Adherence to Mathematics Professional Standards and Instructional Design Criteria for Problem-Solving in Mathematics

ASHA K. JITENDRA
Lehigh University

CYNTHIA GRIFFIN
University of Florida

ANDRIA DEATLINE-BUCHMAN
Lehigh University

CAROLINE DIPIPI-HOY
Easton School District

EDWARD SCZESNIAK
NATALIE G. SOKOL
Lehigh University

YAN PING XIN
Purdue University

ABSTRACT: This study investigated the extent to which teaching the recommended methods for problem-solving presented in third-grade mathematics textbooks adhered to the National Council of Teachers of Mathematics Standards and instructional design criteria. Results indicated that there were more variations than similarities within and across textbooks in meeting the Standards. In terms of the instructional design criteria, most were satisfied in only two textbooks and few discrepancies were evident across the textbooks that were evaluated. For example, instructional design criteria of clarity of objectives, sufficient teaching examples, and nonexamples were met in less than half of the textbooks. Additional findings and implications for practitioners meeting the diverse needs of students with learning problems are discussed.

Extant dissatisfaction with student performance in America's schools during the last decade has led to national reform efforts to improve the content and quality of school curriculum and instruction. Reform initiatives, such as the standards-based reforms, called for more challenging learning standards and school accountability (Chatterji, 2002; Nolet & McLaughlin, 2000). As a result, individ-
uals developed their own versions of subject area standards that shaped the ways in which they responded to curriculum and instruction (Council of Great City Schools, 1996; Gandel, 1997). The standards-based reforms occurred at a time when classrooms were becoming increasingly diverse with students who represent a wide range of economic and cultural backgrounds, whose native language often is not English, or who have a learning difficulty (Carnine, Jones, & Dixon, 1994). We find it interesting that policy and legislative statements at both the state and national levels suggested that standards-based reforms pertain to all students. The reform efforts are clearly reflected in the 1997 reauthorization of the Individuals With Disabilities Act (IDEA, 1999), which contains explicit language that students with disabilities have meaningful access to the general education curriculum and participate in district and state testing.

A key aspect of the standards-based approach is its emphasis on the development of conceptual understanding and reasoning over memorization and rote learning. An important implication of the standards-based reforms is that complex, higher order thinking and problem-solving have become integral to the concept of what constitutes learning in general education classrooms. In mathematics education, particularly, we witnessed ways in which the field responded to the discontent with the traditional curriculum that was criticized for being “relatively repetitive, unfocused, and undemanding” (Hiebert, 1999, p. 11). The discontent led to the “National Research Council issuing Everybody Counts and the National Council of Teachers of Mathematics (NCTM) issuing Curriculum and Evaluation Standards for School Mathematics (the Standards) in 1989” (Schoenfeld, 2002, p. 14) followed by the Principle and Standards for School Mathematics (Principles and Standards) in 2000. A key aspect of the standards-based approach is its emphasis on the development of conceptual understanding and reasoning over memorization and rote learning. Recent developments in theories of how students learn and changes in marketable workplace skills may explain the pedagogical shift from direct instruction, drill, and practice toward more active student engagement with mathematical ideas (Goldman, Hasselbring, & the Cognition and Technology Group at Vanderbilt, 1997; Goldsmith & Mark, 1999; Hiebert et al., 1996).

Despite the notion that effective mathematics instruction should emphasize understanding and reasoning, most students’ mathematical thinking is overly “restricted to the domain of mathematical computation and formula following” (Bransford, Zech, Schwartz, Barron, & Vye, 1996, p. 203). Further, the highly procedural instruction in special education may sustain the characterization of students with learning disabilities as passive learners (Torgesen, 1982). Not surprisingly, student underachievement in mathematics is most severe for students with learning disabilities (Carnine et al., 1994; Parmar, Cawley, & Frazita, 1996) who experience difficulties in several aspects of mathematics (see Mastropieri, Scruggs, & Shiah, 1991; Miller, Butler, & Lee, 1998; Rivera, 1997) and who lag considerably behind their normally achieving peers (Parmar et al., 1996; Zentall & Ferkis, 1993). For these students, the prevalence of problems in mathematics, particularly poor problem conceptualization, has serious implications with regard to accessing the general curriculum, because it interferes with the ability to learn higher level math content and skills (Jitendra & Xin, 1997; Parmar, Cawley, & Miller, 1994).

Although there is consensus that mathematics literacy and the need to provide high quality mathematics instruction is an important goal for all students (Schoenfeld, 2002), the Standards pose several challenges in terms of implementation and meeting the diverse needs of all students (Bybee, Ferrini-Mundy, & Loucks-Horsley, 1997). Teachers find themselves in a quandary as the range of skills students bring to the classroom increases. The thinking and problem-solving skills represented in the Standards require generative instructional strategies (Cobb & Bauersfeld, 1995; Hiebert et al., 1996; Lampert, 1991), in which students are encouraged to construct their own understanding of mathematical concepts and rela-

320

Spring 2005
tionships (e.g., developing the abilities of inquiry, problem-solving, and mathematical connections) to allow them to direct their own learning (Bybee et al., 1997; Goldman et al., 1997; Kinzer, Risko, Vye, & Sherwood, 1988; Schoenfeld, 1991). Although the Standards provide a framework for curriculum development, the primary challenge that teachers encounter is designing instruction that requires synthesizing knowledge regarding the Standards (e.g., mathematical content and processes), student learning, effective instructional models, and appropriate classroom opportunities and experiences (Bybee et al.).

Recent research in instructional approaches based on the concept of situated learning suggests its effectiveness for at-risk students (Cognition and Technology Group at Vanderbilt, 1990; Hiebert, 1999; Woodward & Montague, 2002). Situated instruction refers to authentic, meaningful contexts for applying problem-solving skills. The elaborated instructional context in situated learning serves as an effective support mechanism during learning for at-risk students. However, generative instructional strategies place most of the responsibility on the learner for organizing, sequencing, and integrating information. Learners who lack sufficient prior knowledge may need more supportive instructional strategies in which the teacher scaffolds information processing for the learner by supplying a greater degree of instructional facilitation during the learning process (Smith & Ragan, 1999). As teachers find themselves caught in this balancing act between the demands of the curriculum and the needs of their students, the quality of curriculum materials used in classrooms becomes crucial. In some school districts, the only forms of curriculum that exist are textbooks, which are considered the primary means of imparting new information to students (Armbruster & Ostertag, 1993; Chandler & Brosnan, 1994; Garner, 1992; Parmar, 1992; Porter, 1989; Shutes & Petersen, 1994; Valverde & Schmidt, 1997/1998). For example, Garner points out that “textbooks serve as critical vehicles for knowledge acquisition in school” and can “replace teacher talk as the primary source of information” (p. 53).

Textbook materials continue to be a challenge in mathematics education (Flanders, 1994; Lappan, 1999; Saminy & Liu, 1997). Results of the Third International Mathematics and Science Study (TIMSS) suggest that U.S. textbooks, as compared to those of other countries, are much larger and heavier, cover more topics with less depth, and fail to develop linkages between topics (Valverde & Schmidt, 1997/1998). A comparative analysis of seventh-grade Japanese and U.S. textbooks indicated that U.S. textbooks focused more on “eye catching,” irrelevant illustrations, and unsolved practice problems, whereas Japanese textbooks emphasized relevant illustrations, “worked out” teaching examples, and multiple representations (Mayer, Sims, & Tajika, 1995). The Japanese texts were also better integrated and shorter than U.S. texts.

Although publishers have attempted to conform to the changing trends in the way that educational reforms dictate the content of mathematics textbooks, U.S. mathematics textbooks continue to be criticized for their lack of adherence to the Standards (Mokros, 1994) and their poor instructional design in meeting the needs of students with disabilities (Carnine, Jitendra, & Silbert, 1997; Jitendra, Carnine, & Silbert, 1996; Jitendra, Salmento, & Haydt, 1999). Given that textbooks are considered a de facto national curriculum (Mayer et al., 1995), it seems reasonable to periodically analyze and revise them rather than eliminate their use completely (Crawford & Snider, 2000). Indeed, textbooks should be modified in ways that are clearly intended to affect the valued outcomes, which are defined by the Standards as “what students should know and be able to do” (Bybee et al., 1997, p. 331). General educators, for various reasons, do not readily adapt or modify their curricula or instructional methods.
Textbook materials continue to be a challenge in mathematics education. for students with disabilities (Baker & Zigmond, 1990; Fuchs, Fuchs, & Bishop, 1992; Kagan & Tippins, 1991). As such, access to the general curriculum for these students may require that teachers be provided with the knowledge needed to meet their student needs. Therefore, this study was designed to evaluate the extent to which five 3rd-grade mathematics textbooks adhered to the Standards and instructional design criteria in teaching problem-solving. This study extends earlier research that evaluated the content of mathematics textbooks and contributed to an emerging research base on comparisons of content changes in mathematics textbooks since the publication of the NCTM standards (e.g., Chandler & Brosnan, 1994; Jitendra et al., 1999).

**METH O D**

Five 3rd-grade mathematics textbooks, Harcourt (HC; Maletsky et al., 2002); Houghton Mifflin (HM; Greenes, Leiva, & Vogeli, 2002); McGraw Hill (MH; Carlsson et al., 2002); Scott Foresman-Addison Wesley (SFAW; Charles et al., 2002); and Silver Burdett & Ginn (SBG; Fennell et al., 2001) were evaluated. The selection of these textbooks was based on consultations with mathematics educators, teachers, and school administrators, and because they are representative of mathematics textbooks typically adopted in the United States. Given that problem-solving, reasoning, communicating, connecting, and representing mathematical content are important goals in the Standards across grade levels and content, in this study we decided to focus on these processes as we investigated the content area of whole number operations.

The data source consisted specifically of lessons on addition and subtraction of whole numbers in the five textbooks. Addition and subtraction lessons were chosen because they allow for a sampling across a range of textbooks as well as provide a focus on word problem-solving. Within the lessons on addition and subtraction, we focused on word problems (i.e., problems or questions stated in words) to investigate how the textbooks adhered to the goals of the Standards. In addition, we scrutinized all addition and subtraction lessons with respect to meeting the instructional design criteria (e.g., prior knowledge, explanations). The underlying rationale for focusing on third-grade mathematics textbooks was that by that grade level students have the necessary reading skills to comprehend increasingly complex word problems. In addition, students would have the necessary written and verbal communication skills to engage in mathematical discourse and fundamental mathematical computation skills. Further, if we are to assist students with disabilities to access the general education curriculum, promote later learning, and reduce subsequent gaps between them and their normally achieving peers, it is important to attend to the mathematics content in early grades. In sum, the data source consisted of a total of 141 lessons on addition and subtraction, ranging from 16 to 40 lessons in each of the five textbooks.

**DATA ANALYSIS AND CODING PROCEDURES**

To conduct an analysis of the textbooks with regard to meeting the Standards, we examined all addition and subtraction lessons, including preview, midchapter, and end-of-unit review lessons, in each textbook. The raters broke each lesson into six parts including preview, warm-up/introduce, teach, practice, assess, and review and then coded teacher directions in each section of the lesson using the criteria for evaluating the Standards. We did not include sections on extended activities (e.g., additional review and practice) within a lesson, because these are typically optional exercises provided by the publisher. It must be noted that to quantify the analysis of textbooks in this study, we had to operationally define the criteria for the NCTM Standards. However, the criteria may not be comprehensive as conceptualized; thus, conclusions based on this analysis are tentative.

Two raters read each lesson in its entirety to identify word problems involving addition and subtraction. They also recorded the number of problem-solving lessons that were identified as such by the publisher in each textbook. Next, they compared the information on word problems in the lesson (e.g., explanations, practice) as it related to the Standards (i.e., problem-solving,
reasoning, communication, connection, and representation). All conflicts between the two raters were resolved by consensus. A third rater independently rated 30% of the lessons for the purpose of interrater reliability. Similar procedures were used to evaluate the textbooks with regard to meeting the instructional design criteria. However, the raters evaluated the entire lesson content (e.g., computational problems) rather than word problems only.

An analysis of the Standards focused on problem-solving, reasoning, communication, connections, and representations. The coding procedure we used in our analysis is shown in Table 1. "Problem-solving" refers to problem-solving opportunities that encourage students to solve interesting and challenging problems (Lappan, 1999). For this Standard, the raters coded the number of word problems, unique problem-solving contexts, and different word problem-solving strategies in each section of the lesson. To code for unique contexts, for example, word problems that involved "rafts" or "boats" would be coded as the same context, whereas a word problem that discussed "caves" would be coded as a different context. In addition, a strategy was coded only once in a lesson.

"Reasoning" includes skills such as making mathematical conjectures, exploring phenomena by observing, examining conjectures, justifying findings, and developing mathematical arguments (Goldsmith & Mark, 1999). To code for this Standard, the directions for the word problem task had to specify "guess and check," "justify choice of action," or require solving a problem using a specific strategy (e.g., count up) followed by an alternate strategy (act it out) to get to the same solution.

"Communication" refers to communicating mathematical thinking coherently and clearly to others using both verbal and written communication. To code for this Standard, the raters recorded opportunities for organizing and consolidating mathematical thinking, articulating reasoning to peers and/or the teacher, and using the language of mathematics to express mathematical ideas.

"Connections" refer to relating new material to students' prior knowledge, skills, experiences, and interests (Buchanan & Helman, 1997; Hiebert, 1999; Lappan, 1999). For this standard, we had to infer making connections among mathematical ideas by reading the explanation portion of the lesson, because they were not explicitly stated in the directions. For example, the raters coded as making connections between addition and multiplication when the directions had students solve a word problem using the addition operation followed by the teacher modeling "doubling" (i.e., multiplying by 2s) as an easier way to solve the problem. In addition, we coded connections to other subject areas when the publisher explicitly identified them as such (e.g., "Science Connection"). Finally, we quantified connecting problem-solving to real world contexts by tallying all problems that involved personalized contexts.

"Representations" (i.e., diagrams, graphs, models, tables, pictures, manipulatives, symbolic expressions) that define mathematical relationships help students organize their thinking and interpret mathematical situations (Goldsmith & Mark, 1999). For this standard, the raters coded word problem tasks that required students to generate their own representations, to select among representations, and to apply a representation provided in the text (e.g., information presented and organized in a diagram or chart) to solve word problems.

An analysis of instructional design criteria focused on the teaching of addition and subtraction skills. The criteria examined in this analysis are shown in Table 2. The raters examined the target lesson in the teacher manuals for the presence or absence of specific design criteria (e.g., clarity of objective, explicit teaching explanations, effective feedback) and assigned a rating from 0 to 2 points based on various levels of satisfaction of the criterion. When appropriate, the number of instances of the criterion (e.g., number of concepts or skills, teaching examples and nonexamples, practice and review examples) that occurred in the lesson also was computed. The coding for prerequisite skills was dichotomous as either present or absent for the lesson.

To determine the appropriate number of examples of a specific criterion, we examined the literature for directions to quantify them. For example, research suggests teaching one new concept or skill at a time so that adequate resources of working memory are made available for learn-
<table>
<thead>
<tr>
<th>Standard</th>
<th>Definition</th>
<th>Code</th>
<th>Dependent Data</th>
</tr>
</thead>
</table>
| Problem Solving | • A problem or question stated in words as a word problem  
• A context (e.g., pets, recreation, money) is considered unique if it differs from any of the word problem contexts previously recorded in the lesson  
• A problem-solving strategy (e.g., draw a picture, make a table, look for patterns, make a number sentence, solve a simpler problem, act it out, guess and check) to solve word problems                                                                                                                                                                                                 | No. of word problems in each section     | No. of word problems per lesson                                                                 |
| Reasoning   | • Reasoning referred to (a) making mathematical conjectures (guessing and checking findings), (b) justifying findings, and (c) using alternative solution methods                                                                                                                                                                                                                                                                              | No. of word problem tasks in each section that specified:  
(a) "guess and check"  
(b) "justify choice of action;"  
(c) use of alternative solution methods | No. of word problem contexts per lesson |
| Communication | • Communication included (a) organizing and consolidating mathematical thinking, (b) articulating reasoning to peers and/or the teacher, and (c) using the language of mathematics (e.g., addend, sum, estimate) to express mathematical ideas                                                                                                                                                                                                                                                      | No. of word problem tasks in each section that specified:  
(a) "write about it"  
(b) "explain"  
(c) use of specific mathematical terms | No. of strategies across all lessons |
| Connections | • Connections defined as (a) making connections among mathematical ideas (e.g., addition and multiplication), (b) applying mathematics in contexts outside of mathematics (e.g., geography, science, social studies, language arts), and (c) connecting problem-solving to real world contexts (e.g., adding books in the classroom, counting pets at home, creating own problems)                                                                                     | No. of word problem tasks in each section that specified connections:  
(a) among mathematical ideas  
(b) to other subject areas  
(c) to real world contexts | No. of percentage of lessons that included strategies  
No. of word problem contexts per lesson  
No. of strategies per lesson |
ing new skills and solving complex problems (Carnine et al., 1994; Klausmeier, 1992; Porter, 1989). Further, effective skill explanations should demonstrate application of a strategy using sufficient examples (i.e., 3 to 5) before moving on to guided practice (Cawley, Parmar, Yan, & Miller, 1996; Gagne, 1985). The raters coded practice as sufficient if a lesson provided 10 or more problems. We acknowledge that using 10 or more problems as the cutoff is arbitrary and not guided by any known standard. However, the underlying rationale for this criterion is based on our knowledge of mathematics curricula and third grade students. At the same time, when sufficient problems are available, teachers can assign more or less practice items as needed to accommodate individual needs of students. To quantify the number of sufficient review problems, we rationalized that addition and subtraction skills compose one of the major content areas in third grade and that review of these skills should be frequent. As such, we decided that the minimum for meeting this criterion is at least four review problems per lesson. This decision is based on the results of another study in which the mean proportion of review problems in seven mathematics textbooks was four (Jitendra et al., 1999).

**Training and Reliability**

The evaluators met in a training session with the first author for approximately 4 hr prior to initiation of data collection. Following this training session, interrater agreement was computed between each evaluator and the first author. Reliability was calculated as the number of agreements divided by the number of agreements and disagreements and multiplied by 100. All reliability estimates were above 90%. As a further check on reliability, another rater (a graduate student) independently sampled 30% of the lessons in each textbook and evaluated all dimensions of the lessons in each of the five textbooks. Across textbooks, the mean interrater agreement for coding the NCTM Standards and design of instruction criteria was 91% (range = 87% to 95%) and 88% (range = 86% to 93%), respectively.
<table>
<thead>
<tr>
<th>Criterion</th>
<th>Definition</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective</td>
<td>Objectives listed at the beginning of each addition and subtraction lesson, which specified the goal of instruction by identifying the critical content and functioned as a student performance indicator.</td>
<td>0 = No lesson objective</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = Lesson objective present but stated globally or incompletely (e.g., subtraction)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = Lesson objective complete and stated in specific terms (e.g., subtraction across zeros with regrouping)</td>
</tr>
<tr>
<td>Concept/Skills</td>
<td>Objectives examined for the introduction of mathematical concepts or skills (e.g., subtraction with regrouping), including the newly introduced target skill.</td>
<td>No. of concepts presented per lesson</td>
</tr>
<tr>
<td>Prerequisite Skills</td>
<td>Presentation of basic skills or simpler component skills (e.g., place value, subtraction facts, and regrouping) that are prerequisite for the target higher order skill (e.g., subtraction across zeros problems).</td>
<td>0 = No review of prerequisite skills</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = Review of prerequisite skills present</td>
</tr>
<tr>
<td>Teaching explanation</td>
<td>Explanations for symbolic or manipulative activities that clarify concepts and procedures, promote understanding, and make connections between the activity and the concept. These were examined in sections entitled, &quot;Motivate and Teach,&quot; &quot;Teach,&quot; and &quot;Learn About It.&quot;</td>
<td>0 = No teacher explanations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = Explanations present, but not clear</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = Explanations clear in promoting understanding and linkage</td>
</tr>
<tr>
<td>Teaching examples</td>
<td>Teaching of a new concept or skill (e.g., subtraction with regrouping) includes the presentation of both examples (problems that require regrouping) and nonexamples (problems that do not need regrouping).</td>
<td>No. of examples and nonexamples in the explanation portion of each lesson</td>
</tr>
<tr>
<td>and nonexamples</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Practice problems</td>
<td>Typically, practice examples specified by the publisher in sections entitled, &quot;Practice,&quot; &quot;Independent Practice,&quot; &quot;Try It,&quot; &quot;Practice and Problem Solving,&quot; or &quot;Try It Out&quot; followed teaching examples.</td>
<td>No. of practice problems presented per lesson</td>
</tr>
<tr>
<td>Review</td>
<td>Review problems specified as such by the publisher.</td>
<td>No. of review problems presented per lesson</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0 = No review</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = insufficient review (1 to 3 problems per lesson)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = sufficient review (4 or more problems per lesson)</td>
</tr>
</tbody>
</table>
RESULTS

NCTM STANDARDS

A summary of findings with respect to the NCTM Standards in the five mathematics textbooks is presented in Table 3. We calculated the proportion in each lesson that met the criterion (e.g., reasoning, connection) based on the total number of all addition and subtraction lessons counted in a textbook. Results indicated that although opportunities for problem-solving were present most often across the five textbooks, our analysis did not reveal that to be true for the other Standards. A description of our findings organized by the Standards follows.

Problem-Solving. Although the number of problem-solving lessons per se across the five textbooks was low ($M = 4.4$, range = 2 to 7), the proportion of word problems for the five textbooks was high ($M = 6.00$). The highest score was 8.48 word problems per lesson in MH. Scores for the remaining four textbooks ranged from 4.69 to 5.90. With regard to solving problems in different contexts, the mean score across the textbooks was 2.92 contexts per lesson. Although the highest score of 3.22 was found in MH, scores for the remaining four textbooks were not very different (HM and SFAW = 3.00, HC = 2.94, SBG = 2.44). It is interesting to note that the mean number of problem-solving strategies across textbooks was 11.4 (range = 4 to 16). Specifically, the mean percentage of lessons that applied strategies to solve word problems was 50.8%. The highest percentage was 60% in SFAW followed by 52.0%, 52.0%, 46.0%, and 44.0% in HM, MH, SBG, and HC, respectively. However, the mean score across textbooks for applying specific strategies to solve word problems was 0.39 strategies per lesson. MH scored the highest (0.56), followed by SBG (0.43), HM (0.38), SFAW (0.35), and HC (0.25).

Reasoning. The mean scores for using reasoning in terms of checking findings ($M = 0.24$), justifying findings ($M = 0.12$), and alternate methods of reasoning ($M = 0.22$) were low across the five textbooks. However, when this Standard was met, most textbooks provided explicit guided instruction to facilitate reasoning. The highest score for checking findings was 0.32 word prob-
### TABLE 3
**NCTM Standards**

<table>
<thead>
<tr>
<th>NCTM Standards</th>
<th>HC</th>
<th>HM</th>
<th>MH</th>
<th>SFAW</th>
<th>SBG</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of addition and subtraction lessons</td>
<td>16</td>
<td>21</td>
<td>27</td>
<td>40</td>
<td>37</td>
<td>28.20</td>
</tr>
<tr>
<td>Number of problem-solving lessons</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>7</td>
<td>4.40</td>
</tr>
<tr>
<td>Number of problem-solving strategies</td>
<td>4</td>
<td>8</td>
<td>15</td>
<td>14</td>
<td>16</td>
<td>11.40</td>
</tr>
<tr>
<td>Percentage of lessons with problem-solving strategies</td>
<td>43.75</td>
<td>52.38</td>
<td>51.85</td>
<td>60.00</td>
<td>47.22</td>
<td>51.04</td>
</tr>
</tbody>
</table>

I. **Problem Solving (proportion)**

1. Word problems | 4.69 | 5.90 | 8.48 | 5.35 | 5.59 | 6.00 |
2. Word problems with different contexts | 2.94 | 3.00 | 3.22 | 3.00 | 2.44 | 2.92 |
3. Variety of strategies | 0.25 | 0.38 | 0.56 | 0.35 | 0.43 | 0.39 |

II. **Reasoning (proportion)**

4. Check findings | 0.31 | 0.24 | 0.19 | 0.15 | 0.32 | 0.24 |
5. Justify findings | 0.13 | 0.14 | 0.00 | 0.00 | 0.32 | 0.12 |
6. Use alternate methods of reasoning | 0.19 | 0.05 | 0.19 | 0.35 | 0.30 | 0.22 |

III. **Communication (proportion)**

7. Organize and consolidate mathematical thinking by writing about it | 0.31 | 0.24 | 0.33 | 0.13 | 0.03 | 0.21 |
8. Articulate own reasoning by talking about it | 0.75 | 0.81 | 0.22 | 0.44 | 0.57 | 0.56 |
9. Use the language of mathematics | 0.94 | 0.48 | 0.37 | 0.13 | 0.38 | 0.46 |

IV. **Connections (proportion)**

10. Make connections among mathematical ideas | 00.0 | 0.05 | 0.07 | 0.10 | 0.24 | 0.09 |
11. Apply mathematics in contexts outside of mathematics | 0.38 | 0.00 | 0.30 | 0.68 | 0.13 | 0.30 |
12. Connect problem-solving to real world contexts | 0.19 | 0.05 | 0.11 | 0.25 | 0.30 | 0.18 |

V. **Representations (proportion)**

13. Generate own mathematical representations | 0.06 | 0.00 | 0.33 | 0.53 | 0.35 | 0.25 |
14. Select among mathematical representations | 0.00 | 0.90 | 0.15 | 0.48 | 0.16 | 0.34 |
15. Apply text presented mathematical representations | 1.31 | 1.38 | 2.11 | 1.05 | 1.08 | 1.39 |

*Note.* HC = Harcourt Brace (2002); HM = Houghton Mifflin (2002), MH = McGraw-Hill (2002); SFAW = Scott, Foresman-Addison Wesley (2002); SBG = Silver Burdett & Ginn (2001). Proportion was calculated by dividing the number of instances of the criterion (e.g., checking findings, using the language of mathematics) by the total number of all addition and subtraction lessons in the textbook.
lems per lesson in SBG. For example, in one of the lessons in SBG, students were engaged in reasoning when they were taught to add more than two numbers and then check their answer by adding up. SFAW (0.15) scored the lowest on this criterion. For justifying findings, SBG scored the highest (0.32 problems per lesson). Specifically, an example in SBG that met this criterion had students estimate whether a restaurant had enough pineapples for a party based on a picture showing the number of pineapples in each of two crates of pineapples. Next, students had to determine whether or not estimation could be used to solve the problem and justify their choice. In contrast, scores for the remaining four textbooks were very low (HM = 0.14, HC = 0.13) or not present (e.g., SFAW and MH). For using alternate methods of reasoning, SFAW had the highest score of 0.35 problems per lesson. For example, although students in SFAW were taught to subtract three-digit numbers using place-value blocks, they could also use the “act it out” strategy as an alternate method to solve the problem. Although SBG scored 0.30 on this criterion, scores for the remaining three textbooks were very low (HC = 0.19, MH = 0.19, HC = 0.05).

Communication. The mean scores across textbooks were low for (a) organizing and consolidating mathematical thinking (M = 0.21), (b) articulating own reasoning (M = 0.56), and (c) using the language of mathematics (M = 0.46). MH scored the highest with 0.33 problems per lesson for organizing and consolidating mathematical thinking. Although HC scored 0.31 on this criterion, scores for the remaining three textbooks were very low (HM = 0.24, SFAW = 0.13, SBG = 0.03). Scores for articulating own reasoning were high in HM (0.81) and HC (0.75). Although MH had the lowest score of 0.22, scores for SBG and SFAW were 0.57 and 0.44, respectively, on this criterion. With regard to using the language of mathematics, the highest score of 0.94 was found for HC. For example, terms such as estimate, addend, and sum appeared often in HC when solving word problems. Although HM included 0.48 problems per lesson that met this criterion, scores for SBG (0.38) MH (0.37), and SFAW (0.13) were very low.

Connections. Results for making connections among mathematical ideas indicated a low mean score of 0.09 problems per lesson across the textbooks. The highest score was 0.24 for SBG and the lowest was 0 in HC. With regard to applying mathematics in other contexts, the mean score was 0.30. Although SFAW scored high (0.68) on this criterion, scores for the remaining four textbooks were low (HC = 0.38, MH = 0.30, SBG = 0.13, and HM = 0.00). When opportunities to connect problem-solving to real world contexts were examined, the mean score for the five textbooks was low (0.18 problems per lesson). SBG scored the highest on this criterion (0.30). An example illustrating this criterion in SBG included a lesson in which students had to create their own problem about a sled dog race using the numbers provided in a map in the text and then have a friend solve the problem. Scores for the remaining four textbooks were 0.25 in SFAW, 0.19 in HC, 0.11 in MH, and 0.05 in HM.

Representations. The mean scores across textbooks varied for (a) generating own representations (M = 0.25), (b) selecting among mathematical representations (M = 0.34), and (c) applying mathematical representations (M = 1.39) to solve problems. For generating representations, SFAW scored the highest of 0.53 problems per lesson, followed by a score of 0.35 in SBG, 0.33 in MH, 0.06 in HC, and 0.00 HM. With regard to selecting among mathematical representations, the highest score of 0.90 was found for HM, followed by 0.48 in SFAW, 0.16 in SBG, and 0.15 in MH. HC scored 0 on this criterion. Scores for applying mathematical representations to solve problems were relatively high (M = 1.39), with the highest score of 2.11 in MH followed by 1.38 in HM, 1.31 in HC, 1.08 in SBG. The exception was SFAW, which scored 0.05 on this criterion.

Instructional Design Criteria
A summary of findings with respect to instructional design criteria in the five mathematics textbooks is presented in Table 4. Results indicated that most criteria were satisfied in only two textbooks and few discrepancies were evident across the textbooks. A description of our findings that is organized by instructional design criteria follows.

Criterion #1: Clarity of Objective. Although all five textbooks provided lesson objectives, only
### Table 4
Design of Instruction Criteria

<table>
<thead>
<tr>
<th>Instructional Criteria</th>
<th>Mathematics Textbooks</th>
<th>% met criterion&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HC</td>
<td>HM</td>
</tr>
<tr>
<td>1. Objective (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present&lt;sup&gt;c&lt;/sup&gt;</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Implicit</td>
<td>0.00</td>
<td>80.95</td>
</tr>
<tr>
<td>Explicit&lt;sup&gt;c&lt;/sup&gt;</td>
<td>100.00</td>
<td>19.05</td>
</tr>
<tr>
<td>2. Concept/Skills (proportion) &lt;sup&gt;d&lt;/sup&gt;</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>3. Prerequisite Skills (%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>100.00</td>
<td>76.19</td>
</tr>
<tr>
<td>4. Teaching explanation (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present&lt;sup&gt;c&lt;/sup&gt;</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Implicit</td>
<td>21.43</td>
<td>37.50</td>
</tr>
<tr>
<td>Explicit&lt;sup&gt;c&lt;/sup&gt;</td>
<td>78.57</td>
<td>62.50</td>
</tr>
<tr>
<td>5. Teaching examples (proportion) &lt;sup&gt;e&lt;/sup&gt;</td>
<td>3.08</td>
<td>1.71</td>
</tr>
<tr>
<td>6. Teaching nonexamples (proportion) &lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>7. Practice problems (proportion) &lt;sup&gt;f&lt;/sup&gt;</td>
<td>25.00</td>
<td>25.19</td>
</tr>
<tr>
<td>8. Review</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present&lt;sup&gt;c&lt;/sup&gt;</td>
<td>75.00</td>
<td>80.00</td>
</tr>
<tr>
<td>Review examples (proportion) &lt;sup&gt;f&lt;/sup&gt;</td>
<td>5.75</td>
<td>7.29</td>
</tr>
<tr>
<td>9. Effective feedback (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present&lt;sup&gt;c&lt;/sup&gt;</td>
<td>33.00</td>
<td>57.00</td>
</tr>
<tr>
<td>Not Instructive</td>
<td>0.00</td>
<td>33.33</td>
</tr>
<tr>
<td>Instructive&lt;sup&gt;c&lt;/sup&gt;</td>
<td>100.00</td>
<td>66.67</td>
</tr>
<tr>
<td>Total percent met criterion&lt;sup&gt;b&lt;/sup&gt;</td>
<td>84.61</td>
<td>53.85</td>
</tr>
</tbody>
</table>

<sup>a</sup> Percent meeting criterion was calculated by dividing the number of textbooks that met the criterion (or number of criteria met by a textbook<sup>b</sup>) by the total number of textbooks (or total number of criteria<sup>b</sup>). Criterion was defined as follows: "" = 75% or more; "" = 1; "" = 3 or more; "" = 10 or more; "" = 4 or more.

28% of the objectives in the textbooks analyzed were complete and clear. Across the five textbooks, most included lessons with adding two or three digit numbers. HC was the only program in which 100% of the objectives were clear and complete. For example, an objective in HC read as follows: "To find the sum of three two-digit addends" (Maletsky et al., 2002, p. 36). In contrast, all of the objectives (100%) in SFAW and MH were stated globally. Whereas SFAW reported adding two-digit numbers as a lesson objective, MH noted adding more than two numbers in a lesson objective. The mean for clearly stated objectives in SBG and HM was 26% and 19%, respectively.

**Criterion # 2: Concepts/Skills Taught.** With the exception of SBG, none of the other textbooks specified any additional skills other than the target skill in the lesson objectives. Across the textbooks, an average of one concept was introduced per lesson. As such, all textbooks met this criterion.

**Criterion # 3: Prerequisite Skills Taught.** With regard to the presence of review of prerequisite skills, the mean across textbooks was high (82%). Whereas HC and SBG provided review.
100% of the time, the mean for HM and MH was 76% and 81%, respectively. SFAW was the only textbook that did not meet this criterion (55%).

Criterion #4: Explicit Teaching Explanations. All textbooks provided some form of teacher explanations. The mean percentage of clear explanations across textbooks was 75%. Three of the five textbooks, HC, MH, and SBG, met this criterion of providing explicit teaching explanations more than 75% of the time. For example, in a lesson in SBG, instruction included facilitative questioning and manipulatives (base-10 blocks) to teach subtracting two- and three-digit numbers. An explicit worked-out example was used to show the connection between the manipulation of the base-10 blocks and the concept of regrouping a two-digit number from a three-digit number. In MH, a lesson on adding multi-digit whole numbers with and without regrouping, teaching included a worked-out example using place-value models to illustrate if regrouping is necessary. The worked-out example included a sequence of specific steps to solve the problem.

Another example illustrating explicit explanation was in HC, which used several examples that were worked out in one lesson to guide instruction when teaching addition of three-digit numbers with and without regrouping. Further, the lesson discussed modifying instruction by “modeling the instructional steps with base-10 blocks, individually or in pairs” (Maletsky et al., 2002, p. 42). In contrast, SFAW and HM incorporated explanations that were not clear 47% and 38% of the time, respectively. In these two programs, students were presented with examples, but the teaching of specific steps was not apparent. In general, the instruction presented had students discover the concepts or steps on their own. For example, in an addition lesson in SFAW, teacher directions had students “explore” finding sums (e.g., 27 + 20) using a 100s chart by starting with the larger number (27) and moving forward two 10s on the chart to find the sum 47. In another lesson on estimating sums by rounding numbers in HM, teacher directions had students use a number line to estimate 208 +182 as follows: “Round each number to the greatest place. Then add.” (Greene et al., 2002, p. 110).

Criteria #5 and #6: Sufficient Teaching Examples and Nonexamples. With regard to teaching examples, the mean frequency was low (M = 2.35). Although SFAW (M = 3.67) and HC (M = 3.08) provided sufficient number of examples, HM (M = 1.71), SBG (M = 1.85), and MH (M = 1.44) did not include adequate number of teaching examples. Across the five textbooks, none provided nonexamples to teach the new skill and/or concept.

Criterion #7: Adequate Practice. The mean number of practice examples was high and similar across the five textbooks (M = 24; range = 20 to 25).

Criterion #8: Sufficient Review. Although review was present in the majority of the lessons (M = 78%), the percentage of lessons in which review was present was the highest in SBG (92%) and lowest in SFAW (61%). The proportion of review problems across the five textbooks was high (M = 9.86; range = 5.75 to 14.85).

Criterion #9: Effective Feedback. With respect to feedback, four of the five textbooks included error correction procedures most of the time (M = 65%). The one exception was HC in which error interventions were present in only 33% of the lessons examined. When feedback was present, HC, HM, MH, and SFAW included instructive or elaborated feedback most of the time (100%, 67%, 93%, and 90%, respectively). For example, in a lesson on finding the sum of three 2-digit addends, the error alert in HC noted that, “students may group two addends to make a ten and then forget the third addend” (Maletsky et al., 2002, p. 37). The error intervention specified prompting students to “mark off addends” as they add them. In a subtraction lesson with subtracting four-digit numbers, HM provided the following feedback: “Students can use grid paper or lined paper turned sideways to align ones, tens, hundreds, and thousands. Have students compare the digits in each place. Remind them that they must regroup only when the number being subtracted is greater than the number being subtracted from” (Greene et al., 2002, p. 138). MH recommended providing the following type of feedback in an addition lesson on adding four-digit numbers with and without regrouping: “If students forget to write the regrouped number in the place to the left, then have them stop after
adding each pair of digits to see if the sum was greater than 9" (Carlsson et al., 2002, p. 67).

In a similar addition lesson on four-digit numbers in SFAW, teacher feedback was as follows: "Students may try to add beginning at the left. Remind students to begin with adding the ones and check whether regrouping is needed. Have students add 24 + 27 starting with tens to see why this method won't work" (Charles et al., 2000, p. 117). Although feedback was present in 79% of the lessons in SBG, 74% of that feedback was not instructive. For example, although the program provided an error alert, it did not specify the correction procedure. The error alert in one lesson read as follows: "Watch for students who make ten but then incorrectly subtract from the other addend before finding the sum. For examples, 9 + 5 = 10 + 3 rather than 10 + 4" (Fennell et al., 2001, p. 47).

DISCUSSION

This study evaluated the extent to which five 3rd-grade mathematics textbooks adhered to the Standards and instructional design criteria with respect to the teaching of problem-solving. Results of the analysis indicated that there were more variations than similarities within and across textbooks in meeting the Standards. Although opportunities for problem-solving were present most often and were reasonable, the analysis revealed that was not true for the other Standards. In particular, opportunities for reasoning and making connections were present less than half of the time. The exception was SFAW, which provided many opportunities (0.68 word problems per lesson) for applying mathematics in other contexts. When the communication Standard was examined, with the exception of MH, the other textbooks provided sufficient opportunities for articulating own reasoning when solving problems. However, the opportunity for consolidating mathematical thinking by writing about it was low (M = 0.21) across textbooks. In contrast, only HC (0.94) provided adequate opportunities for using the language of mathematics.

The findings for the representation Standard were discrepant across the three criteria. Although all textbooks provided several opportunities to apply text representations to solve problems, they were less likely to have students generate representations or select among representations to solve problems. The exceptions were SFAW (0.53) that required students to generate representations and HM (0.90) for having students select among representations to solve problems. In sum, these results suggest that if teachers are expected to teach to the Standards, publishers need to continue to attend to the Standards when they revise the textbooks. A recent survey study by Maccini and Gagnon (2002) regarding teacher perceptions and application of NCTM Standards indicated that although special education teachers are less familiar with the goals of the Standards and reported relatively less confidence in implementing them with students with learning and behavioral disabilities as compared to general education teachers, the greatest barrier to successful implementation noted by both general education and special education teachers was a lack of adequate materials.

When the textbooks were evaluated with respect to their adherence to the instructional design criteria, results indicated few discrepancies across the textbooks. Although HC and MH satisfied most of the instructional design criteria, less than half of the textbooks we identified met three instructional design features (see Table 2). These design features included clarity of objectives, sufficient teaching examples, and nonexamples. This finding is encouraging, because the textbooks we analyzed seemed to have improved in their quality of instruction when compared to the less than positive results of earlier research evaluating the content of mathematics textbooks (Carnine et al., 1997; Jitendra et al., 1996). Further, this finding is an improvement over a more recent analysis of a subtraction lesson across seven textbooks (Jitendra et al., 1999). In that study, less than half of the programs adequately met four of the nine design features (i.e., clarity of objective, explicit teaching explanations, sufficient and appropriate teaching examples, and effective feedback). In contrast, our analysis indicated that textbooks have improved in terms of providing more explicit explanations using examples that are worked out as well as presenting effective error intervention strategies to correct potential student errors. However, this improvement in teaching
explanations and feedback represents only 60% of the textbooks in this study. In addition, despite our finding that a sufficient number of practice examples existed in the textbooks, these practice examples were part of the independent practice sections of lessons. Consequently, students had relatively few opportunities for scaffolded instruction (i.e., guided practice) before having to solve problems independently.

Results from this study extend previous research on mathematics textbook analysis in that we evaluated both standards-based practices as well as instructional design principles. Although previous studies have offered valuable insights into cross-national differences in mathematics textbooks (Mayer et al., 1995; Saminy & Liu, 1997) and trends in mathematics textbook content before and after the implementation of the Standards (Chandler & Brosnan, 1994), they have not discussed whether textbooks satisfy the goals of the Standards in promoting higher level thinking (e.g., reasoning, connecting). Further, previous studies have examined mathematics content with respect to meeting the design of instruction principles in isolation of the Standards (e.g., Carnine et al., 1997; Jitendra et al., 1996; Jitendra et al., 1999).

At the same time, the findings in this study must be interpreted with caution given the various limitations. First, our sample included only five mathematics textbooks and was limited to third grade textbooks. Although the textbooks we selected were commonly used, we did not review all other textbooks and supplemental materials (e.g., workbooks). In addition, the criteria for the five Standards do not represent the Standards in their entirety. Although the criteria we employed were based on student outcomes associated with the Standards, not every outcome identified by the Standards was evaluated. At best, they are only a partial sampling of the Standards. Moreover, our analysis of the Standards was limited to word problems, whereas evaluation of instructional design criteria focused on problem-solving in general. However, this analysis is more extensive than previous studies (Jitendra et al., 1999; Mayer et al., 1995) in that we evaluated all addition and subtraction lessons in each textbook. Finally, although we focused on dimensions of the lessons related to problem-solving instruction, we did not examine other aspects of these texts such as readability and illustrations.

**Implications for Practice**

How can we improve mathematics and problem-solving instruction? Based on the findings of our study, we offer potential implications for practice.

Given the discrepancies across textbooks in meeting the Standards, we encourage state and local textbook adoption committees to discuss selection of a textbook based on the needs and goals of instruction. For example, if the instructional goal is to build students’ conceptual understanding of operations by way of adequate word problem-solving opportunities, then selecting a textbook that meets this criterion is critical. At the same time, given that most textbooks scored low on the reasoning and connection Standards, these textbooks provide little support for teachers in either of these areas. Therefore, teachers must have not only a deep understanding of their content area to adequately prepare their students to engage in complex thinking and problem-solving, but also be able to accommodate individual differences. For example, they may need to emphasize reasoning and critical thinking, link the newly introduced concept to students’ previous mathematical knowledge, and facilitate generalizable skill application. One way to prepare teachers for such goals is to ensure that professional development programs in higher education institutions provide stronger connections between content knowledge of the discipline with strategic knowledge and systematic teaching that would benefit all learners.

Based on our research results, we suggest that teachers attend closely to instructional design principles and modify or supplement instruction

If the instructional goal is to build students’ conceptual understanding of operations by way of adequate word problem-solving opportunities, then selecting a textbook that meets this criterion is critical.
in mathematics textbooks to meet the needs of individual students. Teachers of students with learning disabilities may need to specify learning objectives that provide a direction for instruction and an evaluation component that documents student performance (Ornstein & Levine, 1993; Stein, Silbert, & Carnine, 1997). Further, given that "true math deficits are specific to mathematical concepts and problem types" (Zentall & Ferkis, 1993, p. 6), we encourage classroom teachers to modify the curriculum as needed to facilitate solid conceptual understanding. This means providing students with carefully designed and explicit instruction that includes sufficient and varied examples and nonexamples (Gersten & Baker, 1998). Creating learning environments in which adequate time is devoted to unambiguous explanations and strategic application of newly learned skills to promote conceptual understanding is an essential goal for teachers. At the same time, providing instructive feedback that allows students to analyze their performance in relation to the feedback provided is critical to promote skilled error-free performance. In addition, because students with learning difficulties often evidence cognitive disadvantage in attention, organization skills, and working memory (e.g., Gonzalez & Espinel, 1999; Zentall & Ferkis, 1993), it is essential that teachers emphasize the importance of representation techniques to help students understand the problem and organize information, as well as to guide them through the problem-solving process.

In response to the No Child Left Behind Act, we encourage publishers to continue to work on improving the quality of educational materials. If not, the onus for evaluating and modifying instruction remains with the classroom teacher to ensure that all students are successful learners. Modifications may be as simple as clarifying objectives or involving more challenging practices, such as providing opportunities for students to engage in reasoning and connecting mathematical thinking. In sum, we recommend stimulating discussions among education professionals (administrators, practitioners, teacher educators, and researchers), parents, and policymakers about using research-proven practices to improve problem-solving instruction and develop students' mathematical proficiency.

In this article, we discussed the importance of a well-designed public curriculum (commercially prepared teaching materials). However, curriculum entails much more than this limited definition and description (Gehrke, Knapp, & Sirotnik, 1992). The notion of an enacted or implemented curriculum (i.e., what happens in the classroom) may play an even greater role in facilitating student achievement. According to Hafner (1993), "the implemented curriculum refers to how the teacher presents the curriculum, what subject matter is taught, the teacher's belief systems and intentions, together with the context in which instruction occurs" (p. 72). In fact, the "implemented curriculum" may prove to be more important in understanding classroom processes (e.g., the interaction between teacher, student, and materials) than the intended or public curriculum (Ball & Cohen, 1996). Understanding what teachers know about the content they teach and how this content is communicated to students is critical. Therefore, future research in this area should validate and extend the findings from this textbook analysis by conducting actual observations of teachers implementing the curriculum.

REFERENCES


In response to the No Child Left Behind Act, we encourage publishers to continue to work on improving the quality of educational materials.


**ABOUT THE AUTHORS**

ASHA K. JITENDRA, Professor, College of Education, Department of Education and Human Services, Lehigh University, Bethlehem, Pennsylvania. CYNTHIA GRIFFIN, Associate Professor, Department of Special Education, University of Florida, Gainesville. ANDRIA DEATLINE-BUCHMAN, Doctoral Student, Department of Education and Human Services, Lehigh University, Bethlehem, Pennsylvania. CAROLINE DIPPI-HOY, Manager of Off-Campus Work Sites, Easton Area School District, Easton, PA. EDWARD SCZESNIAK, Doctoral Student; and NATALIE G. SOKOL, Doctoral Student, Department of Education and Human Services, Lehigh University, Bethlehem, Pennsylvania. YAN PING XIN, Assistant Professor, Department of Educational Studies, Purdue University, West Lafayette, Indiana.

Correspondence regarding this article should be addressed to Asha K. Jitendra, College of Education, Lehigh University, 111 Research Drive, Bethlehem PA 18015. (e-mail: akj2@lehigh.edu)

This research was supported by Grant No. H324D010024 from the U.S. Department of Education, Office of Special Education Programs awarded to the first and second authors.

Manuscript received March 2003; accepted May 2003.

---

The Summit School
Serving Bright Students with Dyslexia and Other Learning Differences
Grades 1-8

The Summit School offers:
- Evidence based practices in reading, math and written language
- An overall student/teacher ratio of 4:1
- Educational, tutorial, consultative and evaluative services

Edgewater, MD
410/269-0070
www.thesummitschool.org