowing for successive approximation of the target skill (Anderson et al., 1995; Collins, Brown, & Newman, 1989; Renkl, Atkinson, Maier, & Staley, 2002). Scaffolding a difficult component of the target skill (e.g., finding common denominators) can allow students to develop and implement other component skills such as converting to equivalent fractions. As students master some components of the target skill, the scaffolding is faded. Such an approach is often used during cognitive apprenticeships as the skilled adult gradually fades help and support (Collins et al., 1989). Thus, we hypothesize that it is useful to include complete models of expert performance (e.g., worked examples or modeling by teachers), scaffolded problem-solving practice (e.g., providing components of the procedure via hints or partial worked examples), and unscaffolded problem-solving practice. The appropriate balance and transition between these phases requires additional research. For example, in future iterations of our intervention, we are considering providing additional modeling of correct problem solving (especially finding equivalent fractions) by using methods such as partial worked examples (Renkl et al., 2002).

This study raises two additional issues for future research. First, what are the benefits and drawbacks to using multiple instances of a particular knowledge scaffold (e.g., candy bar, measurement, and sharing contexts) rather than a single exemplar? Second, how does integrating different types of scaffolds (e.g., conceptual and procedural) influence problem solving? The methodology outlined in this study will allow us to address these issues in future research.

## CONCLUSION

Combining a variety of methods from cognitive science and education provides valuable tools for designing more effective learning environments. Identifying students' prior knowledge and the components of the target task that are more or less difficult for students are important first steps, and DFA, error and strategy analyses, and knowledge-component modeling provide a powerful package of tools for doing just that. In this study, DFA indicated that contrary to popular opinion, story contexts, not symbolic problems, facilitated problem solving at pretest. Providing visual representations of the quantities or one step in the conventional procedure also improved problem solving. Analyses of strategies and errors combined with knowledge-component modeling suggested mechanisms underlying the effects of each scaffold at pretest. These scaffolds were then used when designing a learning environment on fraction addition and subtraction. Finally, repeating the DFA and supporting analyses at posttest provided valuable data for evaluating and improving the initial learning environment and thus provide valuable tools for the iterative redesign process integral to design experiments.

## ACKNOWLEDGMENTS

This work was supported by NIMH/NRSA training Grant 5 T32 MH19983–02 to Carnegie Mellon University and by a grant from Carnegie Learning. Portion of these results were presented at the biennial meeting of the Society for Research in Child Development, April 2001, Minneapolis, MN, and at the 24th annual meeting of the North American Chapters of the International Group for the Psychology of Mathematics Education, October 2002, Athens, GA. We greatly appreciate the help of Karen Cross with the design and computer programming and the help of the administrators, teachers, and students at Greenway Middle School, Chartiers Valley Middle School, and Perrysville Elementary School who participated in the study. We thank Michael Rose and Heather Frantz for their help with coding the data and Kylie Beck for help creating the graphs. Finally, Megan Saylor, Vincent Aleven, and Jon Star provided valuable input on previous drafts of the article.

## REFERENCES

Aleven, V. A. W. M. M., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science, 26,* 147–179.

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences, 4,* 167–207.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences, 13,* 1–14.

Baranes, R., Perry, M., & Stigler, J. W. (1989). Activation of real-world knowledge in the solution of word problems. *Cognition and Instruction, 6,* 287–318.

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle school years: IEA's Third International Mathematics and Science Study.* Chestnut Hill, MA: Boston College.

Bisanz, J., & LeFevre, J. A. (1990). Strategic and nonstrategic processing in the development of mathematical cognition. In D. F. Bjorklund (Ed.), *Children's strategies: Contemporary views of cognitive development* (pp. 213–244). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2001). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Briars, D. J., & Larkin, J. H. (1984). An integrated model of skill in solving elementary word problems. *Cognition and Instruction, 1,* 245–296.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2,* 141–178.
Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42.
Carpenter, T. P., Corbitt, M. K., Kepner, H. S., & Reys, R. E. (1980, September). Solving verbal problems: Results and implications from national assessment. *Arithmetic Teacher, 28*(1), 8–12.
Carraher, T. N., Carraher, D. W., & Schliemann, A. D. (1985). Mathematics in the streets and in schools. *British Journal of Developmental Psychology, 3,* 21–29.
Carraher, T. N., Carraher, D. W., & Schliemann, A. D. (1987). Written and oral mathematics. *Journal for Research in Mathematics Education, 18,* 83–97.
Chiu, M. H., Chou, C. C., & Liu, C. J. (2002). Dynamic processes of conceptual change: Analysis of constructing mental models of chemical equilibrium. *Journal of Research in Science Teaching, 39,* 688–712.
Cobb, P., McClain, K., & Gravemeijer, K. (2003). Learning about statistical covariation. *Cognition and Instruction, 21,* 1–78.
Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453–494). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology, 88,* 715–730.
Cramer, K. A., Post, T. R., & del Mas, R. C. (2002). Initial fraction learning by fourth- and fifth-grade students: A comparison of the effects of using commercial curricula with the effects of using the Rational Number Project Curriculum. *Journal for Research in Mathematics Education, 33,* 111–144.
Cummins, D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology, 20,* 405–438.
de Corte, E., Verschaffel, L., & de Win, L. (1985). Influence of rewording verbal problems on children's problem representations and solutions. *Journal of Educational Psychology, 77,* 460–470.
Empson, S. B. (1999). Equal sharing and shared meaning: The development of fraction concepts in a first-grade classroom. *Cognition and Instruction, 17,* 283–342.
Fuson, K. C. (1990). Conceptual structures for multiunit numbers: Implications for learning and teaching multidigit addition, subtraction, and place value. *Cognition and Instruction, 7,* 343–403.
Fuson, K. C., & Kalchman, M. (2002). A length model of fractions puts multiplication of fractions in the learning zone of fifth graders. In D. Mewborn, P. Sztajn, D. White, H. Wiegel, R. Bryant, & K. Nooney (Eds.), *Proceedings of the twenty-fourth annual meeting of the North American Chapter for the Psychology of Mathematics Education* (Vol. 4, pp. 1641–1650). Columbus, OH: ERIC Clearinghouse for Science, Mathematics and Environmental Education.
Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 15–46). New York: Macmillan.
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. 25–49). Cambridge, MA: MIT Press.
Harmon, M., Smith, T. A., Martin, M. O., Kelly, D. L., Beaton, A. E., Mullis, I. V. S., et al. (1997). *Performance assessment in IEA's third international mathematics and science study (TIMSS).* Boston: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.
Hiebert, J. (1986). *Conceptual and procedural knowledge: The case of mathematics.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Hiebert, J., & LeFevre, J. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Hiebert, J., & Wearne, D. (1986). Procedures over concepts: The acquisition of decimal number knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 199–223). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Hiebert, J., & Wearne, D. (1996). Instruction, understanding, and skill in multidigit addition and subtraction. *Cognition and Instruction, 14,* 251–283.

Hiebert, J., Wearne, D., & Taber, S. (1991). Fourth graders' gradual construction of decimal fractions during instruction using different physical representations. *Elementary School Journal, 91,* 321–341.

Hudson, T. (1983). Correspondences and numerical differences between disjoint sets. *Child Development, 54,* 84–90.

Kalchman, M., Moss, J., & Case, R. (2001). Psychological models for the development of mathematical understanding: Rational numbers and functions. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty five years of progress* (pp. 1–38). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92,* 109–129.

Klahr, D., & Carver, S. M. (1988). Cognitive objectives in a LOGO debugging curriculum: Instruction, learning, and transfer. *Cognitive Psychology, 20,* 362–404.

Knapp, N. F., & Winsor, A. P. (1998). A reading apprenticeship for delayed primary readers. *Reading Research and Instruction, 38*(1), 13–29.

Koedinger, K. R. (2002). Toward evidence for instructional design principles: Examples from Cognitive Tutor Math 6. In D. Mewborn, P. Sztajin, D. White, H. Wiegel, R. Bryant, & K. Nooney (Eds.), *Proceedings of the twenty-fourth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 21–50). Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science, 14,* 511–550.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8,* 30–43.

Koedinger, K. R., & MacLaren, B. A. (2002). *Developing a pedagogical domain theory of early algebra problem solving.* Pittsburgh, PA: Carnegie Mellon University.

Koedinger, K. R., & Nathan, M. J. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *Journal of Learning Sciences, 13,* 129–164.

Kouba, V. L., Carpenter, T. P., & Swafford, J. O. (1989). Number and operations. In M. M. Lindquist (Ed.), *Results from the fourth mathematics assessment of the national assessment of educational progress* (pp. 64–93). Reston, VA: National Council of Teachers of Mathematics.

Kouba, V. L., Zawojewski, J. S., & Strutchens, M. E. (1997). What do students know about numbers and operations. In P. A. Kenney & E. A. Silver (Eds.), *Results from the sixth mathematics assessment of the national assessment of educational progress* (pp. 87–140). Reston, VA: National Council of Teachers of Mathematics.

Lappan, G., Fey, J. T., Fitzgerald, W. M., Friel, S. N., & Phillips, E. D. (2002). *Connected mathematics.* Glenview, IL: Prentice Hall.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11,* 65–100.

Leinhardt, G. (1988). Getting to know: Tracing students' mathematical knowledge from intuition to competence. *Educational Psychologist, 23,* 119–144.

Lesh, R., Lovitts, B., & Kelly, A. E. (2000). Purposes and assumptions of this book. In A. E. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 17–34). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Lovett, M. C. (1998). Cognitive task analysis in service of intelligent tutoring system design: A case study of statistics. In B. P. Goetti, H. M. Halff, C. C. Redfield, & V. J. Shute (Eds.), *Intelligent tutoring systems. Lecture notes in computer science* (Vol. 1452, pp. 234–242). New York: Springer.
Mack, N. K. (1990). Learning fractions with understanding: Building on informal knowledge. *Journal for Research in Mathematics Education, 21,* 16–32.
Mack, N. K. (1993). Learning rational numbers with understanding: The case of informal knowledge. In T. P. Carpenter, T. A. Romberg, & E. Fennema (Eds.), *Rational numbers: An interpretation of research* (pp. 85–105). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Mayer, R. E., Mautone, P., & Prothero, W. (2002). Pictorial aids for learning by doing in a multimedia geology simulation game. *Journal of Educational Psychology, 94,* 171–185.
Mix, K. S., Levine, S. C., & Huttenlocher, J. (1999). Early fraction calculation ability. *Developmental Psychology, 35,* 164–174.
Moss, J., & Case, R. (1999). Developing children's understanding of the rational numbers: A new model and an experimental curriculum. *Journal for Research in Mathematics Education, 30,* 122–147.
Nathan, M. J., & Koedinger, K. R. (2000). An investigation of teachers' beliefs of students' algebra development. *Cognition and Instruction, 18,* 209–237.
Nathan, M. J., Long, S. D., & Alibali, M. W. (2002). The symbol precedence view of mathematical development: A corpus analysis of the rhetorical structure of textbooks. *Discourse Processes, 33,* 1–21.
Novick, L. R. (2001). Spatial diagrams: Key instruments in the toolbox for thought. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 40, pp. 279–325). San Diego, CA: Academic.
Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development, 6,* 449–468.
Pothier, Y., & Sawada, D. (1983). Partitioning: The emergence of rational number ideas in young children. *Journal for Research in Mathematics Education, 14,* 307–317.
Reese, C. M., Miller, K. E., Mazzeo, J., & Dossey, J. A. (1997). *NAEP 1996 mathematics report card for the nation and the states* (Rep. No. NCES 97–488). Washington, DC: National Center for Education Statistics.
Renkl, A., Atkinson, R. K., Maier, U. H., & Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *Journal of Experimental Education, 70,* 293–315.
Riley, M. S., Greeno, J. G., & Heller, J. J. (1983). Development of children's problem-solving ability in arithmetic. In H. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 153–195). New York: Academic.
Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91,* 175–189.
Rittle-Johnson, B., & Siegler, R. S. (1999). Learning to spell: Variability, choice, and change in children's strategy use. *Child Development, 70,* 332–348.
Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93,* 346–362.
Saxe, G. B. (1988). Candy selling and math learning. *Educational Researcher, 17*(6), 14–21.
Scardamalia, M., & Bereiter, C. (1985). Fostering the development of self-regulations in children's knowledge processing. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Research and open questions* (pp. 563–577). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Schoenfeld, A. H. (1985). *Mathematical problem solving.* New York: Academic.
Schraagen, J. M., Chipman, S. F., & Shalin, V. L. (Eds.). (2000). *Cognitive task analysis.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology, 8,* 481–520.
Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General, 116,* 250–264.

Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Sleeman, D. (1985). Basic algebra revisited: A study with 14-year-olds. *International Journal of Man Machine Studies, 22,* 127–149.
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12,* 257–285.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.



## APPENDIX A
## Equations Used in Alternative-Common-Denominator Model

| *Problem Type* | *Response Category* | *Equation* | *Verbal Explanation* |
|---|---|---|---|
| No scaffold | Correct | $cd \times ef \times cn$ | Find common denominator (*cd*), find equivalent fractions (*ef*), and combine numerators (*cn*) |
| | Combine-both error | $1 - cd$ | Do not find *cd* |
| | Fail-to-convert error | $cd \times (1 - ef) \times cn$ | Find *cd*, do not find *ef*, and *cn* |
| | Other error | $cd \times (1 - cn)$ | Remaining errors: Find *cd* and do not *cn* |
| Procedural scaffold | Correct | $u \times ef \times cn$ | Use given common denominator (*u*), find *ef* and *cn* |
| | Combine both error | $1 - u$ | Do not *u* |
| | Fail-to-convert error | $u \times (1 - ef) \times cn$ | *u*, do not find *ef*, and *cn* |
| | Other error | $u \times (1 - cn)$ | Remaining errors: *u* and do not *cn* |
| Conceptual or Contextual scaffold | Correct | $(cd + (1 - cd) \times m) \times ef \times cn$ | Find *cd* or find magnitude-based common denominator (*m*), and find *ef*, and *cn* |
| | Combine-both error | $1 - (cd + (1 - cd) \times m)$ | Do not find *cd* nor find *m* |
| | Fail-to-convert error | $(cd + (1 - cd) \times m) \times (1 - ef) \times cn$ | Find *cd* or find *m*, and do not find *ef*, and *cn* |
| | Other error | $(cd + (1 - cd) \times m) \times (1 - cn)$ | Remaining errors: Find *cd* or find *m* and do not *cn* |

# APPENDIX B
## Estimation Model: Equations for Conceptual and Contextual Scaffold Problems

| *Problem Type* | *Response Category* | *Equation* | *Verbal Explanation* |
|---|---|---|---|
| Conceptual or Contextual scaffold | Correct | $st + (cd \times ef \times cn) \times (1 - st)$ | Estimate the answer *(st)* or find common denominator (*cd*), find equivalent fractions (*ef*) and combine numerators (*cn*), if don't estimate |
| | Combine-both error | $(1 - st) \times (1 - cd)$ | Do not *st* and do not find *cd* |
| | Fail-to-convert error | $(1 - st) \times (cd \times (1 - ef) \times cn$ | Do not *st* and find *cd*, do not find *ef* and *cn* |
| | Other error | $(1 - st) \times (cd \times (1 - cn)$ | Remaining errors: Do not *st*, and find *cd*, and do not *cn* |

*Note.* Equations for no scaffold and procedural scaffold problems were the same as the alternative-common-denominator model.

# APPENDIX C
## Reject-Implausible-Answers Model: Equations for Conceptual and Contextual Scaffold Problems

| *Problem Type* | *Response Category* | *Equation* | *Verbal Explanation* |
|---|---|---|---|
| Conceptual or Contextual scaffold | Correct | $cd \times ef \times cn$ | Find common denominator (*cd*), find equivalent fractions (*ef*), and combine numerators (*cn*) |
| | Combine-both error | $(1 - cd) \times (1 - r)$ | Do not find *cd* and do not reject implausible (*r*) |
| | Fail-to-convert error | $cd \times (1 - ef) \times cn \times (1 - r)$ | Find *cd*, do not find *ef*, *cn*, and do not *r* |
| | Other error | $cd \times (1 - cn) + (1 - cd) \times r + cd \times (1 - ef) \times cn \times r$ | Remaining errors: Find *cd* and do not *cn*, or make combine-both or fail-to-convert error, but *r* |

*Note.* Equations for no-scaffold and procedural-scaffold problems were the same as the alternative-common-denominator model.