C R E S S T

National Center for
Research on Evaluation,
Standards, and
Student Testing

Technical Report

You can view this document on
your screen or print a copy.

UCLA Center for the
Study of Evaluation

in collaboration with:

- University of Colorado
- NORC, University of Chicago
- LRDC, University of Pittsburgh
- The RAND Corporation

The Influence of Problem Context on Mathematics Performance

CSE Technical Report 346

Noreen Webb and Esther Yasui

CRESST/University of California, Los Angeles

October 1992

**National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Graduate School of Education
University of California, Los Angeles
Los Angeles, CA 90024-1522
(310) 206-1532**

The work reported herein was supported under the Educational Research and Development Center Program cooperative agreement R117G10027 and CFDA catalog number 84.117G as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the position or policies of the Office of Educational Research and Improvement or the U.S. Department of Education.

THE INFLUENCE OF PROBLEM CONTEXT ON MATHEMATICS PERFORMANCE

Noreen Webb and Esther Yasui

CRESST/University of California, Los Angeles

The past decade has seen repeated calls to place high priority on problem solving in mathematics instruction (e.g., California State Department of Education, 1985, 1987, 1989; Conference Board of the Mathematical Sciences, 1983a, 1983b; National Council of Teachers of Mathematics, 1980, 1989). The *Curriculum and Evaluation Standards for School Mathematics* (National Council of Teachers of Mathematics, 1989) makes specific recommendations for modifying middle school instruction, including increased attention to extended problem-solving projects and connecting mathematics to the world outside the classroom, and decreased attention to practicing routine, one-step problems and developing skills out of context.

Mathematics educators and researchers argue that using realistic and complex problem-solving contexts can improve problem-solving skills as well as basic skills, increase students' understanding of the mathematics they use, and increase their attitudes toward mathematics (e.g., Lave, Smith, & Butler, 1988; Lesh, 1985; Nesher, 1980; Noddings, 1988; Schoenfeld, 1988), and discourage students from applying memorized algorithms or manipulating numbers without attempting to understand the problem (see Mayer, 1981; Silver & Kilpatrick, 1988; Sowder, 1988).

While the importance of using problems in realistic contexts makes a great deal of intuitive sense, few researchers have compared learning outcomes in mathematics curricula systematically varying the realism and complexity of the problems given to students to solve, nor students' performance on mathematics problems differing in degree of realism and complexity. The present study was designed to investigate both issues. First, it contrasted two versions of an instructional program holding all aspects of content and instruction constant except for the kinds of problems used. One version followed the textbook, consisting of predominantly numerical exercises and short, one- or two-step word problems. The other version used more extended one- to two-week mathematical projects instead of problems from the textbook. Second, the study contrasted student performance on achievement test problems that corresponded to the two curriculum variations. Some problems were short word problems requiring few arithmetic operations (corresponding to the "traditional" curriculum), and others were extended word problems with more information about the context and more operations to apply (corresponding to the more "realistic" curriculum).

The objectives of the study were (a) to determine whether working with more realistic and lengthier problems during instruction would make students better able to solve similar problems on an achievement test, and (b) to determine whether the different kinds of problems (short vs. extended word problems) would provide different information about students' performance and mathematical problem-solving ability.

Method

Sample

Three seventh-grade general mathematics classes at an urban middle school taught by the same teacher participated in the study. Two classes were randomly assigned to the experimental condition (realistic problems); the third class was assigned to the control condition (textbook problems). Comparisons of the two conditions on the pretest showed a slight tendency for the control class to outperform the experimental classes, but few differences were statistically significant. The total enrollment in the three classes was 99 students, but only 82 students had complete data on the tests analyzed here. Consequently, the analyses presented in this paper focus on the sample of 82 students (50 in the experimental condition, 32 in the control condition).

Both instructional programs used small-group work extensively. The previous year's project with the same teacher and comparable classes showed that cooperative small-group problem solving in which students were given instruction in how to help other students improved students' achievement in basic skills and problem solving (Webb, Qi, Yan, Bushey, & Farivar, 1990). So the present study incorporated the same instruction in how to work effectively in small groups and group work in class. Students were assigned to small groups heterogeneously to reflect the mixture of ethnic background, gender, and ability in the classroom.

Instructional Curricula

The content of the 10-week instructional program was operations with whole numbers and decimals. Classes in the experimental condition worked on basic skills for the first six weeks and problem solving for the next four weeks. The development of basic skills used numerical exercises and simple, one-step word problems. Rather than following the order in the textbook (operations treated sequentially, whole numbers before decimals), the basic skill development phase integrated whole numbers, decimals, and the operations. Students also practiced writing and solving simple word problems during this phase: They wrote a word problem on Mondays, and solved the students' problems on Fridays. The purpose of this activity was to encourage students to think about problems, to improve their ability to express mathematical ideas, and to break up the monotony of practice on basic skills.

During the next four weeks in the experimental program, groups worked on several extended problems or projects, each lasting one or two weeks. One project, for example, involved the school vending machine. Students calculated the money collected by the vending machine for different items for different time periods, and calculated profits and losses resulting from stocking the machine from different stores. Another project focused on buying a car (taking into account mileage, gasoline costs, insurance costs, loan costs, depreciation, etc.). The projects were fairly structured, with specific questions for students to research and answer.

The control class followed the textbook (*Addison-Wesley Math*, Eicholz, O'Daffer, & Fleenor, 1989), with one exception. To parallel the schedule in the experimental classes, the first six weeks focused on the numerical exercises in the textbook; the last four weeks focused on the problems. Some sample problems from the textbook are “Oranges cost $0.44 each. How much would 10 oranges cost?” and “A package of notebook paper costs $1.33. The sales tax is 0.06 of the cost. What is the amount of sales tax, rounded to the nearest cent?”

Preparation for Small Group Work

Because most students in this study had little experience working with peers in cooperative, peer-directed settings, all classes carried out a series of activities designed to introduce students to cooperative learning and prepare them to work in groups. Students carried out classbuilding activities to become acquainted with their fellow students and develop a feeling of being part of the class, and teambuilding activities to build team identity; they also participated in activities designed to improve their communication and cooperation skills and received specific instructions and practice in how to help one another solve mathematical problems. (For further detail about the activities carried out in the classrooms, see Farivar and Webb [1991].)

Sequence of Procedures

At the beginning of the program, students were administered pretests of mathematics achievement and problem solving. Students then spent three weeks on class inclusion, teambuilding activities, and instruction on helping behavior in small groups to familiarize themselves with their classmates and to enable them to work effectively in small groups (see previous description). They worked in small groups for the rest of the unit. Each day was a combination of a whole-class introduction by the teacher and small-group work. At the end of the program, students were administered posttests of mathematics achievement and problem solving, including (a) numerical exercises common to both conditions, (b) word problems based on the textbook, and (c) multistep problems based on those used in the experimental condition. (The pretests were shortened versions of the same tests.) All students were administered all tests.

Tests

Three kinds of problems were selected from the tests for analysis in this study: numerical exercises; short, one-question word problems; and extended word problems. Figures 1 through 3 give the different kinds of problems. All problems used a free-response format, and students were instructed to show all of their work. Ample space on the page was given for students to show their work.

1. 3 x 0.6 =

2.
$$\begin{array}{r} 2.31 \\ \underline{\times\ 0.23} \end{array}$$

3. 3 x 3.5 + 2 x 1.50 =

Figure 1. Numerical exercises.

4. Ray pays $3.00 to ride the bus to school. What will it cost him to ride the bus for 23 days?

5. John bought 2 notebooks that cost $3.50 each and 3 pens that cost $1.50 each. How much did he spend?

6.

Burger Palace

Hamburger	$2.50
Fried Chicken	3.50
Beef Burritos	3.00
Coke	.80
Milk	.65

Josie and five of her friends ordered a hamburger and a coke for lunch. What was the total cost of their lunches?

Figure 2. Short word problems.

7. You run a vending machine that sells candy bars. The two best-selling candy bars in your machine are Snickers bars and Butterfingers bars.

Each week you sell 65 Snickers bars and 85 Butterfingers bars. Your machine sells all the candy bars for $0.60 each.

a. After 1 month (4 weeks), how many of each candy bar have you sold?

Snickers ____________ Butterfinger ____________

b. After 1 month, how much money has the machine collected for each type of candy bar?

Snickers ____________ Butterfinger ____________

c. Which candy bar makes more money in one month?

How much more? _______________

d. You decide to raise the price of the candy bars in your machine to $0.70 each. After this price change, how much more money will you make in a month (for each candy bar) than you did at the old price?

Snickers ____________ Butterfinger ____________

8. Your older brother just bought a used car, and now he has to buy car insurance. He is comparing two insurance companies to find out which one has the lowest prices.

At Farmer's Insurance Company, your brother can insure his car for $780.00 <u>per year</u>.

At Allstate Insurance Company, your brother can insure his car for $70.00 <u>per month</u>.

a. Which insurance company has the lowest price for one year of insurance: Farmer's or Allstate?

How much lower is their price? _______________

Figure 3. Extended word problems.

8. b. Your brother gets all A's and B's in school. Because he has good grades, Allstate will give your brother a discount of $10.00 each month on his car insurance.

With the good student discount, how much will one year of insurance cost at Allstate?

c. With the good student discount at Allstate, which insurance company now has the lowest price for one year of insurance: Farmer's or Allstate?

How much lower is their price? _____________

9. There is a Sony Walkman that has every feature you want: recording, AM-FM stereo, auto reverse, etc. It costs $100.

a. Suppose you are going to earn money to buy the Walkman by babysitting for the families in your neighborhood. They will pay you $2.00 per hour. How many hours of babysitting do you have to do to make enough money to buy the Walkman you chose?

b. You would like to buy the Walkman before Christmas. There are 4 weeks left before Christmas. How many hours do you have to babysit each week to earn enough money to buy the Walkman by Christmas?

c. You tell your parents of your plan to work this many hours per week in order to buy the Walkman, but they won't let you work that often during the school year. You decide to wait to buy the Walkman until your birthday which is 10 weeks after Christmas.

(1) How many weeks do you have in which to earn the money for the Walkman?

(2) How many hours do you have to work each week in order to get the Walkman?

Figure 3. (continued)

All of the word problems analyzed here required students to calculate the cost of multiple units. In some problems, the units were tangible objects (e.g., pens, notebooks, candy bars). In other problems, the units were lengths of time (e.g., days, weeks, months). Because the word problems required students to multiply decimal numbers, numerical exercises involving multiplication of decimal numbers were also included to assess students' computational ability in the absence of a verbal context.

Coding of Test Performance

The coding of students' solutions to the problems focused on two aspects of performance: (a) their ability to generate an arithmetic expression that corresponded to the problem (hereafter called the "setup" for the problem), and (b) their ability to execute their setup. The first category focuses on students' comprehension of the problem and their ability to translate from the verbal representation of the problem to an appropriate arithmetic expression. The second category focuses on students' ability to carry out the numerical calculations themselves. Most of the analyses focused on specific errors that students made in each category rather than a summary score for a problem.

Importance and Implications

This study has important implications for mathematics instruction. It raises questions about the effects of the kinds of problems included in the curriculum on students' achievement. Other features of instruction (such as whether students work collaboratively in small groups) may have greater effects on achievement and problem-solving performance than do the kinds of problems. And, the kinds of problems used in the curriculum may have greater effects on other outcomes, such as problem-solving strategies and attitudes, than on achievement.

Results

Differences Between Control and Experimental Classes

The first set of analyses compared student performance in the class following the textbook (control class) and the classes using the experimental curriculum (experimental classes). The performance in the two instructional conditions was remarkably similar across the three types of problems. The slight differences appearing between conditions did not show a consistent pattern nor were they statistically significant. Detailed breakdowns of performance on each problem for the two instructional conditions appear in Appendices A to I. Because no systematic differences in student performance emerged between instructional conditions, the instructional conditions were combined for further analyses.

Comparisons Between Problem Types: Overall Performance

Table 1 gives the overall rate of accuracy on the test problems analyzed here. The first column gives the accuracy of the arithmetic expression that the students generated to represent the problem, without regard to the computational accuracy of their arithmetic operations. Students could set up the problem correctly, that is, generate the correct arithmetic expression, and then perform the arithmetic manipulations incorrectly. They would be scored as correct for the column marked "setup." The second column gives the computational accuracy of students' work without regard to the accuracy of their setup. Students could set up an arithmetic expression that did not accurately represent the problem, but carry out their arithmetic manipulations correctly. They would be scored as correct for the column marked "computation." Later tables give frequencies of specific kinds of setup and computational errors.

Table 1 shows quite clearly that the kind of problem (numerical exercise, short word problem, extended word problem) had little overall impact on students' performance. No category of problem was more difficult than another overall. Rather, students' performance was more variable within a kind of problem than between problem types. The wide variability in student performance is especially evident in the accuracy of their setups: students showed considerable variability in their ability to generate a correct arithmetic expression for both short and extended word problems.

The only overall difference between problem types was in computational accuracy of numerical exercises compared to that of word problems (both short and extended). Computational accuracy was higher for word problems than for numerical exercises. Students made fewer errors when carrying out numerical manipulations in word problems (e.g., multiplying numbers and placing the decimal point) than in numerical exercises. This result makes sense because the context supplies clues that allow students to judge the reasonableness of their answers. For example, since many of the word problems involved amounts of money, there was no ambiguity about where to place the decimal point in an answer. Furthermore, the money context of the

Table 1

Accuracy of Performance on All Problems
(percentage of students correct)

Problem	Setup	Computation	Total Problem
Numerical Exercises			
1	n/a	72	72
2	n/a	24	24
3	57	40	39
Short, One-Question Word Problems			
4	85	80	76
5	76	88	68
6	32	79	30
Extended Word Problems			
7a1	77	87	73
7a2	77	89	74
7b1	56	68	52
7b2	56	67	51
7c	77	71	67
7d1	21	63	18
7d2	21	63	20
8a	56	74	54
8b	46	78	48
8c	62	72	60
9a	72	87	72
9b	28	52	26
9c1	48	72	48
9c2	23	37	21

problem may have enabled students to evaluate the magnitude of their answers and make corrections where necessary. Some students probably had some concept of the order of magnitude to expect in their answer. An answer too large or too small may have been questioned, checked, and corrected.

Specific Errors for Each Problem Type

Numericalexercises. Tables 2 and 3 give the breakdowns of specific errors for the numerical exercises. It should be noted that in all tables, the percentages of students making specific types of errors (e.g., numerical calculation, placement of decimal point) do not sum to the percentages of students in the summary categories (e.g., computation errors) because some students made errors in multiple categories and so are included in more than one percentage.

Quite a few students had difficulty multiplying numbers correctly and even more had difficulty placing the decimal point in the answer correctly. Furthermore, some students carried out arithmetic operations in the wrong order when there were multiple operations to perform (e.g., carrying out addition before multiplication, see Table 3).

Some students also gave uninterpretable answers. From what they wrote on their papers, it was impossible to determine what operations they were trying to carry out. Some uninterpretable answers were simply numbers that had no correspondence to the problem (e.g., the date that the test was administered, or a very large number such as 1,000,000).

Table 2

Breakdown of Errors on Problems 1 and 2
(percentage of students making each error)

	Problem	
Error	1	2
COMPUTATION ERRORS	28	74
Numerical calculation	21	39
Placement of decimal point	15	61
BLANK	0	7

Note. Because some students made more than one kind of error, percentages in specific error categories do not sum to percentages in overall error category.

Table 3

Breakdown of Errors on Problem 3
(percentage of students making each error)

Error	Percentage of Students
ARITHMETIC OPERATION ERRORS	28
Incorrect order of arithmetic operations (e.g., addition before multiplication)	22
Incorrect arithmetic operation	2
Skipped one or more arithmetic operation(s)	4
COMPUTATION ERRORS	46
Numerical calculation	34
Placement of decimal point	39
UNINTERPRETABLE	6
BLANK	7

Short word problems. Tables 4 through 6 give the breakdowns of specific errors for the short word problems. A common error in two of the three problems was to treat multiple units as a single unit. Problem 5 asked for the cost of multiple pens and notebooks, and problem 6 asked for the cost of multiple hamburgers and cokes. Instead of calculating the cost of multiple items, some students included the cost of only one item (for example, treating three pens as one pen, or treating six cokes as one coke, see Tables 5 and 6). Either students failed to recognize that the problem asked for the cost of multiple units (failure to encode the information initially), or they encoded the information but neglected to include it in their calculations.

Surprisingly, more students used the incorrect arithmetic operation in the problem that involved only one kind of unit (days) and so required only one calculation (number of days multiplied by the cost per day, see Table 4) than in the problems with multiple sets of units (pens and notebooks; hamburgers and cokes) which required multiple calculations (see Tables 5 and 6). The difference in the nature of the units in the problems may explain this surprising result. The arithmetically simpler problem involved units of time

Table 4

Breakdown of Errors on Problem 4
(percentage of students making each error)

Error	Percentage of Students
SETUP ERRORS	10
Failed to recognize set of items (treated 23 days as 1 day)	1
Incorrect arithmetic operation	10
Extracted incorrect given information (e.g., 24 days instead of 23)	2
COMPUTATION ERRORS	15
Numerical calculation	12
Decimal point	6
BLANK	5

Table 5

Breakdown of Errors on Problem 5
(percentage of students making each error)

Error	Percentage of Students
SETUP ERRORS	21
Omitted one type of item entirely (notebooks or pens)	1
Treated multiple units in a set as one unit (e.g., treated 3 pens as 1 pen)	17
Used incorrect arithmetic operation	2
Extracted incorrect given information (e.g., 2 pens instead of 3)	2
COMPUTATION ERRORS	9
Numerical calculation	7
Decimal point	1
BLANK	4

Table 6

Breakdown of Errors on Problem 6
(percentage of students making each error)

Error	Percentage of Students
SETUP ERRORS	77
Omitted one type of item entirely (e.g., coke)	9
Failed to recognize set of items (treated multiple cokes as 1 coke)	17
Used incorrect arithmetic operation	5
Extracted incorrect given information (e.g., 4 persons)	2
Failed to include Josie in total number of persons	45
COMPUTATION ERRORS	16
Numerical calculation	15
Decimal point	2
BLANK	5

(days), whereas the arithmetically more complicated problem involved tangible items (e.g., pens and notebooks). Students may have more difficulty conceptualizing problems with units of time than problems with tangible objects. Corroborating evidence appears in students' performance on some of the extended word problems which involved units of time, as will be discussed below.

In the problems with two types of items, some students omitted one type of item entirely (for example, failing to calculate the cost of multiple cokes in problem 6). These students may have been accustomed to solving problems requiring a single calculation. Once they completed one calculation (e.g., the cost of multiple hamburgers), they assumed that they were finished with the problem.

A few students extracted the wrong number from the information given (e.g., using 24 days instead of 23 in problem 4 or using 2 pens instead of 3 in problem 5). Compared to other kinds of errors, this was fairly insignificant and rarely occurred.

A large number of students made an error specific to problem 6, which is probably due to the phrasing of that particular problem. The problem asks for the cost of hamburgers and cokes for Josie and five of her friends. Nearly half of the students failed to include Josie in their calculations, and so calculated the cost for five persons, instead of six. Perhaps students did not encode the numerical meaning of “Josie and five of her friends” but merely focused on the number stated (five) and assumed that number was the total number of persons. Or students may have misinterpreted the question asked. In “What was the total cost of their lunches?”, students may have interpreted “their” to refer to Josie's friends, not Josie.

Finally, compared to the numerical exercises, students had much less difficulty carrying out their computations in the short word problems. This is seen most clearly in a comparison between problem 3 (Table 3) and problem 5 (Table 5). The two problems were designed to be comparable in all respects (operations to be performed, numbers used) except for the presence of a verbal context: Problem 3 is a numerical exercise and problem 5 is a short word problem.

Without the aid of a verbal context, students made many more computational errors (Table 3 vs. Table 5). As was pointed out earlier, the money context gave an obvious clue about the placement of the decimal point. Less obvious is the role of the context in aiding numerical calculations. In the numerical exercise, many students made multiplication errors such as 1.5 x 2 = 2 or 2 x 1.5 = 5. But few students made such errors when solving problem 9, even though the numbers were quite similar (e.g., 3 x 1.5). How they avoided making similar multiplication errors in the word problem is not clear.

Extendedwordproblems. Tables 7 through 9 give the breakdowns of specific errors for the extended word problems. Students made some of the same errors that they had made on the short word problems. They sometimes treated a set of items as a single item (e.g., using the cost of a single candy bar instead of the cost of a week's worth of candy bars sold), used incorrect arithmetic operations, extracted incorrect numbers from the information given, and made computation errors.

New types of errors emerged in the extended word problems that did not occur in the short word problems. One error could have occurred on the short

Table 7

Breakdown of Errors on Problem 7
(percentage of students making each error)

Error	7a1	7a2	7b1	7b2	7c	7d1	7d2
SETUP ERRORS	15	16	29	28	9	61	61
Misinterpreted type of answer required (cost vs. number of units)	10	10	6	6	6	2	5
Failed to recognize set of items (treated set as single item)	0	0	2	2	0	13	16
Failed to recognize that change in price applies to each item in set (e.g., treated as change in price for set)	n/a	n/a	n/a	n/a	n/a	15	17
Failed to recognize change in length of time from given information (1 week) to question being asked (1 month): Used 1 week.	7	7	15	13	2	21	22
Incorrectly translated length of time from given information (1 week) to question being asked (1 month) (e.g., interpreted as number sold per day and used 28 days per month)	5	5	2	1	2	1	1
Failed to recognize that problem asked for a difference in costs (e.g., gave larger cost or both costs)	n/a	n/a	n/a	n/a	4	52	54
Used incorrect arithmetic operations	2	2	9	9	6	6	9
Extracted incorrect given information (e.g., 14 Snickers instead of 65)	0	0	2	4	0	0	0
COMPUTATION ERRORS	6	4	18	17	15	18	20
Numerical calculation	6	4	4	5	10	11	10
Decimal point	0	0	16	13	9	9	12
UNINTERPRETABLE	5	4	6	6	9	5	4
BLANK	4	4	7	10	6	13	13

word problems but did not. On the extended word problems, some students misinterpreted the type of answer required: confusing the cost for a set of units with the number of units. For example, in problem 7, some students gave the cost of the candy bars sold instead of the number of candy bars sold, or vice

Table 8

Breakdown of Errors on Problem 8 (percentage of students making each error)

Error	8a	8b	8c
SETUP ERRORS	34	38	12
Failed to recognize change in length of time from given information (1 month) to question being asked (1 year)	17	13	6
Incorrectly translated length of time from given information (1 month) to question being asked (1 year) (e.g., interpreted as cost per day and used 365 days)	23	17	7
Failed to recognize that change in length of time applies to one cost (Allstate): Applied time change to both costs or did not apply to either cost	20	23	0
Failed to recognize that problem asked for a difference in costs	13	n/a	7
Failed to consider discount	n/a	0	12
Failed to apply time change to discount	n/a	26	0
Misinterpreted question as asking for difference between discounted and undiscounted cost	n/a	0	4
Used incorrect arithmetic operations	9	4	4
Extracted incorrect given information (incorrect value of insurance)	0	0	1
COMPUTATION ERRORS	16	6	2
Numerical calculation	10	4	2
Decimal point	7	2	0
UNINTERPRETABLE	5	6	6
BLANK	5	10	20

versa (Table 7). Students never made this error in the short word problems, even though they also involved multiple units and costs. Because the extended word problems presented more information about the context of the problem (to make it more realistic), the question asked may have been less salient to students.

Table 9

Breakdown of Errors on Problem 9
(percentage of students making each error)

Error	9a	9b	9c1	9c2
SETUP ERRORS	16	32	24	18
Misinterpreted type of answer required (hours vs. amount of money)	10	2	4	6
Misinterpreted amount of time given (e.g., used $2.00/minute instead of $2.00/hour)	6	6	20	2
Confused different types of units (e.g., confused dollars and hours)	7	27	1	15
Included extraneous information (tax)	7	6	9	9
Failed to recognize that the problem asked about an amount per week	n/a	6	n/a	5
Failed to consider or apply the number of hours	n/a	29	n/a	17
Failed to use information given or calculated in previous problem (4 weeks or total number of weeks)	n/a	n/a	20	7
Confused different units of time (e.g., hours vs. weeks)	n/a	n/a	16	n/a
Used incorrect arithmetic operations	6	7	5	2
COMPUTATION ERRORS	1	7	9	5
Numerical calculation	1	7	9	5
Decimal point	0	0	0	0
UNINTERPRETABLE	9	21	9	26
BLANK	4	20	20	33

Other new types of errors corresponded to specific features of the extended word problems that had no counterpart in the short word problems. The major new feature was the introduction of time into the context. All of the extended word problems involved units of time and cost over changing units of time (e.g., cost for candy bars sold during 1 week vs. cost for candy bars sold during 4 weeks; cost of insurance for 1 month vs. cost of insurance for 1 year). The initial information presented about the context gave one length of time (e.g., one week) and the questions typically posed questions for a different length of time (e.g., one month, which students were told to treat as 4 weeks). Changing lengths of time were considered important features of the contexts of the extended word problems (e.g., operation of a vending machine, cost of car insurance).

Students had great difficulty dealing with units of time and changing lengths of time. As can be seen in Tables 7 through 9, most of the errors involved time. Sometimes students ignored the change in length of time altogether and operated on the problem as if no change in time occurred (e.g., performing all calculations in problem 7 for one week instead of four weeks, Table 7). Sometimes they recognized that the length of time changed, but translated incorrectly from one unit to another. For example, problem 7 required students to multiply the cost of one week's worth of candy bars by four to obtain the cost of a month's worth. Some students, however, multiplied by 28 as if they interpreted the given information to pertain to the candy bars sold during one day and then translated from one day to 28 days (the number of days in a month if a month has 4 weeks, Table 7). Some students even confused other kinds of information with units of time when setting up an arithmetic expression. For example, in problem 9, some students used the number of dollars in an expression where they should have used the number of hours (Table 9).

Two of the extended word problems asked students to calculate a difference in costs between kinds of items, not only the cost for each kind of item (e.g., difference between cost of two kinds of candy bars in problem 7; difference between two kinds of car insurance in problem 8).

Some students misinterpreted this question and merely gave the cost of one kind of item, whichever was greater. This was especially a problem in the last part of problem 7 which embedded several questions in one. That part of the problem asked “After this price change, how much more money will you make in a month (for each candy bar) than you did at the old price?” This question requires students to calculate the cost of the candy bars at the new price and then compare the new cost to the old cost. Over half of the students ignored the comparison question and simply gave the new costs (Table 7).

Finally, a substantial number of students left the extended word problems blank or gave uninterpretable answers. The amount of material presented to students may have been too intimidating for some of them.

Comparison of Errors Across Problem Types

Table 10 compares error rates for the kinds of errors that are common to short and extended word problems. With the exception of the categories of BLANK and UNINTERPRETABLE responses (discussed further below), the differences in performance between short and extended word problems seem to be differences in kind rather than differences in degree.

Overall, the percentages of students who made setup errors were similar across the two types of word problems. But the overall mean similarity masks differences in opposite directions for specific setup errors. As mentioned earlier, a major difference between types of word problems is that students sometimes misinterpreted the type of answer required in the extended word problems but never did in the short word problems (cost vs. number of items). The problem may have been one of cognitive overload. The extended word problems presented more information than the short word problems, and the question itself was a smaller portion of the total information. The amount of information presented in the extended word problems was larger in several respects. There was more verbal description of the context itself. More words were used to express the given information and to connect the given information to the context. And, finally, more numerical information was presented. Because students had more information to digest in the extended word problems, they may have been less able to correctly encode the question being asked or to remember it while they attended to specific calculations.

Another difference between short and extended word problems, but in the opposite direction, is that on the short word problems a few students omitted one type of item completely (e.g., calculating the cost of hamburgers and omitting the cost of cokes). This never occurred on the extended word problems. This result may have been due to the difference in format between the types of word problems. The extended word problems gave separate spaces for students to write the cost of each item (separate spaces for each candy bar

Table 10. Types of Errors Common to Short and Extended Word Problems (percentage of students making each error)

	Short Word Problems			Extended Word Problems													
ERROR	**4**	**5**	**6**	**7a1**	**7a2**	**7b1**	**7b2**	**7c**	**7d1**	**7d2**	**8a**	**8b**	**8c**	**9a**	**9b**	**9c1**	**9c2**
SETUP ERRORS																	
Misinterpreted Answer Type	0	0	0	10	10	6	6	6	2	5	0	0	0	10	2	4	6
Omitted Type of Item	0	1	9	0	0	0	0	0	0	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Treated Set as Single Item	1	17	17	0	0	2	2	0	13	16	17	13	6	n/a	n/a	n/a	n/a
Incorrect Operation	10	2	5	2	2	9	9	6	6	9	9	4	4	6	7	5	2
Extracted Incorrect Number	2	2	2	0	0	2	4	0	0	0	0	0	1	0	0	0	0
COMPUTATION ERRORS																	
Numerical Calculation	12	7	15	6	4	4	5	10	11	10	10	4	2	1	7	9	5
Decimal Point	6	1	2	0	0	16	13	9	9	12	7	2	0	0	0	0	0
UNINTERPRETABLE	0	0	0	5	4	6	6	9	13	13	5	6	6	9	21	9	26
BLANK	5	4	5	4	4	7	10	6	13	13	5	10	20	4	20	20	33

in problem 7, see Figure 3). So students did not have to rely on memory to include both candy bars in their calculations. (Interestingly, this format was incorporated into the test at the teachers' request—they thought that the extended problems would be too difficult without giving separate answer blanks. Students performed better than the teachers expected, however, so the separate answer blanks could probably have been omitted. This would have allowed a better comparison of students' tendency to omit a type of item.)

The remaining setup errors common to short and extended word problems—treating a set of units as a single unit, using incorrect arithmetic operations, and extracting an incorrect number from the information given—occurred just as frequently on the two types of word problems. In any case, the large variability in error rates across problems of one type swamps the small differences in means between problem types.

The tendency of students to make computational errors was similar for short and extended word problems. As before, the problem-to-problem variation in computational errors was much greater than any mean difference between problem types.

The most dramatic differences between short and extended word problems are the frequencies of uninterpretable or blank responses. This is the one result that points to a difference in degree of performance between the two types of word problems. Whereas students' answers on the short word problems were always interpretable, quite a few students gave answers to the extended word problems that were impossible to interpret. They seemed to bear no relationship to the problem asked. Moreover, large numbers of students left parts or all of the extended word problems blank. As suggested earlier, perhaps they were intimidated by the large amount of information presented. Some students had limited proficiency in English and may simply have decided not to exert the considerable effort required to digest the information given.

This comparison of common errors across short and extended word problems gives a very different picture from the rates of overall accuracy given in Table 1. The overall accuracy of short and extended word problems was very similar (Table 1). But the analysis of common errors shows important qualitative differences between the word problem types. The major differences

are that students were more likely to misinterpret the kind of answer required, to give uninterpretable answers, or to leave the item blank on extended word problems than on short word problems. In these respects, short word problems overestimate students' ability to interpret information in problems presented in realistic contexts.

Erosion of Performance Over Extended Problems

The extended word problems show another aspect of students' problem-solving ability that short word problems do not: their ability to sustain effort and accuracy over an extended problem. As before, looking only at overall accuracy rates (as in Table 1) gives only limited information. Table 1 shows some indication of decreasing performance over an extended item, but few clearcut patterns, and little information about the nature of eroding performance across parts of an item.

The breakdowns of specific errors on parts of the extended problems give more detailed information about the trends in student performance across parts of an item. As can be clearly seen in Tables 7 through 9 (last row in each table), the number of blank responses increases fairly steadily on all problems. Few students left the first part of a problem blank, but more students left succeeding parts of the problems blank. This is not the pattern often seen on traditional tests, where students may reach a certain point in the test and do little or no work on succeeding problems. Rather, nearly all students started each extended problem, but substantial numbers of students failed to complete the problem.

Was this drop-off in attempts to solve a problem due to increasing difficulty with the problem, a decrease in motivation, or both? One way to try to answer this question is to examine students' performance immediately preceding a part of a problem left blank. If a student had difficulty solving the preceding part of the problem, then leaving the subsequent part of the problem blank could be attributed to student difficulty. On the other hand, if a student solved the previous part of the problem correctly, it would seem unlikely that the student suddenly experienced so much difficulty that he or she could not approach the problem. Rather, we might suspect that the student simply was not motivated to apply the effort to solving the problem.

Table 11 shows, for students who left a part of the problem blank, their performance on the preceding part of the problem. The results show both factors at work. The first row shows that some students solved one part of a problem successfully and then left the next part blank. The second row shows that other students had difficulty with one part of a problem and left the next part blank. The final row shows that still others left a part of a problem blank because they left the previous part blank. The surprising result in Table 11 is that, overall, the percentages of students in the three categories are quite similar. It wasn't only the students having difficulty solving the problem who left parts blank; some students who appeared to have little or no difficulty solving the problem also left parts blank. So it seems that extended problems measure students' motivation to continue working on a problem, quite apart from their ability to do so.

On all other specific errors, there is no consistent trend of increasing numbers of errors from beginning to end of an extended problem. Rather, on two problems, there seems to be distinction between the first part of the problem and the rest of the problem. Error rates on the first part of the problem were lower than on the rest of the problem for most types of errors (see Tables 7 and 9). There did not seem to be a steady, cumulative effect of the amount of information to be processed, but rather a difference between students' performance on the first part of a problem and the rest of the problem taken as a whole. Perhaps students approached the first part of a problem

Table 11

Performance on the Part of Extended Word Problem Immediately Preceding a Blank Response (percentage of students)

Accuracy of Setup on Previous Part	Problem								
	7b1	7b2	7d1	7d2	8b	8c	9b	9c1	9c2
Correct	2	5	4	4	1	4	10	1	1
Incorrect	4	2	4	4	5	7	10	6	12
Blank	1	3	5	5	4	9	1	13	20

differently from subsequent parts of the problem, taking more time to encode the information, set up their arithmetic representation, and carry out their calculations on the first part of the problem than on later parts. The greater numerical calculation errors on later parts than on the first part on these two problems suggests that students were being less careful with their work.

Ability to Build on Previous Work in Extended Problems

The extended problems also give the opportunity to examine students' ability to build on previous work in a problem. Some parts of the problems were designed so that students could simply perform an additional calculation on an answer already calculated in a previous part of the problem. If students recognized the connections between different parts of the same problem, they would logically build on their previous result, rather than performing their previous calculations from scratch or starting some other calculation from scratch. For example, in problem 7, once students calculated the number of candy bars sold in one month (7a1, 7a2), they could use their answers to calculate the amount of money collected in one month (7b1, 7b2), and the amount of money collected in one month at the new price (7d1, 7d2). It was not necessary to repeat the initial calculation.

Table 12 shows that large percentages of students did not build on their previous work. They often recalculated previous correct answers or failed to recognize the relevance of their previous correct answers and performed completely new (usually incorrect) work. A few students even repeated previous calculations that were conceptually incorrect (based on misunderstandings of the problem).

It is impossible to know whether students who repeated their previous calculations chose to do this to make sure that they did not make an error the first time or whether they did not realize that their previous answer could be used in the subsequent part of the problem. The large frequencies of students who started completely new work on later parts of the problem even though they had obtained correct answers on earlier parts makes the second explanation more likely. The vast majority of students who started new work on later parts of a problem generated an incorrect setup for the problem that was different from what they had done earlier. It seems, then, that a

Table 12

Failure to Use Previous Answers in Later Parts of Extended Problems (percentage of students)

	Problem					
Behavior	7b1	7b2	7d1	7d2	9b	9c2
DID NOT USE PREVIOUS ANSWER	30	28	41	40	44	16
Repeated previous work	17	16	10	10	0	0
Previously correct	11	11	9	9	0	0
Previously incorrect	6	5	1	1	0	1
Started new and different work	13	12	31	30	44	16
Previously correct	13	12	18	17	29	15
Previously incorrect	0	0	13	13	15	1
DID USE PREVIOUS ANSWER	51	52	34	34	36	24
Previously correct	44	43	33	33	34	23
Previously incorrect	7	9	1	1	2	1

Note. For problem 7, "previous answer" refers to the number of candy bars sold in one month (7a1, 7b1). For problem 9, "previous answer" refers to the number of hours of babysitting needed to buy a Walkman (9a).

substantial number of students failed to recognize the connections in the problem and treated parts of a problem as distinct problems to be addressed anew.

It should be noted that Table 12 surely underestimates students' difficulty in recognizing the connections between parts of a problem. A number of students who performed correct work on earlier parts of the problem gave uninterpretable answers to later parts. Those students were not included here, however, because it was not clear what they did on the later parts. It is quite likely that some (or even all) of them failed to build on their previous work. That is, they did not use what they had done before.

Interestingly, the teachers in the study had advocated separating the problems into more steps instead of fewer to make the problems easier for the students. Unwittingly, this may have increased the amount of work for students who treated each part of a problem as a separate problem, not connected to the previous ones. Students' tendency to ignore previous work may also shed some light on the increasing number of parts of the problem that students left blank. They may have felt overwhelmed by the perceived amount of work required for each subsequent part, even though much of it was unnecessary.

A related question is whether students are able to go back to previous parts of the problem to retrieve information they need to solve the problem. Some parts of the problems *required* students to use information given or calculated in previous parts of the problem. Problem 8, part c, for example, requires students to use the information or their answer from part b. And each part of problem 9 requires the student to retrieve information or answers from previous parts of the problem. Were students able to build on the connections from one part to the next in an extended problem?

Table 13 gives the percentages of students who did not go back to previous parts of the problem to retrieve necessary information. The results show that on problems requiring students to refer to previous parts of the problem for necessary information, many failed to do so. They simply operated on incomplete information. This is further evidence that students failed to see the connections among the parts of an extended problem.

Table 13

Percentage of Students Failing to Use Necessary Information in Previous Parts of Extended Word Problems

Error		Percentage of Students
8c:	Did not use discount given in part b	17
9b:	Did not use hourly wage given in part a	43
9c1:	Did not include 4 weeks given in part b	28
9c2:	Did not include 4 weeks given in part b	41

The results presented here raise the question of whether separating a complex problem into separate steps as was done here makes the job easier or more difficult for students. It was done here to simplify their work and reduce the cognitive demands at each step, but instead may have made their task more difficult. The extended problems administered in this study could easily be condensed into fewer steps. For example, in problem 7, parts a and c could be eliminated without changing the goal of the problem. The problem could be condensed even further into a single, albeit complex, question by asking students to compare the amount of money their vending machine would collect in one month at two different prices. Comparing performance across complex problems with more and fewer intermediate steps made explicit would show the advantages and disadvantages of separating a complex problem into steps.

Discussion

The first major result of the study was the similarity in performance for the two instructional conditions. Students performed the same on all problems regardless of the type of problem they solved during instruction. More specifically, students' experience with extended, realistic problems during instruction in the experimental classes did not give them an advantage when solving such problems on the posttest. Similarly, the lack of experience with such problems did not seem to disadvantage students in the control condition. Working on traditional word problems from textbooks and working on extended problems from real life seemed to equip students with similar problem-solving skills.

An alternative interpretation of the similarity in results for the two instructional conditions is that the test problems were too structured to detect differences in students' problem-solving skills. That is, the high degree of structure in the extended word problems on the test may have compensated for control students' lack of experience with more complex problem contexts. Perhaps presenting test problems without indicating the intermediate steps would have shown more differences between instructional conditions.

The second major result of the study was that the different kinds of word problems revealed different information about students' problem-solving skills. Moreover, these differences were detected only using detailed scoring of specific kinds of errors. Scoring problems globally, concentrating on overall

accuracy or even separating conceptual understanding (as indicated by the accuracy of the arithmetic expression students generated to represent the problem) from computational ability, was not sufficient to detect differences. Only by coding specific errors were differences between problem types revealed.

Detailed coding of specific errors showed qualitative differences in performance between short and extended word problems. On the extended word problems, some students misinterpreted the type of answer required (e.g., number of units vs. cost of the set of units), whereas such misinterpretation never occurred on the short word problems. Conversely, on the short word problems, some students omitted a type of item entirely, whereas this error never occurred on the extended word problems. So short word problems would overestimate students' ability to interpret the type of answer required, and extended word problems would overestimate students' ability to include all relevant given information in the problem. Because these effects worked in different directions, they would cancel out in global scoring of the problems.

Extended word problems also showed aspects of problem-solving performance that could not be measured with short word problems. Most important was the erosion of student performance over the course of an extended word problem. Increasing numbers of students gave uninterpretable or blank responses as they progressed through an extended word problem. Moreover, many students showed an inability to build on their previous work in extended problems and a failure to recognize that the parts of a problem were interconnected. These difficulties were not evident in the first part of extended problems. Students showed the same tendency to give uninterpretable or blank responses in the first part of an extended word problem as in short word problems. The difficulties were revealed only by tracking students' performance over the course of extended word problems.

Questions remain about the exact reasons for differences in performance between short and extended word problems. The short and extended problems used in this study differed in several major respects: format of the presentation (one question asking for one answer vs. questions presented in steps, each requiring an answer), complexity of the context (lengthier verbal presentation, more information about the context), and amount of numerical information

provided. Which feature or features are responsible for the differences between problem types found here is not clear.

To clarify which features of the word problems accounted for the differences in performance observed here, it would be necessary to design problems which vary one feature at a time. For example, multistep problems with simple and complex contexts could be compared to determine whether the multiple-step format or the complexity of the problem context was responsible for the erosion in student performance and failure to build on previous work. As another example, single answer and multistep problems with similar levels of complexity in the problem context could be compared to determine whether separating the problem into multiple steps aids or hinders students' ability to solve it.

In any case, the comparisons in this study suggest that there are important aspects of students' ability to solve structured problems that are not measured with traditional, short, one-question word problems and that can be measured with extended problems with more realistic and complex contexts. Further studies are needed to examine the most efficient ways to measure these additional aspects of students' problem-solving ability.

References

California State Department of Education. (1985). *Mathematics framework for California public schools: Kindergarten through grade twelve.* Sacramento: California State Department of Education.

California State Department of Education. (1987). *Mathematics model curriculum guide, kindergarten through grade eight.* Sacramento: California State Department of Education.

California State Department of Education. (1989). *A question of thinking.* Sacramento: California State Department of Education.

Conference Board of the Mathematical Sciences. (1983a). *The mathematical sciences curriculum K-12: What is still fundamental and what is not* (Report to the National Science Board Commission on Precollege Education in Mathematics, Science, and Technology). Washington, DC: Conference Board of the Mathematical Sciences.

Conference Board of the Mathematical Sciences. (1983b). *New goals for mathematical sciences education* (Report of a conference sponsored by Conference Board of the Mathematical Sciences, Airlie House, Warrenton, Virginia, November 1983). Washington, DC: Conference Board of the Mathematical Sciences.

Eicholz, R. E., O'Daffer, P. G., & Fleenor, C. R. (1989). *Addison-Wesley math.* Menlo Park, CA: Addison-Wesley Publishing Company.

Farivar, S., & Webb, N. M. (1991). *Helping behavior activities handbook.* Los Angeles: University of California, Graduate School of Education.

Lave, J., Smith, S., & Butler, M. (1988). Problem solving as everyday practice. In R. I. Charles & E. A. Silver (Eds.), *The teaching and assessing of mathematical problem solving* (Vol. 3, pp. 61-81). Reston, VA: The National Council of Teachers of Mathematics, Inc.

Lesh, R. (1985). Conceptual analyses of problem-solving performance. In E. A. Silver (Ed.), *Teaching and learning mathematical problem solving: Multiple research perspectives* (pp. 309-329). Hillsdale, NJ: Lawrence Erlbaum.

Mayer, R. E. (1981). Frequency norms and structural analysis of algebraic story problems into families, categories and templates. *Instructional Science, 10,* 135-175.

National Council of Teachers of Mathematics. (1980). *An agenda for action: Recommendations for school mathematics of the 1980s.* Reston, VA: National Council of Teachers of Mathematics, Inc.

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

Nesher, P. (1980). The stereotyped nature of school word problems. ***For the Learning of Mathematics, i,*** **41-48.**

Noddings, N. (1988). Preparing teachers to teach mathematical problem solving. In R. I. Charles & E. A. Silver (Eds.), ***The teaching and assessing of mathematical problem solving*** **(Vol. 3, pp. 244-258). Reston, VA: The National Council of Teachers of Mathematics, Inc.**

Schoenfeld, A. H. (1988). Problem solving in context(s). In R. I. Charles & E. A. Silver (Eds.), ***The teaching and assessing of mathematical problem solving*** **(Vol. 3, pp. 82-92). Reston, VA: The National Council of Teachers of Mathematics, Inc.**

Silver, E. A., & Kilpatrick, J. (1988). Testing mathematical problem solving. In R. I. Charles & E. A. Silver (Eds.), ***The teaching and assessing of mathematical problem solving*** **(Vol. 3, pp. 178-186). Reston, VA: The National Council of Teachers of Mathematics, Inc.**

Sowder, L. (1988). Choosing operations in solving routine story problems. In R. I. Charles and E. A. Silver (Eds.), ***The teaching and assessing of mathematical problem solving*** **(Vol. 3, pp. 148-158). Reston, VA: The National Council of Teachers of Mathematics, Inc.**

Webb, N. M., Qi, S., Yan, K. X., Bushey, B., & Farivar, S. (1990, April). ***Cooperative small-group problem solving in middle school mathematics.*** **Paper presented at the annual meeting of the American Educational Research Association, Boston.**

Appendices

Appendix A

Accuracy of Performance on All Problems by Treatment (percentage of students correct)

Problem	Setup		Computation		Total Problem	
	Control	Exper.	Control	Exper.	Control	Exper.
Numerical Exercises						
1	n/a	n/a	69	74	69	74
2	n/a	n/a	31	20	31	20
3	50	62	28	48	28	44
Short, One-Question Word Problems						
4	94	80	91	74	88	68
5	75	76	88	88	66	70
6	28	34	72	84	25	34
Extended Word Problems						
7a1	81	74	91	84	78	70
7a2	81	74	94	86	81	70
7b1	69	48	88	56	66	44
7b2	72	46	88	54	66	42
7c	88	70	81	64	75	62
7d1	22	20	69	60	19	18
7d2	22	20	69	60	22	18
8a	59	54	75	74	59	50
8b	50	44	72	82	50	46
8c	63	62	75	70	59	60
9a	72	72	91	84	69	74
9b	25	30	63	46	25	26
9c1	47	48	69	74	47	48
9c2	25	22	44	32	25	18

***Note*. Exper. = Experimental group.**

Appendix B

Breakdown of Errors on Problems 1 and 2 by Treatment (percentage of students making each error)

	Problem			
	1		2	
Error	C	E	C	E
COMPUTATION ERRORS	31	26	69	78
Numerical calculation	22	20	34	42
Placement of decimal point	19	12	63	60
BLANK	0	0	0	2

***Note*. Because some students made more than one type of error, percentages in specific error categories do not always sum to the percentages in the overall error category. C = Control, E = Experimental.**

Appendix C

Breakdown of Errors on Problem 3 by Treatment (percentage of students making each error)

Error	Control	Experimental
ARITHMETIC OPERATION ERRORS	31	26
Incorrect order of arithmetic operations (e.g., addition before multiplication)	28	18
Incorrect arithmetic operation	3	2
Skipped one or more arithmetic operation(s)	0	6
COMPUTATION ERRORS	56	40
Numerical calculation	44	28
Placement of decimal point	44	36
UNINTERPRETABLE	9	4
BLANK	9	6

Appendix D

Breakdown of Errors on Problem 4 by Treatment (percentage of students making each error)

Error	Control	Experimental
SETUP ERRORS	6	14
Failed to recognize set of items (treated 23 days as 1 day)	0	2
Incorrect arithmetic operation	3	10
Extracted incorrect given information (e.g., 24 days instead of 23)	3	2
COMPUTATION ERRORS	6	20
Numerical calculation	6	16
Decimal point	0	10
BLANK	0	6

Appendix E

Breakdown of Errors on Problem 5 by Treatment (percentage of students making each error)

Error	Control	Experimental
SETUP ERRORS	22	20
Omitted one type of item entirely (notebooks or pens)	3	0
Treated multiple units in a set as one unit (e.g., treated 3 pens as 1 pen)	16	18
Used incorrect arithmetic operation	0	4
Extracted incorrect given information (e.g., 2 pens instead of 3)	6	0
COMPUTATION ERRORS	9	8
Numerical calculation	9	6
Decimal point	0	2
BLANK	3	4

Appendix F

Breakdown of Errors on Problem 6 by Treatment (percentage of students making each error)

Error	Control	Experimental
SETUP ERRORS	75	78
Omitted one type of item entirely (e.g., coke)	6	10
Failed to recognize set of items (treated multiple cokes as 1 coke)	16	18
Used incorrect arithmetic operation	3	2
Extracted incorrect given information (e.g., 4 persons)	3	6
Failed to include Josie in total number of persons	47	44
COMPUTATION ERRORS	22	12
Numerical calculation	19	12
Decimal point	3	2
BLANK	6	4

Appendix G. Breakdown of Errors on Problem 7 by Treatment (percentage of students making each error)

	7a1		7a2		7b1		7b2		7c		7d1		7d2	
ERROR	C	E	C	E	C	E	C	E	C	E	C	E	C	E
SETUP ERRORS	13	16	13	18	22	34	22	32	6	10	72	54	72	54
Misinterpreted type of answer required (cost vs. number of units)	6	12	6	12	3	8	3	8	3	10	0	4	0	8
Failed to recognize set of items (treated set as single item)	0	0	0	0	3	2	3	2	0	0	19	10	16	16
Failed to recognize that change in price applies to each item in set (e.g., treated as change in price for set)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	13	16	13	20
Failed to recognize change in length of time from given information (1 week) to question being asked (1 month): Used 1 week.	3	10	3	10	13	16	13	14	3	4	19	24	19	26
Incorrectly translated length of time from given information (1 week) to question being asked (1 month) (e.g., interpreted as number sold per day and used 28 days per month)	6	4	6	4	3	2	0	2	3	4	3	0	3	0

Note. C = Control, E = Experimental.

Appendix G (continued)

	7a1		7a2		7b1		7b2		7c		7d1		7d2	
	C	E	C	E	C	E	C	E	C	E	C	E	C	E
Failed to recognize that problem asked for a difference in costs (e.g., gave larger cost or both costs)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0	6	59	48	59	50
Used incorrect arithmetic operations	0	4	0	4	3	12	3	12	3	8	6	6	6	10
Extracted incorrect given information (e.g., 14 Snickers instead of 65)	0	0	0	0	3	4	3	2	0	0	0	0	0	0
COMPUTATION ERRORS	3	8	0	6	6	26	6	24	13	16	25	14	25	16
Numerical computation	3	8	0	6	3	4	3	6	9	10	16	8	13	8
Decimal Point	0	0	0	0	3	24	3	20	6	10	9	8	13	12
UNINTERPRETABLE	6	4	6	2	6	6	6	6	6	10	0	8	3	4
BLANK	0	6	0	6	0	12	0	16	0	10	3	20	3	20

Note . C = Control, E = Experimental.

Appendix H
Breakdown of Errors on Problem 8 by Treatment
(percentage of students making each error)

	8a		8b		8c	
Error	**C**	**E**	**C**	**E**	**C**	**E**
SETUP ERRORS	31	36	31	42	16	10
Failed to recognize change in length of time from given information (1 month) to question being asked (1 year)	13	20	13	14	9	4
Incorrectly translated length of time from given information (1 month) to question being asked (1 year) (e.g., interpreted as cost per day and used 365 days)	22	24	22	20	6	4
Failed to recognize that change in length of time applies to one cost (Allstate): Applied time change to both costs or did not apply to either cost	22	18	25	22	0	0
Failed to recognize that problem asked for a difference in costs	9	16	n/a	n/a	13	4
Failed to consider discount	n/a	n/a	0	0	16	10
Failed to apply time change to discount	n/a	n/a	16	32	3	4
Misinterpreted question as asking for difference between discounted and undiscounted cost	n/a	n/a	0	0	3	4
Used incorrect arithmetic operations	9	8	3	4	6	2
Extracted incorrect given information (incorrect value of insurance)	0	0	0	0	0	2
COMPUTATION ERRORS	16	16	9	4	3	2
Numerical calculation	6	12	3	4	3	2
Decimal point	13	4	6	0	0	0
UNINTERPRETABLE	3	6	6	6	6	6
BLANK	6	4	13	8	16	22

Note. C = Control, E = Experimental.

Appendix I

Breakdown of Errors on Problem 9 by Treatment
(percentage of students making each error)

	9a		9b		9c1		9c2	
Error	**C**	**E**	**C**	**E**	**C**	**E**	**C**	**E**
SETUP ERRORS	22	12	38	28	22	26	16	22
Misinterpreted type of answer required (hours vs. amount of money)	9	10	6	0	6	2	9	4
Misinterpreted amount of time given (e.g., used $2.00/minute instead of $2.00/hour)	9	4	9	4	22	18	3	2
Confused different types of units (e.g., confused dollars and hours)	13	4	34	22	0	2	13	16
Included extraneous information (tax)	3	0	9	4	13	6	9	8
Failed to recognize that the problem asked about an amount per week	n/a	n/a	9	6	n/a	n/a	6	4
Failed to consider or apply the number of hours	n/a	n/a	38	24	n/a	n/a	16	18
Failed to use information given or calculated in previous problem (4 weeks or total number of weeks)	n/a	n/a	n/a	n/a	25	16	0	12
Confused different units of time (e.g., hours vs. weeks)	n/a	n/a	n/a	n/a	13	18	n/a	n/a
Used incorrect arithmetic operations	6	6	9	6	3	6	3	2
COMPUTATION ERRORS	3	0	0	12	0	0	0	8
Numerical calculation	3	0	0	12	0	0	0	8
Decimal point	0	0	0	0	0	0	0	0
UNINTERPRETABLE	6	10	19	22	13	6	28	24
BLANK	0	6	19	20	19	20	28	36

***Note*. C = Control, E = Experimental.**