DOCUMENT RESUME

ED 425 910		SE 061 852

AUTHOR	Wiest, Lynda R.
TITLE	The Role of Fantasy and Real-World Problem Contexts in Fourth- and Sixth-Grade Students' Mathematical Problem Solving.
PUB DATE	1998-04-17
NOTE	28p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Fantasy; Grade 4; Grade 6; Intermediate Grades; *Mathematics Education; *Problem Solving; *Word Problems (Mathematics)
IDENTIFIERS	*Authenticity

ABSTRACT

This study was an investigation of the influence of word-problem context--the nonmathematical, verbal aspect of a problem--on fourth- and sixth-grade students' preferences for and problem solving performance on word problems. Problems were parallel in mathematical structure but differed by major context category--fantasy and real-world. Subcategories were low fantasy, high fantasy, children's real-world, and adults' real world. Students' responses to problems varied more markedly by individual problem than by context category. Problem-solving scores in general were somewhat low. Girls and boys fared equally well, but students in a small town scored significantly lower than those from a small city. (Contains 64 references and 8 tables.) (Author)

The Role of Fantasy and Real-World Problem Contexts in Fourth- and Sixth-Grade Students' Mathematical Problem Solving

Lynda R. Wiest

University of Nevada, Reno

Curriculum and Instruction/282
College of Education
University of Nevada
Reno, NV 89557-0214

Phone: 775-784-4961 x2022
Fax: 775-327-5220
E-mail: wiest@unr.edu

ABSTRACT

This study was an investigation of the influence of word-problem context—the nonmathematical, verbal aspect of a problem—on fourth- and sixth-grade students' preferences for and problem-solving performance on word problems. Problems were parallel in mathematical structure but differed by major context category—fantasy and real-world. Subcategories were *low fantasy, high fantasy*, *children's real-world*, and *adults' real-world*. Students' response to problems varied more markedly by individual problem than by context category. Problem-solving scores in general were somewhat low. Girls and boys fared equally well, but students in a small town scored significantly lower than those from a small city.

Paper presented at the Annual Meeting of the American Educational Research Association, April 13-17, 1998, San Diego, CA

The Role of Fantasy and Real-World Problem Contexts in Fourth- and Sixth-Grade Students' Mathematical Problem Solving

Lynda R. Wiest

Word problems are often used in an effort to apply classroom learning to the "real world." Few would disagree that "how children transfer knowledge between school and the outside world may be the central problem in education" (Baranes, Perry, & Stigler, 1989, p. 287). Standard word problems, however, have not lived up to their promise. They do not serve as devices for thinking and may in fact inhibit it, leading Reusser (1988) to this unfortunate finding: "The major result observed in most of our studies is the extent to which textbook problem solving contexts can impair the quality of comprehension" (p. 334).

Because word problems are likely to endure as school mathematics tasks and because their contexts—verbal aspects of the problem statement, such as the story line—have much to do with the problem-solving experience, it is worthwhile to seek to improve word-problem contexts in ways that maximize their potential for fostering thinking, which is "at the heart of mathematical problem solving" (Willoughby, 1983, p. 58). Further, because various student subgroups have been shown to react differentially to problems based on their context, studying and improving word-problem context relates to equity in mathematics education.

One problem context noticeably absent from present-day textbooks is fantasy, even though it is basic to human thinking at all age levels, particularly that of children, and despite its value for generating interest and fostering creativity. The specific topic of this research, therefore, is how solving fantasy versus real-world word problems influences: (a) children's preferences for solving various problems, and (b) children's problem-solving performance.

RELATED LITERATURE

Effects of Word-Problem Context on Students' Problem Solving

Word problems are "notoriously difficult to solve" (Cummins, Kintsch, Reusser, & Weimer, 1988, p. 405), particularly nonroutine and multistep types (López & Sullivan, 1991; Murphy & Ross, 1990). One aspect of word problems some researchers have found to impact student performance is problem context, or "nonmathematical meanings present in the problem statement" (Kulm, 1984, p. 17), for example, the "story" in which the mathematics problem is set. Problem context may help problem solvers give meaning to the mathematical content in a problem and it is likely to influence, in particular, the problem-solving stages of understanding a problem and planning its solution (Boaler, 1993a; Kulm, 1984). Bickmore-Brand (1990/1993) says context is foundational to mathematical activity: "Context is paramount to the construction of meaning the whole way through. It is the backdrop against which the parts have to make sense" (p. 3).

Problem context of interest in this study was the type in which major aspects of word problems are held constant while the story idea is varied. This review mainly includes research involving students in the elementary and middle grades, because it subsumes the grade levels of interest for the research reported in this article and because of the importance of these years for maintaining student interest in mathematics. Most existing research on problem context centers about the middle grades. The majority of studies reviewed here occurred at the fourth- through eighth-grade levels, with sixth grade represented most often.

One way problem context might influence problem-solving performance is through the degree of interest and hence motivation it sparks in solvers, perhaps inspiring them to engage to a greater degree in the problem and to persevere longer in solution attempts (cf. Boaler, 1993b, 1994; Murphy & Ross, 1990). Thus, affective variables, in addition to problem-solving success measured by correctness of solutions and answers, are considered.

Hembree (1992) conducted a meta-analysis of 44 studies, involving fourth-grade through undergraduate students, in which word-problem context was varied while mathematical structure was held constant. Better performance was most strongly associated with familiar contexts.

Concrete (vs. abstract) and imaginative (vs. ordinary) problems, the latter using fantasy or unusual circumstances, showed borderline significance in their impact on problem-solving performance.

Numerous studies have investigated the impact of personalizing problems—inserting students' names or information from their background experiences into the problems they solve—on student interest and problem-solving success. All studies found positive effects on the measured variable of interest or achievement (e.g., Davis-Dorsey, Ross, & Morrison, 1991; Hart, 1996; Kloosterman, 1992; López & Sullivan, 1991, 1992). López and Sullivan (1992) found significant differences favoring personalization on problem-solving scores for two-step but not for one-step problems (although students also scored higher on the latter) in comparison with nonpersonalized problems. They say personalization may be particularly important for more demanding (e.g., unfamiliar or mathematically complex) cognitive tasks. Personalized word problems may be more motivational, meaningful, concrete, and familiar (Davis-Dorsey et al., 1991; López & Sullivan, 1992). Group personalization—using dominant interests of a group of students—has also been shown to benefit students' problem-solving scores compared with nonpersonalization (López & Sullivan, 1992), but individual personalization is more effective in impacting students' attitudes and preferences (López & Sullivan, 1992; Murphy and Ross, 1990).

Some research has centered about using various topics, such as animals, sports, or fashion, for word problems. Problem-solving response to these is believed to relate to students' interests and background knowledge. Boaler (1994) found gender-related differences in achievement that appeared to relate to problem-context topic. Renninger (1992) concluded from her work with fifth- and sixth-grade students that individually identified interests and noninterests embedded in contexts influence the way students engage in and complete tasks (e.g., in terms of entry into and perseverence on a task), although this was much less the case with better problem solvers.

Students have completed problem-solving tasks at the computer for some research on manipulated problem context. This has been the case for most of the very few studies that claim to have used fantasy contexts in their investigations. However, these studies do not merely compare differing verbal contexts of problems. For example, the problems may include visual displays, or they might compare problems with and without contexts, such as constructing a simple abstract drawing versus imagining the task as a search for buried treasure (e.g., Parker & Lepper, 1992). Results for these studies might relate to other factors, such as level of abstraction of the task. Technology use itself can have a confounding effect upon study results, particularly by user gender (Mark, 1992; Middleton, Littlefield, & Lehrer, 1992).

Boaler (1994) says students are tied to the situation-specificity of word-problem contexts and are therefore unable to see mathematical connections among them. She attributes this, in part, to the types of contexts used in mathematics classrooms, which tend to be superficial in meaning and to cue students to look for and apply a "correct" procedure. She (1993a) suggests, "If students can learn mathematics in such a way that enables them to see the underlying similarities between questions set in different contexts they will probably also develop enhanced capability in transferring their school learned mathematics to 'real world' situations" (p. 342). Boaler's (1993a) study involving 12- and 13-year-olds in two schools with different instructional orientations provides support for her contentions. Students in the more process-oriented school tended to use similar procedures across word problems of different contexts, whereas those in the more traditional mathematics classroom were more responsive to context, using varied procedures with problems of similar mathematical structure but differing contexts.

Some research shows that word-problem context has differential impact on certain groups of students in terms of interest and achievement. Sex-stereotyped contexts favoring males (e.g., science themes) have been offered as one possible factor in female underachievement in solving word problems (Jones & Smart, 1995; Murphy & Ross, 1990), although sex-stereotyped contexts favoring females also have been associated with harming females' achievement by engaging them more fully in familiar or interesting contexts (see Boaler, 1993a, 1994). Some researchers and theorists, therefore, consider sex-stereotyped contexts to be a source of test bias that must be considered in designing assessment tasks (Boaler, 1993a, 1994; Chipman et al., 1991). In addition to gender, differential impact of word-problem context has also been suggested in student subgroups by age (Davis-Dorsey et al., 1991; Parker & Lepper, 1992), ability (Marshall, 1995; Renninger, 1992), and community type (Murphy & Ross, 1990).

Use of Fantasy Contexts in School Mathematics

Many children like fantasy themes and story elements (Gillespie & Naden, 1994; Huck, Hepler, & Hickman, 1993). Nevertheless, in recent years the genre is hard to find in mathematics textbooks. I surveyed individuals associated with textbook authoring and publishing, mainly elementary mathematics textbook authors. Seventeen people—approximately two-thirds of those contacted—responded to some or all of my questions. Of the 14 respondents who commented on genres currently used for elementary mathematics textbook word-problem contexts, all 14 stressed use of realistic themes and real-world data. One school textbook author said, "The themes that are avoided are those that are 'unrealistic.' It is the age of realism: no humor, no fantasy, no fiction."

Survey respondents offered two main reasons for the trend away from fantasy and toward real-world problem contexts. First, they assigned the Curriculum and Evaluation Standards for School Mathematics (NCTM, 1989) a key role in the move toward realism. "Real-world" mathematics is continually mentioned in the landmark document, although the term is not defined clearly. Second, concern about parental objection looms large. Most such objection is religious in nature and attacks printed verbiage throughout the school curriculum. In general, fundamentalist religious groups contend that opening the mind to imagination may allow all sorts of ideas to rush in and create a potential for questioning God and being distracted from His Word (DelFattore, 1992). Publishers tend to avoid controversial topics (Webb, 1995).

Proposed benefits associated with the fantasy genre are many. These include instrinsic motivation (Lepper & Hodell, 1989), mental and emotional health (Cullinan & Galda, 1994; Tunnell, 1994), and moral, social, and intellectual development (Cullinan & Galda, 1994; Stewig, 1988). One vital consideration for including fantasy in people's lives is its probable impact upon developing creativity, imagination, and problem-solving abilities (Cullinan & Galda, 1994; Huck et al., 1993; Tunnell, 1994). "Fantasy helps to develop imagination. The ability to imagine, to conceive of alternative ways of life, to entertain new ideas, to create strange new worlds, to dream dreams are all skills vital to the survival of humankind" (Huck et al., 1993, pp. 394-395). Many authors have noted the importance of fantasy and imagination to invention and problem solving in mathematics and science (Prawat, 1993; Shepard, 1988; Tunnell, 1994).

RESEARCH GOAL

The purpose of this research was to investigate the effects of fantasy versus real-world problem contexts upon children's mathematical problem solving, specifically, word problems. The goal was to determine how these problem contexts influence students' preferences for and problem-solving performance on word problems. Because some studies have found gender differences in working with word problems with different problem contexts, I also looked for differential female-male response in the research data.

RESEARCH METHODS

Sample

Students in twelve classrooms—six at each of grades four and six—in a Midwestern state participated in the study. Six classrooms were located in the only elementary school in a small town (Ashland) and six were in three (of thirteen) elementary schools in a small city (Jonesburg) about 50 miles away.[1] Ashland is a lower-middle-class area with students of below-average to average ability, and the Jonesburg schools draw from lower-to upper-middle-class areas and house students of below- to above-average ability. Students in the four schools are predominantly white. Interviews with the twelve classroom teachers indicate that Jonesburg teachers had students work on word problems more often than Ashland teachers. The total number of students in the twelve classrooms was 308; however, 35 did not participate in the study or were excluded from data

[1]Town names are pseudonyms.

analysis for various reasons, such as having very low mathematics or reading ability. Therefore, there were 273 participants, with grade, gender, and town distributions as shown in Table 1.

Design and Procedures

Word-problem structure. Word problems were constructed to be parallel in mathematical structure but different in story content, or context. "Mathematical structure" means the nature of the mathematical relationships expressed in a problem, or situation. The mathematical structure of the problems in this study is one in which smaller groups (subsets) combine to form a larger group (superset), a situation that has been variously termed "Group" (Marshall, 1995), "part-part-whole" (Carpenter & Moser, 1982), and "combine/subset-unknown" (Riley, Greeno, & Heller, 1983). The problems were two-step translation problems that contained extraneous information. They each had three subsets, one of which was unknown, and a superset. In abstract form, the problems have the following general structure, where italicized terms represent numerals:

> There are a (superset name) and b (extraneous set name). C are (subset name 1), d are (subset name 2), and unknown quantity are (subset name 3). How many (superset name) are (subset name 3)?

Context categories. The two literary genres used for story content—fantasy and real-world—were subdivided into two categories each. The resulting context categories were low fantasy, high fantasy, children's real-world, and adults' real-world. Six problems were created in each of these four categories for a total of 24 problems. Six adults sorted the 24 problems, mixed randomly on cards, into the four categories, for which definitions were provided. Their feedback was used to adjust category definitions and problem wording, following which ten other adults completed the same exercise to establish intercoder reliability for problem categorization. All ten categorized fourteen of the problems according to the intended classifications, and eight or nine did so with another five problems. Seven people agreed with four other problem placements, and six people confirmed the remaining problem (for which, interestingly, four people had the erroneous belief that dental braces could pick up radio waves, thus classifying it as a real-world problem). Most difficulty classifying the story ideas was in distinguishing between the two fantasy categories. A final coder, a specialist in children's literature, concurred 100% with the context-category definitions and the intended classification of all 24 problems.

For this research, fantasy was defined as "any departure from consensus reality" [italics in original] (Hume, 1984, p. 21), as one might expect the vast majority of people across cultures to define such a departure. "It contains some element not found in the natural world" (Cullinan & Galda, 1994, p. 199).

Low fantasy involves an otherwise normal world that contains an unexplained, nonrational aspect or event. Sample problem:

> One day in December, 131 people shopping in a toy store were stunned when 157 toys on the shelves suddenly came to life. 49 of the toys danced around the store, 46 of them chatted with each other, and the others sang the song "Toyland." How many toys in the store sang when they came to life?

High fantasy involves an imaginary, but internally consistent, secondary world that often includes mythical or unreal characters or creatures. Sample problem:

> The Secret Forest has 159 redwood trees and is the home of 134 animals that like to keep to themselves. Of the animals that live there, 37 are unicorns, 54 are fire-breathing dragons, and the others are horses with wings. How many animals living in The Secret Forest are horses with wings?

Real-world contexts are those that are "everyday," or that life as it is presumed to be experienced, or potentially experienced, by most people.

Children's real-world contexts involve everyday situations presumed to be experienced or potentially experienced by most fourth- and sixth-grade students. Sample problem:

> Jordan and Shonda counted 141 dollar bills and 158 coins that they had earned from collecting cans all year and selling them to a recycling center. Of the coins they had, 38 were quarters, 59 were dimes, and the rest were nickels. How many of Jordan and Shonda's coins were nickels?

Adults' real-world contexts involve everyday situations presumed to be experienced or potentially experienced by most adults. The theme or topic of the actual story line is presumed to be one that is—for the most part—of greater concern, interest, familiarity, or relevance to adults. Sample problem:

> 159 people visited the Westfield Art Show, held one weekend in August, to see the 136 works of art for sale. 58 works of art were sold on Saturday, 33 on Sunday, and the others were not sold, so they were stored for another art show. How many works of art were not sold at the Westfield Art Show?

Word-problem construction. Most story topics and elements used for the fantasy and children's real-world problems were drawn from literature on children's interests. Efforts were made to choose story content of fairly equal interest to girls and boys. Topics for the adults' real-world contexts were gleaned from fourth- and sixth-grade mathematics textbooks, chosen because they were deemed to be of greater interest or relevance to adults than to children.

To keep the problems as parallel as possible mathematically, numerous aspects were controlled. The first sentence of each problem contains the superset number and an extraneous number, the extraneous number appearing first in half of the problems in each context category. Verbal labels for the superset and extraneous numbers differ (e.g., "toys" and "people"). The second sentence begins with the two known subsets and ends with the unknown subset. The third—and shortest—sentence poses a question soliciting the quantity of the unknown subset, and it includes the verbal label that should accompany the numerical answer. The number of words in the problems ranges from 50 to 60 and the number of characters from 283 to 309. The numbers in each problem range from 131-159 in the first and 33-59 in the second sentence. Answers range from 42-67. These quantitative factors were believed to be held within a reasonably narrow range. Use of grade-level-specific terms, such as "fourth graders," was adapted to match students working with the problems. Attempts to minimize gender bias included choosing story topics believed to interest both girls and boys and using the same number of female and male names in each problem that contained people's names, equalizing their order of appearance across problems.

Flesch-Kincaid readability scores ranged from grade levels of 5.2 to 8.8. Readability scores are higher and represent a broader range than had originally been intended. This resulted from the mathematical structure used (e.g., a somewhat lengthy and complex second sentence) and efforts to create problems with fuller stories than those with lower readability scores. However, many researchers question use of standard readability measures—which usually consider word and sentence length—for gauging comprehension of text (cf. Shelby, 1992). Further, the limitations of these measures for determining readability of mathematical text are particularly great (Paul, Nibbelink, & Hoover, 1986; Shuard & Rothery, 1984). Standard readability formulas cannot account for such factors as student background knowledge and degree of interest in a text. Nevertheless, the issue of readability was addressed in several ways. Feedback from five fourth- and sixth-grade teachers who read the problems, as well as information from a pilot study, was used to adjust wording and change vocabulary in some problems. Pronunciation and meaning of selected vocabulary was discussed with students before they completed forms, and reading and comprehension help was provided during task completion. Finally, as noted earlier, students with low reading ability either did not participate in the study or their data were excluded from analysis.

Study phases. After a pilot study to refine materials and procedures, three study phases—followed by individual interviews—were conducted near the end of a traditional school year. Students completed the first three phases individually in a whole-class setting, each phase occurring one week after the previous. Students participating in interviews did so privately with the researcher. In Phase 1 students indicated their preferences, using a five-item, Likert-type scale ("really dislike," "dislike," "okay," "like," "really like"), for nonquantitative "story ideas" that formed the basis for word problems solved in Phase 3. For example, the children's real-world problem shown earlier appeared in this form:

> Many children rushed to their favorite ride at the yearly carnival held by the business owners in their town. Some children hurried to the ferris wheel, some went to the merry-go-round, and the rest chose the roller coaster.

Students used a similar scale in Phase 2 ("very bad," "bad," "okay," "good," "very good") to rate the research problems in terms of their preferences for solving the problems, without actually solving them. In Phase 3, students were asked to solve five problems. Four problems were research problems of interest, one in each of the four context categories. The fifth, located third in the problem sequence, appeared on the surface to be similar to the others but called for a different solution method (one-step subtraction). Students then used the same rating scale as in Phase 2 to indicate how they felt about working the problem. In each of the three phases, a comment space was available for students to explain their ratings, which most did.

Each student received 12 of the 24 story ideas in Phase 1 and 8 of the 24 problems in Phase 2 to rate so that she or he would work with a manageable number of items. To minimize the effects of interest in individual problems, only the three most highly rated story ideas in each of the four categories in Phase 1—by grade level—were used for problems solved in Phase 3. Therefore, Phase 3 included 12 of the original 24 problems, with each student solving four. In each phase, different forms were created that included different story ideas or problems so that all research items were addressed by a portion of the sample. Because the order in which a problem is presented within a problem set has been shown to affect problem-solving performance (Zollman, 1986/1987), the order of fantasy and real-world problems, as well as placement of individual items, varied across forms. An equal number of each form created for a phase was randomly distributed within each classroom.

Finally, 31 students who had participated in the three phases were interviewed. An attempt was made to include a diverse group of students in terms of grade level, gender, town, ability, and personal interests (as discerned from forms completed earlier). The subsample interviewed was not as representative of the study sample as intended because of the constraint imposed by including only those who returned consent forms, but it still was fairly diverse. The group consisted of 17 girls (nine Ashland, eight Jonesburg) and 14 boys (six Ashland, eight Jonesburg). These included 15 fourth graders and 16 sixth graders. Overall, these students' abilities were weighted more toward average and above-average than was representative of the study sample. Interviewed students solved and rated problems as in Phase 3, while explaining their thinking, and they answered questions concerning the influence of problem context. Sessions were audiotaped.

<u>Data Analysis</u>

Ratings for items in Phases 1 through 3 were assigned a score from 1 to 5, 1 being the most negative of the five response choices ("really dislike" or "very bad") and 5 being the most positive ("really like" or "very good"). Phase 3 problem solutions were scored as 0 points—inappropriate solution plan, and 1 point—appropriate solution plan. "Appropriate" meant the plan would lead to a correct answer if implemented correctly. To establish intercoder reliability, a mathematics education doctoral student scored one classroom set of problems: four problems for each of 28 students for a total of 112 problems. Our scores matched for 107 problems. After discussing the remaining five items, we agreed on the scores assigned to all 112 problems.

Descriptive statistics were obtained for aggregated data for each item. T-tests for independent samples were run to test for significant differences between problem-solving scores (aggregated for all 12 problems) for the two grade levels and also within each grade level between female and male students and Ashland and Jonesburg students. Written comments were stored—verbatim—in word-processed form for each item, and these were separated by grade and within grade by gender. Interviews were transcribed verbatim from the audiotapes. After numerous readings of the verbal data to establish and adjust categories, categories of responses that occurred naturally with some frequency were determined.

Table 2 shows the range in number of students who completed items in each study phase.

Insert Table 2 about here.

RESULTS

Phase 1 Ratings: Preferences for Story Ideas

Table 3 shows Phase 1 summary ratings by grade level for the six story ideas in each context category. Fourth graders rated items more highly than sixth graders. For both grade levels, preferences for low fantasy, high fantasy, and children's real-world story ideas were fairly comparable, with lower preferences expressed for adults' real-world themes. Standard deviations (which are not included in the table) indicated that students' ratings at both grade levels varied more for fantasy than for real-world story ideas.

Insert Table 3 about here.

Phase 2 Ratings: Pre-Solution Preferences for Solving Problems

Ratings of word problems in Phase 2, summarized in Table 4, showed no strong pattern of change—for either grade level—from Phase 1 ratings for low fantasy, high fantasy, and children's real-world problems. However, both fourth and sixth graders rated all six adults' real-world problems more highly. Sixth graders' ratings of the children's real-world problems also increased somewhat.

Insert Table 4 about here.

Phase 3: Problem-Solving Scores and Post-Solution Ratings

Table 5 shows Phase 3 problem-solving scores, by grade, aggregated across all twelve problems solved. Overall, about three-fifths of fourth graders and three-quarters of sixth graders solved the problems "correctly," that is, earned 1 point by using a plan that would lead to a correct answer if implemented properly. Prominent differences in problem-solving scores exist between grade levels and between towns, particularly in grade four for the latter, whereas female-male scores were not very discrepant at either grade level. One-way analysis of variance shows the differences to be significant at the $p < .0001$ level for grade level, $p < .0001$ for town at grade four, and $p < .001$ for town at grade six. No significant differences exist for gender at either grade level. Girls' scores improved more than boys' from grade four to grade six (.23 of a point compared with .13 of a point), as did Ashland compared with Jonesburg students' scores (.26 of a point increase for the former, .12 of a point for the latter).

Insert Table 5 about here.

Problem-solving scores (see Table 6) show little distinction by context category. Considerable variation exists among individual problems within and across categories (see Table 7), particularly at the fourth-grade level, where scores range by .42 of a point compared with .32 at the sixth-grade level.

Insert Tables 6 and 7 about here.

Gender differences in scores for the same items are not large in most cases. However, the discrepancy is quite large for a few items, exceeding .20 of a point for the three problems below.

Favoring males (both grade four)

> The Secret Forest has 159 redwood trees and is the home of 134 animals that like to keep to themselves. Of the animals that live there, 37 are unicorns, 54 are fire-breathing dragons, and the others are horses with wings. How many animals living in The Secret Forest are horses with wings?
>
> One day 134 Hocus-Pocus Club members cast 138 magic spells before the Good Fairy

changed things back again. 35 of the spells turned children's hair into noodles, 38 spells turned adults' noses into bananas, and the other spells turned babies into piglets. How many magic spells turned babies into piglets?

Favoring females (grade six)

The braces on Pat's teeth, which would be worn for 138 more days, suddenly started picking up a radio program from 136 miles away. Pat listened to music the first 53 days and sports events the next 38, then the braces stopped working. For how many days did Pat's braces no longer pick up the radio program?

More problems with large score differences favored males, which occurred in five of the seven problems with a point spread of at least .15. Six of the most discrepant seven pairs of scores between females and males occurred in grade four. Children's real-world items, which received the most consistent ratings of the four context categories across the three study phases, also showed the fewest differences overall—of the four context categories—between girls' and boys' problem-solving scores on individual problems.

All large differences between Ashland and Jonesburg students' problem-solving scores for individual problems favored Jonesburg students. Of the seven most extreme examples, four problems are at the sixth-grade level (differences of .22 or greater) and three are at the fourth-grade level (differences of .34 or greater). High fantasy problems are strongly overrepresented, accounting for four of the seven problems. Despite Jonesburg students' significantly better problem-solving scores at both grade levels, in three cases Ashland students outscored Jonesburg students at the sixth-grade level, by only .02 and .03 of a point for two problems but by .14 of a point for the third.

Students were more likely in solving certain problems to use the extraneous number instead of the superset number. Choosing the extraneous number ranged from no use for some problems to use in about one-seventh of the solution plans for others. Predominant type of appropriate solution plan used for a problem varied. These choices did not fall along fantasy and real-world lines. An add-subtract approach was most popular, in which students first added the known subset numbers before subtracting the sum from the superset number. Students also used a subtract-subtract approach, in which they subtracted one subset number at a time from the superset number. Some individuals used different solution plans among the four problems of interest they solved in Phase 3. On average across all problems, grade four students used the add-subtract method for 68% and the subtract-subtract method for 32% of the combined total of these two predominant methods; grade six figures are 85% add-subtract and 15% subtract-subtract. However, fourth graders used the add-subtract method 100% of the time for one problem and 92% for another and the subtract-subtract method more than 50% of the time for two problems. Sixth graders showed similar variability in their choice of solution plans for different problems.

Analysis of students' solution methods and written comments for problems combined with observational and interview data revealed factors that might contribute to problem-solving response: interest; readability; familiarity and complexity of concepts; distinctness of meaning of superset and extraneous set names; size, physical proximity, and visual similarity of numerals; degree of associability of subsets with each other and with the superset. These elements might be categorized into four major areas: readability, verbal structure, story concepts, and personal factors. Context elements did not stand alone to influence students' problem solving, but they seemed to work in combination to contribute to response.

Post-solution ratings of problems tended to vary less and were somewhat higher than ratings in earlier phases. Means for all problems in the adults' real-world category for grade four and children's and adults' real-world categories for grade six increased from Phase 2 to Phase 3.

Relation of Story-Idea Preference Ratings to Problem-Solving Scores

Pearson correlations between Phase 1 story-idea preference ratings and Phase 3 problem-solving scores show that only one of the 24 problems solved in Phase 3 (12 at each grade level) significantly correlates with its Phase 1 preference rating, an occurrence that might be attributed to chance. A comparison of group means rather than individual pairs of scores also indicates that little consistency exists between Phase 1 preference ratings and Phase 3 problem-solving scores, and that no patterns emerge by context category. See, for example, Table 8, which uses rank-ordered data to show group comparisons for grade six. Similar but more exaggerated patterns appear for grade four, with rankings for the two types of data differing by two to ten ordinal positions.

Insert Table 8 about here.

Phase 1 story-idea preference ratings do not associate with Phase 3 problem-solving scores in any sex- or community-typed pattern. For example, although sixth-grade boys rated one problem as .40 of a point (on a five-point scale) more interesting than girls, girls scored much higher solving the problem (.86 of a point compared with .63 of a point on a one-point scale). Similar examples occurred by town.

Influence of Readability on Problem Solving

Group data for the twelve problems each grade level solved showed no consistency in ordinal position between readability and problem-solving scores, which would be expected to be inversely related. For example, one problem solved by fourth graders had the lowest readability level (5.8) and the highest mean problem-solving score (.83), but the problem with the highest readability level (8.8) had the third highest problem-solving score (.72). Patterns also did not emerge by context category, nor in relation to student subsamples by gender and town.

Insight into readability of these research problems was gained by analyzing students' written and oral comments and by listening to students read problems aloud. Most pronunciation difficulties appeared in reading names, such as Kline, Chen, and Shonda. These difficulties seemed to relate to unfamiliarity, which also affected gender attributions. For example, a sixth-grade girl called Jamal—intended to be a male—"she." Word-pronunciation difficulties did not seem to hinder problem-solving performance as measured in interviews. Two readability issues—understanding of story-content concepts and complexity of wording—appeared to be more important to problem-solving success. For example, many students did not know the meaning of "write-in vote," even though I explained the term to students before each study phase. Students asked to explain the term in interviews described it incorrectly.

Influence of Word-Problem Context on Problem Solving

Many students' written and oral comments illustrate their belief that word-problem context affects their interest in and manner of solving a problem. Numerous students wrote that they had liked or would like solving particular problems because the story was interesting. One factor they say makes stories enjoyable is humor. Some students claimed to try harder or "think worse" based on their like or dislike for problem contexts. Unpredictability of content also had an impact, as one student said a problem "made him think" because he was expecting a different ending. Some students maintained particular problems were harder or easier than others on a form, implying context effects, given the parallel nature of mathematical relationships in the problems.

Three-quarters of interviewed students thought different stories (one liked, one disliked) accompanying the same mathematics problem would affect the way they solved it. This included three of the six sixth-grade males, making them less likely than sixth-grade females and fourth graders to believe they would be influenced by problem context. Half of the students who believed problem context would influence them said they would exert differential effort on the problems. Almost half said they would perceive the difficulty level of the problems differently. About one-fourth said contexts of differing interest would affect their desire to do the problems, their success in solving them, their attentiveness, or how much they like the problems or like solving them. One or two students added—for each of the following context effects—that preferred contexts would encourage them to look back at or "check" problems, to develop greater interest in mathematics and in doing more problems, and to "get more out of" the solved problem, whereas disliked contexts might distract or confuse them. Students who said different contexts would not affect them believed they can or typically do ignore the story in getting information needed to solve a problem. Two students' oral comments demonstrate their beliefs about the effects of story content:

> You would, you'd like, "Aw, this [disliked problem] is dull. It's boring. I can't do it." It's gonna be <u>harder</u>, because it's just boring. But once you get into, like, fantasy, and it's really

> fantasy, and it really jumps at ya, you're gonna wanna solve it, instead of thinking about how big it is, how much the numbers are, and if it's too hard. You're gonna wanna think that, well, "Wow, I wanna read this. I wanna try to get, solve this, see what happens." (grade four female)
>
> If there's just, like, fairy tale or something [disliked context], I just do it. If it's like, um, like, science fiction, I like to read it, too. . . . If it's a fairy tale . . . I don't try to read it all. . . . I just skim it, 'n get what I suppos--, that's what we're supposed to do. But if it was science fiction or something, I would read it. (grade six male)

DISCUSSION

Phase 1 Ratings: Preferences for Story Ideas

Students expressed differential preferences for various story ideas or verbal contexts. These data indicate that students showed roughly equivalent preferences that tended toward the "like" side of the rating scale for low fantasy, high fantasy, and children's real-world story ideas. Preferences for adults' real-world story ideas contrast with the other three types, falling somewhat on the "dislike" side of the rating scale, perhaps due to use of topics that are unfamiliar, irrelevant, or simply uninteresting to children. Conversely, the appeal of children's real-world story ideas might be based to some degree on familiarity, relevance, and interest. Interest in the two fantasy contexts might be based on familiarity for some students through fantasy being common to children's manner of thinking and to the large body of fantasy literature, movies, video games, and so forth geared toward children. The novelty aspect of fantasy also evoked interest in many children. However, some students' comments revealed that novelty also can be a disliked element.

Additional research—for example, that which varies only one factor at a time—is necessary to test whether or not differences in expressed preferences measure the actual constructs on which the categories are based, whether they reflect reactions to specific elements within contexts, regardless of general context category, or whether they represent a combination of factors. Written comments indicate that students sometimes based ratings on context categories, evidenced by positive statements (ones accompanied by "like" or "really like" ratings) such as, "Because I like fantasys," "Because it has to do with things that are make-beleive," and "I like it because it could really happen," and negative comments such as, "This is too imaginary," "To fairy talelish," and "It wasn't very realistic."[2] The particularly low ratings of adults' real-world story ideas also suggest a categorical response, that is, a general, consistent response to a category that is mainly independent of individual item content. Somewhat more often, however, students' comments addressed context specifics, suggesting a need for a more in-depth look at individual contexts. Students' preferences, therefore, appear to relate not only to the major category type (e.g., adults' real-world), but also to individual item content (e.g., animals as a topic or collecting cans to earn money as a story line), meaning that the content of individual items used specifically for this research might have influenced results by category. Aggregated data by category, too, can mask the variability of response to individual items, which the data in many cases show to be fairly large among items not only across but also within categories.

It is important to recall that these data provide a picture of a group of students and not of individuals. Even the most highly rated contexts among the 24 that students rated were disliked by a relatively small percentage of each grade-level sample. Therefore, it is important to consider the preferences of individual students and various student subgroups. Data obtained in Phase 1 show that some differences appear by grade level, gender, and the community in which a student lives (corroborating findings by Summers & Lukasevich, 1983).

The greater variability of students' response to fantasy compared with real-world story ideas falls in line with other research evidence about students' tendency either to like or dislike fantasy and to respond to it categorically (Cullinan & Galda, 1994; Huck et al., 1993), whereas interest in real-world contexts is more likely to fall along a continuum with less extremes and to be more

---

[2]Students' written comments appear in their original form without editorial correction.

analytic (related to specific content within contexts). Perhaps real-world contexts are familiar to students because of life experiences or because of their predominant use in school materials, making response to them less subject to individual variation. Exposure to fantasy contexts may be more responsive to personal (e.g., home) background than are real-world contexts.

Phase 2 Ratings: Pre-Solution Preferences for Solving Problems

The only strong categorical differences between Phases 1 and 2 for grades four and six, in each case an increase in mean scores, occur in the two real-world categories (mostly adults' real-world). This is not surprising. Children have been enculturated into often less-than-thoughtful, but sometimes successful, approaches to solving word problems, most of which they expect to have real-world contexts. Students may be more concerned about a problem's perceived solvability than the interestingness of its context. Seventh graders in Lester, Garofalo, and Kroll's (1989) research wrote word problems they considered to be interesting, which the class subsequently rated from 1 (very boring) to 5 (very interesting). They then solved a problem of their choice from those they had rated. In hopes of improved likelihood of getting a right answer, three-quarters chose a problem they had assigned a 1 or 2. The researchers attribute this phenomenon to classroom grading practices. Perhaps students perceive more boring problems, in this research the adults' real-world problems, as being easier because they are less distracting, or because they are what "real" (school) word problems should look like, or because they are more familiar with and have had more experience solving real-world problems, which dominate school mathematics materials. It may also be true that students hold different standards and expectations for stories read for their own sake versus those tied to mathematical problems to be solved. In any case, these perceptions might affect students' problem-solving attitudes and performance.

Most likely, ratings of children's real-world problems did not increase from Phase 1 to Phase 2 as much as the adults' real-world ratings because the former were already fairly high in Phase 1. For neither grade, however, did mean ratings for adults' real-world problems raise to the level of any of the other three categories, hinting that story content still influenced students' response.

Phase 3: Problem-Solving Scores and Post-Solution Ratings

General context category does not appear to influence—at least by itself—problem-solving success as measured by ability to devise an appropriate solution plan, because problem-solving scores within each category show considerable variation. Further, the means and score ranges for each category are comparable. Possible exceptions are the children's and adults' real-world problems at the sixth-grade level, the former exhibiting higher scores and the latter lower scores compared with the two fantasy categories. As noted earlier, story-idea preferences do not correlate with problem-solving scores, negating the sole influence of interest in accounting for these differences. Appeals to familiarity as "the" contributing factor probably cannot be made, either, because a similar discrepancy between children's and adults' real-world problems does not appear in grade four data. One would expect, anyway, that sixth graders would be more familiar with the adult world than fourth graders. The difference in scores between the two real-world categories is interesting for grade six, because sixth graders have had several years' experience solving adults' real-world problems in school mathematics. Perhaps this is the very reason for the lower scores. The sixth graders might have given less attention to overly familiar problem types they assumed they knew how to solve. Or maybe sixth graders, who are somewhat harder to please than fourth graders (as shown by their tendency to assign items lower ratings and by their decreased motivation in school, as reported, for example, in Anderman & Maehr, 1994), have a negative reaction to solving adults' real-world word problems that manifests itself differently than in merely expressing preferences for story ideas. Finally, perhaps the specific adults' real-world problems used in this research were somehow more difficult to solve than problems in the other categories.

Students attained very different scores across the twelve problems they solved, such that individual problem contexts might influence students' problem solving more than categorical ones, or broad general types (see also Webb & Yasui, 1992). Among the twelve solved problems, fourth graders have a score range that exceeds that of sixth graders by .10 of a point, and fourth-grade boys' range of scores exceeds that of fourth-grade girls by about the same. This might mean that fourth graders are more influenced by problem context than sixth graders, and the same might be true for fourth-grade boys compared with fourth-grade girls. That fourth graders, poorer problem solvers than sixth graders, should be so affected is not surprising, as it resonates with research on novice versus expert problem solvers (e.g., Silver, 1987). However, it is uncertain why fourth-grade boys might be more influenced by problem context than girls their age. Girls generally are believed to be more verbally oriented and to be more contextual/global in their thinking (cf. Jones & Smart, 1995), as in their greater tendency to be classified as field-sensitive (e.g., Van Blerkom, 1988). However, perhaps boys' greater tendency toward risk-taking behavior in problem solving allowed them to engage the contexts more fully than problem solvers whose greater focus is on correct procedures, or perhaps they were more affected by the varying language demands than the females, who generally are more proficient than boys in language areas (e.g., reading) in their schooling (cf. Han & Hoover, 1994). Speculations by gender are somewhat difficult to make at this age level. Additional research in this arena would be interesting and worthwhile.

On the surface, the word problems in my study appear to be parallel. However, more detailed analysis of students' comments and solution processes revealed this was not so. Meaning embedded in individual problem contexts influenced success in choosing appropriate solution plans, type of solution plan used, and choice of extraneous versus superset number. These phenomena occurred not only in aggregated group data but also across the four problems individual students solved.

Phase 3 problem-solving scores are disappointing, given that the problems are ones end-of-the-year fourth and sixth graders would be expected to be able to solve. Further, scores only reflect choosing an appropriate plan without penalty for computation, labeling, copying, or other error (although choosing a correct procedure—as Morales, Shute, & Pellegrino, 1985, point out—is more often a source of student error than are these other types of errors). Two-step problems such as those used in this research are particularly problematic for students because of their greater length and complexity due to containing more information to consider, a greater number of operations to identify and to compute correctly in the right order with the right choice of numbers, a surface resemblance to one-step problems (encouraging a tendency to rather hastily identify key words, choose an algorithm, and compute), and reduced familiarity and experience with these types of problems (Kouba et al., 1988; López and Sullivan, 1992; Quintero, 1983; Sherrill, 1983). Inclusion of extraneous information and use of fairly large (or at least "not small") numbers further raise the difficulty level of these problems for students (Hembree, 1992; Kouba et al., 1988). Although these problems are lengthier than those found in typical school textbooks for these age levels, all problems fell within the same fairly narrow word- and character-count range, yet the problem-solving scores were quite variable. Similarly, readability scores did not correlate with problem-solving scores.

The problem-solving scores attained in this research, therefore, support continued concern about students' ability to solve word problems. Heightened concern is warranted for rural/small-town students. This may be an issue of lower socioeconomic status, which has been strongly associated with problem-solving success (e.g., Marshall, 1984). The comparable scores attained by females and males at both grade levels is encouraging and confirms recent evidence regarding gender comparisons at the elementary level (Davis-Dorsey et al., 1991; Hyde, Fennema, & Lamon, 1990). Scores on some individual items, however, are fairly discrepant between certain student subgroups (e.g., females and males), suggesting a need to examine data for various student groupings in addition to doing global analyses, and again implying possible effects of problem context in a more specific sense than that of general context categories. Across the two-grade-level span, problem-solving scores increased more for girls than for boys, as they did for Ashland compared with Jonesburg students. However, this might partially reflect the fact that girls and Ashland students had the lower scores in grade four and thus had "farther to go."

Post-solution ratings of problems probably were higher and had a narrower range than earlier phases because Phase 3 problems were those most highly rated in Phase 1. Also, students might have been overconfident about the correctness of their answers. Relatively higher scores for children's and adults' real-world categories may have occurred for reasons similar to those speculated for Phase 2.

Discrepant Problem-Solving Scores by Gender and Town

Even more difficult to explain than whole-grade score variations among problems are large score discrepancies for the same problems between either females and males or Ashland and Jonesburg students of the same grade level (accounting for ability differences in the latter case). Despite problem-context variables being constant, no single apparent factor surfaces to account consistently for differing degrees of problem-solving success that appear between students of different gender or from different community types for the same problem. Some factors that might differ categorically among the students are their knowledge bases, interests (preferences), and reading abilities, so that familiarity, interest, and readability might be among factors that interact in various ways to impact a subgroup's problem-solving performance. Familiarity of problem contexts can only be speculated as a factor influencing problem-solving success, since it was not controlled or investigated in this study. If interest is a factor, it might act upon students differently for different problems, perhaps at times a distractor and at others an engager, or perhaps interest is difficult to measure in its full complexity and the Phase 1 scores in this study do not serve as true indicators of students' interests.

Five of the seven problems in which females and males attained problem-solving scores that differed by .15 of a point or greater favored males. Four of these might be construed to have content more traditionally associated with males, in terms of interest and familiarity: dragons, magic tricks, politics, and business. Some studies show that girls have greater interest in animals and in biology within the field of science (e.g., Fisher & Ayres, 1990; Summers & Lukasevich, 1983), which might play into—along with a high readability score for the item—their outscoring of boys on a problem about sea animals. Strong differences in interest in animals, however, did not appear in this research on the whole, although familiarity of content remains an issue. Further, I see no apparent reason for the discrepant scores in the remaining two problems. Nevertheless, it is worth considering Chipman et al.'s (1991) finding that females identified stereotypically female topics as being personally familiar, whereas males expressed no differential familiarity with topics along sex-typed lines. The researchers attribute this to males' familiarity not only with masculine topics but also with many feminine ones (e.g., domestic themes) in their life experiences, which does not necessarily work in reverse. If this were so, it would mean that males have a broader knowledge base from which to draw in engaging problem contexts.

The fact that most of the problems with particularly large female-male problem-solving score differences occurred at grade four might mean that context becomes less influential by gender as students grow older and acquire more life and problem-solving experiences (which might again reverse beyond the elementary grades as sex roles become more salient and sex-typed knowledge increasingly divergent). Since children's real-world items received the most consistent ratings across phases and the fewest notable gender differences in problem-solving scores for individual problems, these problem contexts might be given particular consideration for creating more equitable word problems.

One speculation for the larger score discrepancies between Ashland and Jonesburg students, favoring the latter, for high fantasy problems compared with other types is the disparity in socioeconomic base between the two communities. Perhaps Jonesburg students have more access to high fantasy, a type of fantasy frequently featured in popular culture in computer and video games, movies, literature, and so forth. Given the significantly better problem-solving scores of Jonesburg students at both grade levels, it is perplexing that in three cases Ashland students outscored Jonesburg students at the sixth-grade level. For one about music cassettes, a particularly low readability score relative to the other problems might have combined with familiar content to give Ashland students equal footing with Jonesburg students, illustrating the importance of carefully crafted problem context. The other two problems were ones identified as having either confusing wording or content. Perhaps Jonesburg students, as fairly good problem solvers, were overconfident and erroneously recognized the problems at a glance as addition without deeper investigation. Perkins and Salomon (1989) describe how experts solve "typical" problems: "Expert performance entailed (a) a large knowledge base of domain-specific patterns . . .; (b) rapid recognition of situations where these patterns apply; and (c) reasoning that moves from such recognition directly toward a solution by working with the patterns" (p. 18). This point is interesting, because it might mean research findings that poor problem solvers attend to surface (contextual) details of problems whereas good problem solvers look to the deeper mathematical structure (Marshall, 1995; Renninger, 1992) are overgeneralized. These problem-solving behaviors might be conditional, depending on perceived problem difficulty and problem similarity to standard word-problem forms. Perhaps attention to contextual details gave Ashland students an "edge" in the more obscure problems, engaging real meaning without jumping too quickly to a perceived familiar pattern.

CONCLUSIONS

Traditionally, the textbook has been the most frequently used—in some cases, the *only*—instructional material in the classroom (Ornstein, 1992). It is likely to continue to wield great influence upon teaching and learning. I predict that most textbooks and other instructional materials will continue to rely upon word problems as one type of application- and thinking-oriented medium for mathematics learning at all age levels, and standardized and other testing measures will include word problems in their assessments of mathematical knowledge and abilities. Well constructed word problems can have an important place in the mathematics curriculum, provided they are utilized with effective instructional methods. Despite their confinement to an artificial setting, they can serve as sources of mathematical thinking, which is central to transfer of learning, or useability of mathematical knowledge. As Hiebert et al. (1996) and Renninger (1992) say, the critical factor is not so much a problem as the problematizing of it. Kloosterman (1992) found in his research that "students enjoyed the non-routine problems because they were unique and challenging, not because they saw extensive real-life applications in them" (p. 36).

We must continue to investigate the many aspects of word-problem context that impact students' problem-solving performance and mathematics learning, including differential effects upon subpopulations that might be particularly advantaged or disadvantaged by certain types of contexts. Developers of instructional materials, as well as teachers who use these materials and construct their own, stand to benefit much from these research efforts, as do—most importantly—students. However, information about problem-context effects alone is not sufficient. Teaching method is crucial to using word problems successfully as a learning aid. We must continue to find ways to improve instructional use of word problems and to institute effective measures for teacher development in this area.

I suggest that fantasy contexts be included among those we use for word problems in mathematics teaching. Fantasy contexts interest many children, and they lend themselves to abstract and creative thinking. Because they create greater cognitive demands and contribute to increased variety of problem type, fantasy contexts might encourage students to attend to the mathematical structure of problems, thus enhancing ability to make connections among problems of different contexts, or to transfer learning. The variety these contexts add can yield affective as well as cognitive benefits. Even though a preference for fantasy contexts did not result in higher problem-solving scores in this study, these contexts might yield benefits that are not readily identifiable. For example, inclusion of fantasy problems might influence some students to enjoy solving word problems and, therefore, want to do more, or to develop a more favorable attitude toward word problems or mathematics in general. Also, students are not as familiar with solving these problem types as they are with real-world problems, yet they performed equally well on both, which may mean fantasy contexts serve a useful function.

Sixth graders have shown that, given a choice of word problems to solve, they will select problems with a variety of contexts and not merely those demonstrating their expressed interest areas (Morrison, Ross, & Baldwin, 1992). Therefore, we might include variety and allow students some choice in the problems they solve, which would provide alternatives for students who object to or have a low preference for particular problems.

REFERENCES

Anderman, E. M., & Maehr, M. L. (1994). Motivation and schooling in the middle grades. Review of Educational Research, 64, 287-309.

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.

Barrow, R. (1988). Some observations on the concept of imagination. In K. Egan & D. Nadaner (Eds.), Imagination and education (pp. 79-90). New York: Teachers College Press.

Bickmore-Brand, J. (1990/1993). Implications from recent research in language arts for mathematical teaching. In J. Bickmore-Brand (Ed.), Language in mathematics (pp. 1-9). Portsmouth, NH: Heinemann. (Original work published 1990)

Bishop, A. J. (1988). Mathematical enculturation. Boston: Kluwer Academic Publishers.

Boaler, J. (1993a). Encouraging the transfer of 'school' mathematics to the 'real world' through the integration of process and content, context and culture. Educational Studies in Mathematics, 25, 341-373.

Boaler, J. (1993b). The role of contexts in the mathematics classroom: Do they make mathematics more "real"? For the Learning of Mathematics, 13(2), 12-17.

Boaler, J. (1994). When do girls prefer football to fashion? An analysis of female underachievement in relation to 'realistic' mathematic contexts. British Educational Research Journal, 20, 551-564.

Carpenter, T. P., & Moser, J. M. (1982). The development of addition and subtraction problem-solving skills. In T. P. Carpenter & J. M. Moser (Eds.), Addition and subtraction: A cognitive perspective (pp. 9-24). Hillsdale, NJ: Erlbaum.

Cullinan, B. E., & Galda, L. (1994). Literature and the child (3rd ed.). Fort Worth: Harcourt Brace College Publishers.

Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. Cognitive Psychology, 20, 405-438.

Davis-Dorsey, J., Ross, S. M., & Morrison, G. R. (1991). The role of rewording and context personalization in the solving of mathematical word problems. Journal of Educational Psychology, 83, 61-68.

DelFattore, J. (1992). What Johnny shouldn't read: Textbook censorship in America. New Haven: Yale University Press.

Egan, K. (1989). Teaching as story telling: An alternative approach to teaching and curriculum in the elementary school. Chicago: The University of Chicago Press.

Fisher, P. J. L., & Ayres, G. (1990). A comparison of the reading interests of children in England and the United States. Reading Improvement, 27, 111-115.

Gillespie, J. T., & Naden, C. J. (1994). Middleplots 4: A book talk guide for use with readers ages 8-12. New Providence, NJ: Bowker.

Greene, M. (1988). What happened to imagination? In K. Egan & D. Nadaner (Eds.), Imagination and education (pp. 45-55). New York: Teachers College Press.

Han, L., & Hoover, H. D. (1994). Gender differences in achievement test scores [CD-ROM]. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans, LA. Available: SilverPlatter File: ERIC Item: 369 816

Hart, J. M. (1996). The effect of personalized word problems. Teaching Children Mathematics, 2, 504-505.

Hembree, R. (1992). Experiments and relational studies in problem solving: A meta-analysis. Journal for Research in Mathematics Education, 23, 242-273.

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Human, P., Murray, H., Olivier, A., & Wearne, D. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. Educational Researcher, 25(4), 12-21.

Huck, C. S., Hepler, S., & Hickman, J. (1993). Children's literature in the elementary school (5th ed.). Fort Worth: Harcourt Brace Jovanovich College Publishers.

Hume, K. (1984). Fantasy and mimesis. New York: Methuen.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. Psychological Bulletin, 107, 139-155.
Jones, L., & Smart, T. (1995). Confidence and mathematics: a gender issue? Gender and Education, 7, 157-166.
Kaufmann, G., & Helstrup, T. (1985). Mental imagery and problem solving: Implications for the educational process. In A. A. Sheikh & K. S. Sheikh (Eds.), Imagery in education (pp. 113-144). Farmingdale, NY: Baywood.
Kloosterman, P. (1992). Non-routine word problems: One part of a problem-solving program in the elementary school. School Science and Mathematics, 92, 31-37.
Kouba, V. L., Brown, C. A., Carpenter, T. P., Lindquist, M. M., Silver, E. A., & Swafford, J. O. (1988). Results of the Fourth NAEP assessment of mathematics: Number, operations, and word problems. Arithmetic Teacher, 35(8), 14-19.
Kulm, G. (1984). The classification of problem-solving research variables. In G. A. Goldin & C. E. McClintock (Eds.), Task variables in mathematical problem solving (pp. 1-21). Philadelphia: The Franklin Institute Press.
Lepper, M. R., & Hodell, M. (1989). Intrinsic motivation in the classroom. In C. Ames & R. Ames (Eds.), Research on motivation in education: Vol. 3. Goals and cognitions (pp. 73-105). San Diego: Academic Press.
Lester, F. K., Jr., Garofalo, J., & Kroll, D. L. (1989). Self-confidence, interest, beliefs, and metacognition: Key influences on problem-solving behavior. In D. B. McLeod & V. M. Adams (Eds.), Affect and mathematical problem solving: A new perspective (pp. 75-88). New York: Springer-Verlag.
López, C. L., & Sullivan, H. J. (1991). Effects of personalized math instruction for Hispanic students. Contemporary Educational Psychology, 16, 95-100.
López, C. L., & Sullivan, H. J. (1992). Effect of personalization of instructional context on the achievement and attitudes of Hispanic students. Educational Technology Research and Development, 40(4), 5-13.
Mark, J. (1992, June). Beyond equal access: Gender equity in learning with computers. Women's Educational Equity Act Publishing Center Digest [no volume number], pp. 1-2, 6-7.
Marshall, S. P. (1984). Sex differences in children's mathematics achievement: Solving computations and story problems. Journal of Educational Psychology, 76, 194-204.
Marshall, S. P. (1995). Schemas in problem solving. Cambridge: Cambridge University Press.
Middleton, J. A., Littlefield, J., & Lehrer, R. (1992). Gifted students' conceptions of academic fun: An examination of a critical construct for gifted education. Gifted Child Quarterly, 36, 38-44.
Morales, R. V., Shute, V. J., & Pellegrino, J. W. (1985). Developmental differences in understanding and solving simple mathematics word problems. Cognition and Instruction, 2, 41-57.
Morrison, G. R., Ross, S. M., & Baldwin, W. (1992). Learner control of context and instructional support in learning elementary school mathematics. Educational Technology Research and Development, 40(1), 5-13.
Murphy, L. O., & Ross, S. M. (1990). Protagonist gender as a design variable in adapting mathematics story problems to learner interests. Educational Technology Research and Development, 38(3), 27-37.
National Council of Teachers of Mathematics. (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: Author.
Ornstein, A. C. (1992). The textbook curriculum. Educational Horizons, 70(4), 61-63.
Paul, D. J., Nibbelink, W. H., & Hoover, H. D. (1986). The effects of adjusting readability on the difficulty of mathematics story problems. Journal for Research in Mathematics Education, 17, 163-171.
Parker, L. E., & Lepper, M. R. (1992). Effects of fantasy contexts on children's learning and motivation: Making learning more fun. Journal of Personality and Social Psychology, 62, 625-633.

Perkins, D. N., & Salomon, G. (1989). Are cognitive skills context-bound? Educational Researcher, 18(1), 16-25.

Prawat, R. S. (1993). The value of ideas: Problems versus possibilities in learning. Educational Researcher, 22(6), 5-16.

Quintero, A. H. (1983). Conceptual understanding in solving two-step word problems with a ratio. Journal for Research in Mathematics Education, 14, 102-112.

Renninger, K. A. (1992). Individual interest and development: Implications for theory and practice. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), The role of interest in learning and development (pp. 361-395). Hillsdale, NJ: Erlbaum.

Reusser, K. (1988). Problem solving beyond the logic of things: Contextual effects on understanding and solving word problems. Instructional Science, 17, 309-338.

Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem-solving ability in arithmetic. In H. P. Ginsburg (Ed.), The development of mathematical thinking (pp. 153-196). Orlando: Academic Press.

Shelby, A. N. (1992). Readability formulas: One more time. Management Communication Quarterly, 5, 485-495.

Shepard, R. (1988). The imagination of the scientist. In K. Egan & D. Nadaner (Eds.), Imagination and education (pp. 153-185). New York: Teachers College Press.

Sherrill, J. M. (1983). Solving textbook mathematical word problems. The Alberta Journal of Educational Research, 29, 140-152.

Shuard, H. & Rothery, A. (Eds.). (1984). Children reading mathematics. London: John Murray.

Silver, E. A. (1987). Foundations of cognitive theory and research for mathematics problem-solving instruction. In A. H. Schoenfeld (Ed.), Cognitive science and mathematics education (pp. 33-60). Hillsdale, NJ: Erlbaum.

Stewig, J. W. (1988). Children and literature (2nd ed.). Boston: Houghton Mifflin.

Summers, E. G., & Lukasevich, A. (1983). Reading preferences of intermediate-grade children in relation to sex, community, and maturation (grade level): A Canadian perspective. Reading Research Quarterly, 18, 347-360.

Tunnell, M. O. (1994). The double-edged sword: Fantasy and censorship. Language Arts, 71, 606-612.

Van Blerkom, M. L. (1988). Field dependence, sex role self-perceptions, and mathematics achievement in college students: A closer examination. Contemporary Educational Psychology, 13, 339-347.

Webb, N. (1995, July/August). The textbook business: Education's big dirty secret. The Harvard Education Letter, 11, 1-3.

Webb, N., & Yasui, E. (1992). The influence of problem context on mathematics performance (Project 2.1: Alternative Approaches to Assessment in Mathematics and Science) [CD-ROM]. Available: SilverPlatter File: ERIC Item: 349 331

Willoughby, S. S. (1983). Will real math please stand up? Instructor, 92(8), 58-60.

Zollman, A. (1986/1987). Aspects of transfer of learning in mathematical problem solving with respect to the order of problem presentation (Doctoral dissertation, Indiana University, 1986). Dissertation Abstracts International, 47, 2936A-2937A.

Table 1

Grade, Gender, and Town Distributions for Study Sample

	Ashland	Jonesburg
Grade		
Four		
Females	32	36
Males	27	32
Six		
Females	32	38
Males	31	45
Town totals	122	151

Note. Total N (grades 4 plus 6) is 273, which includes 127 fourth graders and 146 sixth graders, 138 females and 135 males.

Table 2

Subsample Size for Phases 1-3 by Grade, Gender, and Town

	N	Female	Male	Ashland	Jonesburg
Phase 1 Ratings: Preferences for Story Ideas					
Grade 4	58 - 59	30 - 32	27 - 28	27 - 29	30 - 32
Grade 6	69 - 70	31 - 35	35 - 38	31	38 - 39
Phase 2 Ratings: Pre-Solution Preferences for Solving Problems					
Grade 4	34 - 40	17 - 23	12 - 19	17 - 20	15 - 21
Grade 6	42 - 49	16 - 25	22 - 26	17 - 21	23 - 28
Phase 3: Problem-Solving Scores					
Grade 4	32 - 36	17 - 22	14 - 19	15 - 17	17 - 21
Grade 6	44 - 46	16 - 25	20 - 30	19 - 22	24 - 25
Phase 3 Ratings: Post-Solution Feelings About Solved Problems					
Grade 4	31 - 36	15 - 22	13 - 20	15 - 17	16 - 21
Grade 6	44 - 46	16 - 25	20 - 30	19 - 22	23 - 25

Note. Number ranges indicate the minimum and maximum number of students who completed the items on the research instruments for each study phase (variation results from omitted items, etc.). The study sample included 127 fourth graders and 146 sixth graders, each of whom completed one-half of the research items in Phase 1 and one-third of the items in Phases 2 and 3.

Table 3

Phase 1 Ratings: Preferences for Story Ideas

	Low Fantasy		High Fantasy		Children's Real-World		Adults' Real-World	
	Range	M	Range	M	Range	M	Range	M
Grade								
Four	3.12 - 3.68	3.32	3.32 - 3.82	3.58	3.24 - 3.73	3.50	2.44 - 2.94	2.76
Six	2.97 - 3.46	3.20	2.72 - 3.76	3.24	3.23 - 3.50	3.31	2.26 - 2.72	2.44

Note. The range and mean are derived from raw scores for the six items in each context category. The number of raw scores is 354 for grade four and 417 for grade six (each student rated three items per category). Raw scores range from 1 ("really dislike") to 5 ("really like").

Table 4

Phase 2 Ratings: Pre-Solution Preferences for Solving Problems

	Low Fantasy		High Fantasy		Children's Real-World		Adults' Real-World	
	Range	M	Range	M	Range	M	Range	M
Grade								
Four	2.88 - 3.53	3.24	3.14 - 4.03	3.53	3.17 - 3.67	3.50	2.80 - 3.45	3.08
Six	2.76 - 3.98	3.33	3.14 - 3.64	3.39	3.33 - 3.67	3.49	2.80 - 3.40	3.11

Note. The range and mean are derived from raw scores for the six items in each context category. The number of raw scores is 220 for grade four and 268 for grade six (each student rated two items per category). Raw scores range from 1 ("very bad") to 5 ("very good").

Table 5

Phase 3 Problem-Solving Scores Aggregated Across All Problem-Context Categories

	Grade-Level Totals			Gender				Town			
				Females		Males		Ashland		Jonesburg	
	N	M	SD	M	SD	M	SD	M	SD	M	SD
Grade											
Four	415	0.58	0.49	0.55	0.50	0.61	0.49	0.43	0.50	0.70	0.46
Six	541	0.76	0.43	0.78	0.42	0.74	0.44	0.69	0.46	0.82	0.39

Note. Means and standard deviations are derived from the sum of raw scores for the twelve Phase 3 items. Raw scores are 0 (inappropriate solution plan) or 1 (appropriate solution plan).

Table 6

Phase 3 Problem-Solving Scores Aggregated by Problem-Context Category

	Low Fantasy		High Fantasy		Children's Real-World		Adults' Real-World	
	M	SD	M	SD	M	SD	M	SD
Grade								
Four	0.63	0.49	0.55	0.50	0.57	0.50	0.57	0.50
Six	0.75	0.44	0.79	0.41	0.83	0.38	0.68	0.47

Note. Means and standard deviations are derived from the sum of raw scores for the three problems in each context category (students only solved the three problems rated most highly by their grade level in Phase 1). The number of raw scores is 104 for grade four and 135 for grade six (each student solved one problem per category). Raw scores are 0 (inappropriate solution plan) or 1 (appropriate solution plan).

Table 7

Phase 3 Problem-Solving Scores by Individual Problem

	Low Fantasy			High Fantasy			Children's Real-World			Adults' Real-World		
Grade												
Four												
M	.83	.45	.57	.44	.72	.47	.53	.41	.75	.69	.54	.47
SD	.38	.51	.50	.50	.45	.51	.51	.50	.44	.47	.51	.51
Six												
M	.64	.87	.73	.80	.84	.72	.77	.82	.89	.57	.82	.64
SD	.49	.34	.45	.41	.37	.46	.42	.39	.31	.50	.39	.48

Note. Figures are given for the three problems solved in each problem-context category in Phase 3. Because these were the most highly rated items from Phase 1 for each grade level, fourth and sixth graders did not solve the same three problems within each category; therefore, figures within the same column for the two grade levels are not necessarily for the same problem. The number of raw scores per problem ranges from 32-36 for grade four and 44-46 for grade six (each student solved one problem per category). Raw scores are 0 (inappropriate solution plan) or 1 (appropriate solution plan).

Table 8

Grade 6 Story-Idea Preference Rating and Problem-Solving Score Comparison

	Low Fantasy			High Fantasy			Children's Real-World			Adults' Real-World		
Story-Idea Rating	3.46	3.14	3.45	3.33	3.76	3.59	3.27	3.50	3.35	2.72	2.49	2.56
Rank Order	4	9	5	7	1	2	8	3	6	10	12	11
Problem-Solving Score	0.64	0.87	0.73	0.80	0.84	0.72	0.77	0.82	0.89	0.57	0.82	0.64
Rank Order	10-11	2	8	6	3	9	7	4-5	1	12	4-5	10-11

Note. Mean story-idea preference ratings are obtained from Phase 1 data and mean problem-solving scores from Phase 3. Phase 1 raw scores range from 1 ("really dislike") to 5 ("really like"). Phase 3 raw scores are 0 (inappropriate solution plan) or 1 (appropriate solution plan). Rank orders within a few numbers of each other may be assumed to be fairly equivalent, particularly those representing means separated by small differences. Two identical means each are assigned the next two numbers in rank order (e.g., 4-5).

Author Note

This article is based on a doctoral dissertation by the same name completed at Indiana University under the direction of Frank K. Lester, Jr.

Sec046

ER

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: The Role of Fantasy and Real-World Problem Contexts in Fourth- and Sixth-Grade Students' Mathematical Problem Solving	
Author(s): Lynda R. Wiest	
Corporate Source: University of Nevada, Reno	Publication Date: Paper presented at AERA 4/17/98

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY ______ Sample ______ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY ______ Sample ______ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY ______ Sample ______ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ↑ [X]	Level 2A ↑ [ ]	Level 2B ↑ [ ]
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Lynda R. Wiest	Printed Name/Position/Title: Lynda R. Wiest, Ph.D. / Asst. Professor of Elem. Educ.	
Organization/Address: Curriculum & Instruction/282 [redacted] College of Education University of Nevada, Reno Reno NV 89557-0214	Telephone: 702-784-4961 x2022	FAX: 702-327-5220
	E-Mail Address: wiest@unr.edu	Date: 8/22/98

(over)

ERIC® Clearinghouse on Assessment and Evaluation

University of Maryland
1129 Shriver Laboratory
College Park, MD 20742-5701

Tel: (800) 464-3742
(301) 405-7449
FAX: (301) 405-8134
ericae@ericae.net
http://ericae.net

March 20, 1998

Dear AERA Presenter,

Congratulations on being a presenter at AERA[1]. The ERIC Clearinghouse on Assessment and Evaluation invites you to contribute to the ERIC database by providing us with a printed copy of your presentation.

Abstracts of papers accepted by ERIC appear in *Resources in Education (RIE)* and are announced to over 5,000 organizations. The inclusion of your work makes it readily available to other researchers, provides a permanent archive, and enhances the quality of *RIE*. Abstracts of your contribution will be accessible through the printed and electronic versions of *RIE*. The paper will be available through the microfiche collections that are housed at libraries around the world and through the ERIC Document Reproduction Service.

We are gathering all the papers from the AERA Conference. We will route your paper to the appropriate clearinghouse. You will be notified if your paper meets ERIC's criteria for inclusion in *RIE*: contribution to education, timeliness, relevance, methodology, effectiveness of presentation, and reproduction quality. You can track our processing of your paper at http://ericae.net.

Please sign the Reproduction Release Form on the back of this letter and include it with **two** copies of your paper. The Release Form gives ERIC permission to make and distribute copies of your paper. It does not preclude you from publishing your work. You can drop off the copies of your paper and Reproduction Release Form at the **ERIC booth (424)** or mail to our attention at the address below. Please feel free to copy the form for future or additional submissions.

Mail to:

AERA 1998/ERIC Acquisitions
University of Maryland
1129 Shriver Laboratory
College Park, MD 20742

This year ERIC/AE is making a **Searchable Conference Program** available on the AERA web page (http://aera.net). Check it out!

Sincerely,

Lawrence M. Rudner

Lawrence M. Rudner, Ph.D.
Director, ERIC/AE

---

[1]If you are an AERA chair or discussant, please save this form for future use.

CUA

The Catholic University of America