DOCUMENT RESUME

ED 243 719 SE 044 493

AUTHOR	Marshall, Sandra P.
TITLE	Sex Differences in Solving Story Problems: A Study of Strategies and Cognitive Processes. Final Report.
INSTITUTION	California Univ., Santa Barbara. Dept. of Psychology.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	Dec 82
GRANT	NIE-G-80-0095
NOTE	111p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Computation; *Elementary School Mathematics; *Grade 6; Individual Differences; Intermediate Grades; Interviews; *Mathematics Achievement; Mathematics Education; Performance Factors; *Problem Solving; *Sex Differences
IDENTIFIERS	*Error Analysis (Mathematics); *Mathematics Education Research; Story Problems (Mathematics)

ABSTRACT

A 2-year study of mathematics performance of sixth-grade students was conducted to: (1) determine the extent to which previously reported sex differences in California data exist and influence student response; (2) examine the predominant successful and unsuccessful strategies used by sixth-grade students in solving story problems; (3) discover to what extent boys and girls use different strategies in solving story problems; and (4) determine whether there are underlying cognitive differences in typical strategies of boys and girls and to identify these differences if they exist. Particular emphasis was given to analysis and classification of errors made by students solving a variety of mathematical items. The research consisted of analyzing responses from all sixth-grade students on previously existing data and analyzing children's (N=93) problem-solving behaviors obtained from individual interviews. Among the results reported are those indicating that girls were more likely to make errors leading to illogical responses and errors resulting from guessing strategies and that girls performed better on computation and worse than boys on story problems. Appendices include items used in interviews; children's responses to these items; and manuscripts, prepared for journal or book publication, which contain more technical detail about the research. (JN)

ED243719

Sex Differences in Solving Story Problems: A Study of Strategies and Cognitive Processes

NIE-G-80-0095
Final Report

Sandra P. Marshall
Department of Psychology
University of California
Santa Barbara, California 93106

December 1982

SE 044 493

Table of Contents

INTRODUCTION

This report summarizes a two-year study of mathematics performance of sixth-grade students. Particular emphasis has been given to analysis and classification of errors made by students solving a variety of mathematics items. Individual and sex differences are documented here.

The research was conducted in two parts. In the first a large, already-existing data set was analyzed. The data were responses from all sixth-grade children in California. (The students and the instrument are described more fully below.) The second part of the study consisted of individual interviews with 93 sixth-grade students. The problem-solving behaviors of these students were studied in depth.

This report has four sections. In the first, the general outline and objectives for the study are reviewed. The second part describes the results of the secondary analyses of the California data, and the third section describes the results of the interview study. General discussion and conclusions are offered in the final section.

Several appendices are also attached. Appendices C, D, and E are manuscripts prepared for journal or book publication and contain more technical detail about the research described in this report. Appendix A provides the items used in the interview portion of the study, and Appendix B contains the responses of all children to the items of Appendix A.

PROJECT OBJECTIVES

Four objectives were specified for the project. They are described below, followed by brief statements describing success in reaching the objectives.

(1) To determine the extent to which previously reported sex differences in the California data exist and influence student response. The California Assessment Program (CAP) noted in its Annual Report of 1979 that girls generally performed better than boys on computations in the three areas of whole numbers, fractions and decimals. It also reported that boys generally performed better than girls on story problems in the same areas. Statistical confirmation is obtained in the present study. Gender interacts significantly with performance on both story problems and computations.

(2) To examine the predominant successful and unsuccessful strategies used by sixth grade children in solving story problems. Initial identification of unsuccessful strategies comes from examination of the large set of California data. Responses to each item (not just story problems) are examined and common elements of distractor choices are determined. A taxonomy of errors is developed, based upon recognition of problem structure, selection of a problem-solving strategy, and implementation of the chosen strategy.

(3) To discover to what extent boys and girls use different strategies in solving story problems. The taxonomy developed under objective (2) is applied to the California data, and major differences in types of errors made by boys and girls are determined. Further evidence is gathered in an interview study of 93 children, who solved a restricted class of story problems.

(4) To determine whether there are underlying cognitive differences in typical strategies of boys and girls and to identify these differences if they exist. Production systems simulate the performance of children in solving the story problems used in the interview study. The production systems reveal similarities and differences in steps taken to solve problems under three general strategies of problem solving.

STUDY 1: ERRORS ON MULTIPLE-CHOICE TESTS

The first study is an analysis of sixth-grade students' responses to a standardized achievement test.

Population

The data analyzed are achievement test responses from all sixth grade children enrolled in California public schools in 1976-79. The testing instrument is the Survey of Basic Skills: Grade 6, administered in May of each year by the California Department of Education under the California Assessment Program (CAP). Between 275,000 and 300,000 students answered the Survey in each of the four years studied.

Instrument

The Survey of Basic Skills: Grade 6 is designed to assess sixth-grade student performance in reading, writing, spelling, and mathematics. Each subject area is scored separately. The purpose of the test is to estimate average pupil performance on a variety of concepts. As part of its assessment, CAP gathers data about each child's language fluency, primary language at home, age, birthdate, estimated socioeconomic status, and sex.

Scores are reported for school, district, and state levels. The test is constructed in such a way that no student answers all possible items. Individual results are not released because they are not comparable.

There are 16 forms of the test, each containing unique items. An individual form contains 30 multiple-choice items: 10 of mathematics, 8 of reading, 8 of writing, and 4 of spelling.

The following areas comprise the 160 mathematics items:

Area	Number of Items
Number Concepts	28
Whole Numbers	28
Fractions	20
Decimals	20
Geometry	20
Measurement	32
Probability	12

Additional information about the Survey of Basic Skills: Grade 6 can be found in the Annual Report of the California Assessment Program (1979).

General Results

Comparison of Achievement: Computations and Story Problems

For this analysis, attention was restricted to the mathematics areas of whole numbers, fractions, and decimals. Only in these areas were there items of both computation and application (story problems). Of the 160 mathematics items, 68 are relevant here: 41 computations and 27 story problems. Each test form contains either 4 or 6 items from these two categories.

There are two aspects to the analysis of performance on computations and story problems. First, patterns of success and failure by boys and girls on each of the 16 test forms were compared. Second, relationships were investigated between success on either computations or story problems and variables of reading achievement, socioeconomic status, primary language and age.

Patterns of Achievement. Log-linear analysis was used to determine the relationship between gender and patterns of success on story problems and computations. (Details of log-linear analysis may be found in Bishop, Fienberg, and Holland, 1975.) For each test form, a binary classification of success was developed. The classification captured the order of success (i.e., first, second, or third problem correct) for story problems and computations separately. For three story problems, there were eight possible patterns:

+++
++-
+-+
-++
+--
-+-
--+
---

where + denotes successful solving of the item and - denotes unsuccessful. There are eight similar patterns for three computations.

Log-linear analysis provides a means of multi-way contingency table analysis. The factors here are sex, performance on story problems, and performance on computations. Each factor has several levels. Sex has two categories, male and female. Success on story problems has 2**n categories, where n is the number of story problems on a particular form. Success on computations also has 2**n categories, with n being the numbers of computations on the form.

The statistical analysis reveals that a single model predicts the results of all but one test form.

Significant components of that model are interactions between sex and computation success, sex and story-problem success, and story- problem success and computation success. The model does not contain the three-factor interaction.

It is evident from the statistical analysis that one must consider all three factors when predicting student success. Knowing how a student performed on computations is not sufficient to predict his/her performance on story problems. Gender must also be considered.

Other Influences on Problem Solving. Information about students' reading achievement, socioeconomic status, primary language and fluency, and age is collected by the California Department of Education at the time of testing. This information is used in conjunction with the mathematics achievement test to identify subgroups that are more likely to excel or lag in problem-solving success.

The variables socioeconomic status, primary language and age do not appear to be sex-related factors. Boys consistently perform better than girls on story problems, and girls consistently perform better than boys on items of computations (with minor exceptions for older children of both sexes). No subgroup of socioeconomic status of either sex is significantly weaker or stronger than the others. This does not mean socioeconomic status is not highly related to mathematics performance; it means that boys and girls of each SES classification appear to be influenced in approximately the same way.

Similar results are found for age and primary language. Only children identifed by teachers as being fluent in English are included in the language analysis. One intriguing finding is that children whose primary language was Chinese or Japanese (e.g., the language spoken at home) are much more successful in solving both types of items than children who spoke English only.

The comparison of reading achievement and successful solving of computations and story problems indicates that girls with high reading achievement, as measured by the reading test of the Survey of Basic Skills: Grade 6, are relatively weaker in solving story problems than any other group. This finding suggests that factors other than reading influence story problem performance. (Details of this and other analyses are given in Marshall, Note 1, attached as Appendix C.)

Error Analysis

Two issues concerning error analysis are addressed in this study: (1) consistency of distractor selection by boys and girls and (2) classification of predominant errors. Four years of CAP data are used to investigate these issues: 1976-1979. /.sp

Consistency. Again, log-linear techniques are used to analyze student performance. Three factors are of interest: sex, choice of distractor, and year of response. To evaluate consistency, it is necessary to determine whether boys consistently make one error and girls another, regardless of instruction, textbook, or peer group. Thus, comparisons are made of four groups of students: those in the sixth grade in 1976, 1977, 1978, and 1979.

All items of the Survey are used in this study. For each item, a three-factor analysis was carried out. It was hypothesized that interaction between sex and distractor selection would be a significant component of the model and, further, that the three-factor interaction of sex-by-distractor-by-year would not be a significant component. The three-factor interaction would suggest that boys and girls varied in their choice of response in different years and would negate the consistency hypothesis.

A model of best fit was determined for each of the 160 items. One model accounts for eighty percent of all items. There are consistent sex differences for all but 19 of the 180 items.

Classification of errors. A simple model of problem solving is developed as the foundation of the error classification. In this model, there are three stages: (1) identification of the problem type from features within the item; (2) selection of an appropriate strategy for use with the recognized problem type; (3) implementation of the strategy. Errors may be made in any of these stages. In particular, errors may be those of:

1. Attention to spatial and visual cues
2. Attention to verbal cues
3. Selection of inaccurate or inappropriate algorithm
4. Guessing strategies
5. Translation from words to arithmetic expressions
6. Perseverance in use of a chosen algorithm

Errors 1 and 2 are errors of the first stage of problem solving. They reflect inability to identify necessary features of the item. Errors 3 and 4 correspond to difficulties in selecting a strategy or approach for solving a problem. Finally, errors 5 and 6 reflect problems in implementation of a chosen strategy. Details of classification and example items illustrating each category of error are given in Marshall, Note 2, attached as Appendix D.

Both sexes make errors of all types. However, girls appear more likely to err in recognizing relevant features of an item, whether spatial or verbal. Boys, on the other hand, have more difficulty in implementing a chosen strategy.

STUDY 2: ERROR ANALYSIS AND COGNITIVE PROCESSES

The second study is an investigation of sixth-grade students' problem-solving techniques. These techniques were observed during individual interviews.

Population

Ninety-three students were involved in this phase of the study, 49 girls and 44 boys. Sixth grade students in two public elementary schools were asked to participate. A large majority of students in both schools were interviewed. All but two of the sixth graders enrolled in one school took part. Parents of these students requested that their children be excluded. In the second school, one of three teachers did not want his students to be involved in the study. Additionally, parents of five other students requested that their children also be excluded. The resulting sample consisted of approximately 95 percent from the first school and 60 percent of all sixth graders from the second school.

Both schools participating in the study are predominantly white, middle class schools. Each has an enrollment between 350-400 students in grades kindergarten through sixth grade.

The school with the highest rate of participation is traditional. One teacher instructs his/her students in all subject areas. The students are not grouped by ability. Each classroom contains students of varying ability.

The second school operates under the 'pod' system. The school itself consists of several small buildings (pods). Each building houses two or three classrooms. Grades 1-3 form the lower level, and grades 4-6 form the upper level. Students are grouped by ability in English and mathematics at each level. Students may have different teachers for different subject areas, and their classroom peers vary accordingly.

We assume no systematic bias in the samples for the two schools. The excluded classroom from the traditional school did not contain students grouped by ability. Therefore, we assume these students were similar to those who did participate. Students in the second school were grouped by ability, and different classes do reflect differing levels of ability. However, we were able to interview almost all students in the second school and thus did not exclude any group formed by similar capabilities.

INTERVIEW PROCEDURES

Format

Each student participated in a single 45-minute interview. The interviewer was a female graduate student in psychology with previous experience in interviewing students. At the beginning of the interview, the examiner explained to the child that our purpose was to look at the ways sixth-grade children solved story problems. It was particularly stressed that we did not expect the individual to solve each problem. Each student was also told that his/her performance was confidential, and the results would not be given to teachers or parents.

During the first ten minutes of the interview, the student solved a set of story problems in the usual paper-and-pencil format. The purpose of this pretest was twofold. First, we needed an example from each student containing typical work. We wished to estimate the extent to which our interview procedures unduly changed a student's method of problem solving. Comparison of pretest performance with interview performance on selected items allowed us to monitor the effects of the interview. Second, we felt the pretest served as a "warm-up" for the student. Solving problems in a manner similar to usual classroom procedures helped put the student at ease.

Following the pretest, each student was asked to solve at most ten story problems. For five of the problems, the student was asked to describe how he or she would make a plan to solve the problem. Four questions were raised about each item. First, the student was asked to recall as much of the item as possible from memory. Second, the student was asked to specify the question in the item. Third, the student was asked to make a plan about how to solve the item and to describe each step in the plan. Finally, the student was asked to identify the important features in the item that helped him/her develop the plan.

For the remaining five items, the student was asked to imagine a friend could not solve the item and had asked him/her to explain it. The student was given a piece of paper with only the item printed on it. The paper was divided into two parts; on one part the student was able to write if he/she needed to do so. On the other part, the interviewer recorded what the student said and the justifications given for any operations.

Items

Twenty items were used in the interviews. Half of the students responded to a set of ten items in the pretest and then answered questions about a second set of ten in the interview. The remaining students were given the second set of items as a pretest and answered the first set as interview items. The items are given in Table 1 of Appendix A.

In general, no difference in types of errors was noted between pretest behavior and interview behavior. The major errors on the pretest items were also the major errors when the items were used in the interview.

Twelve of the items were matched. That is, six of the items in the first set were matched with six of the items in the second set. These items allowed us to compare performance on the usual paper-and-pencil format with performance under the more unusual interview format. The matched items are indicated in Table 2 of Appendix A.

One item was not included in any analysis (Item 20). This item proved to be too difficult for our sample of sixth-grade children. Most of the students did not attempt to solve it; the remaining ones made idiosyncratic errors that could not be easily interpreted.

Recording of Data

Each interview was tape-recorded. In addition, the interviewer recorded as any observations as possible on special coding sheets. Using the procedures described above, we found that the interviewer could transcribe almost word for word the student's description of his/her problem solving. The tapes served primarily as back-up data recordings.

Since the interviewer recorded most of the student's statements during the interview, complete transcription from the tape recordings was not necessary. Instead, each tape was played, and any information not already recorded by the interviewer was added to the coding sheets. This procedure was quite a bit faster than the usual transcription of interview protocols. In general, the interviewer was able to record almost every statement of the students. At a later time, the final 35 minutes of each tape was played (omitting the 10 minutes used in the pretest), and any additional comments were easily added to the original coding sheets. Few transcriptions were needed.

One particular advantage of our data recording is that we avoided the problem of trying to interpret a student's statement that did not record clearly. In practice, this procedure meant that much of the data were available immediately for analysis.

General Results

Pretest Responses

In general, students found the nineteen items to be difficult. The average proportion of students correct per item was 0.364. Boys were somewhat more successful than girls (0.405 versus 0.3272) but the difference is not statistically significant. Average proportions correct for boys and girls on the two sets of items at the two schools are given below:

	School 1		School 2	
	Set 1	Set 2	Set 1	Set 2
Boys	0.359	0.509	0.333	0.420
Girls	0.343	0.342	0.296	0.329

Proportions are analyzed rather than total test scores because the elimination of one item created an inequality in the number of items on the two test forms. No statistical difference between sex, schools, or forms was found. Thus, for the remaining discussion, no distinction is made between forms or schools.

Interview Responses

It is infeasible to compute average proportion correct for the interview items because not all students attempted to solve all items. Several students attempted only seven, eight, or nine items. In general, their responses to the items were more complete and lengthy than other students, and they simply ran out of time. In these cases, the interviewer used her discretion in selecting items for the student to answer. If the student had shown little difficulty in solving the initial items, the interviewer selected those items with greater difficulty when time was limited. Most students who were unable to solve the items had no time constraint because they merely said they could not solve an item, and the interviewer produced the next item. Little time was spent on items that were too difficult.

It is possible to determine the average probability of success per item for boys and for girls by dividing the numbers of successful individuals by the number of students attempting to solve the item. These probabilities are given for both pretest and interview in Table 3 of Appendix A. Separate probabilities for each sex are also given in this table.

One objective here is to identify characteristic errors of students over the set of story problems. The responses by sex to the items are given in detail in Appendix B. It is clear that there are some errors that are more prevalent in girls' responses than in boys', and there are errors that are more common in boys' responses than in girls'.

For example, consider Item 16:

Mary has 15 hair ribbons. She gives 1/5 of them to Alice. How many hair ribbons does Mary have left?

One noticeable error made by boys is to compute: 1/5 x 15 = 5. These students then continue to solve the problem with the incorrect value of 5. This error, according to the students' statements, is not simply an error in arithmetic but is an error in the students' perception of fraction. A common explanation for this result is that "There are three fifths in 15. If she gave one of those away, she would have two of the fifths left." These students go on to state that a fifth is equal to five and therefore the quantity 'two-fifths' is ten.

The most noticeable error in girls' responses is the subtraction of the fraction 1/5 from the quantity 15. Many more girls than boys attempted this computation, and they gave as a final answer 14 and 4/5 hair ribbons.

There are similar errors on other items. However, it should be pointed out that boys and girls make similar errors, and those identified as characteristic of one sex are not exclusively mistakes of that sex.

Responses to Matched Items

Six of the items on each pretest were matched with six items given in the interview (see Table 2 of Appendix A). A comparison of boys' and girls' rates of success on these items allows us to infer the degree to which our interview procedures interfere with students' problem solving.

Not all students had sufficient time to respond to all items. Only those attempting all twelve problems are included in the analysis. Sixty students did so: 32 boys and 28 girls.

The average numbers of items answered correctly by boys and girls are given below:

		Pretest	Interview
Boys	$\bar{X}$	3.1	3.8
	SD	1.33	1.32
Girls	$\bar{X}$	2.6	2.9
	SD	1.40	1.46

An unweighted-means analysis of repeated measures yields significant F values for tests of sex and of repeated testing. There are significant differences between boys and girls, $F\ (1,58) = 4.567,\ p < .05$, and between the pretest performance and interview performance of all children, $F\ (1,58) = 7.689,\ p < .01$. 0.05 level.

The difference between boys and girls is less than the difference between pretest and interview responses. It was expected that both boys and girls would be somewhat better in their responses to the interview items for two reasons: (1) the interview procedures focused attention on the features of the item and (2) repetition of items having similar formats frequently improves performance. In the interview, a student is more likely to avoid careless errors because he/she is asked to explain each step of the problem-solving process. Many students corrected themselves as they made these explanations.

Post-hoc analyses indicate that the significant findings of the repeated-measures tests are occasioned by the difference in boys' pretest and interview responses. These are significantly different, $t\ (42) = 3.144, < .01$. A similar comparison for girls yields a value of t less than unity. Additionally, no difference is found between boys and girls on the pretest items. We conclude that boys make significant gains in the interview format, either because it reduces their errors of arithmetic or because it enhances their abilities to identify proper strategies for solution.

Sex Differences in Errors

Most of the errors made by students during the interview fall into three classifications: (1) errors in recognizing important information in the statement of the problem; (2) errors in selecting the proper arithmetic operation to perform; or (3) errors of arithmetic facts. Only the first two are of interest here.

Boys and girls appear to be equally likely to recognize or fail to recognize the underlying structure of an item. All children were relatively weak in identifying the general form of a problem. This is apparent in the lack of consistency in students' responses to matched items. Frequently, a student chose one operation on one item in the pretest and selected a different operation for the related item in the interview. This is discussed more fully in Marshall, Note 3, attached as Appendix E.

Three strategies of problem solving are evident in the students' responses. The first approach requires processing of all information in the problem and using it effectively. The second approach focuses only on particular key words in the item, and the choice of operation is directed by the words that are noticed. The third approach is a form of guessing: selecting an arithmetic operation on the basis of the relative size of the numbers in the item.

Consider Item 17 of Table 1, Appendix A. The general structure of this item is given in Figure 1. The difficulty of the item is in determining the amount to be removed. The correct solution depends upon recognition of the fractional component and the transformation of the fractional part of the whole to units of the original quantity. Thus, 1/3 of the cards means 1/3 of the quantity 12, which is 4. This amount is removed from 12 and the resulting 8 is the correct response. Students operating under the first strategy defined above would work the problem similarly.

The second approach focuses on key words. There are two choices of key words in this item: <u>of</u> and <u>have left</u>. Students attending to the word <u>of</u> seek to perform the operation of multiplication. If questioned about the choice, many respond with statements such as "My teacher says OF always means multiply." Thus, the general structure of the problem is never perceived. These students simply multiply 1/3 times 12. The other choice of key words leads to the operation of subtraction. Students focusing on the words "have left" apparently search for two numbers. They find the quantities 1/3 and 12, and they subtract the smaller from the larger, finding an answer of 11 2/3. Although this is an illogical choice (one does not generally have 2/3 of a baseball card), the students are satisfied with the result.

For this particular item, the guessing strategy based upon the size of the numbers usually leads to the response of 1/3 x 12. A student using this strategy makes a response such as, "Oh, . . ., I don't know. . . I guess I would multiply." When asked why he/she multipled, the student is unable to state a reason. Apparently, students realize that when given a simple fraction and a whole number, the most likely operation to be required is multiplication.

A Model of Student Performance

Students exhibited differences in cognitive processing during the interview. One of our research goals was to model these differences. A useful method of modeling individual performance on a complex task is a production system. Such a system is made up of a set of rules, each having two components: condition and action. The condition of a production rule must be met before the action specified in the rule can be taken. The system consists of the set of rules, a mechanism called working memory, and a mechanism for maintaining goals. The set of rules reflects the individual's knowledge of arithmetic and his/her general strategies of problem solving. Working memory contains bits of information gleaned from the problem currently being solved. The goal structure reflects subgoals set by the individual as he/she solves a problem. Computations to be taken or transformations to be made are examples of goals. The conditional statement of a production rule refers to components of working memory or of the goal list.

There are 64 production rules. A subset of the rules is given in Table 4 of Appendix A. The system is programmed in LISP and runs on a VAX computer under the UNIX operating system. The list of rules given in Table 4 is translated from the LISP notation to a more readable form. In practice, each production consists of two LISP statements. The first statement contains the conditional statement. If the evaluation of this statement is positive, the second statement of the rule is executed.

During execution of the system, each production rule has some probability of being called. However, it is inefficient to have the system scan each production rule every time a rule may be needed because many rules will be incapable of firing in the particular situation. For example, if there are no numbers contained in working memory and the immediate goal is the identification of numbers from text, it is a waste of computing (processing) time to scan all productions that specify how two numbers can be combined. Since two numbers do not exist already in working memory, these rules cannot possibly be used at this time. Thus, rather than scanning every rule in the set, only subsets of pertinent rules are evaluated. A rule having a satisfied conditional statement is selected and the appropriate action for the rule is taken. This action generally modifies either working memory or the goal list for the problem. Thus, there exists a new configuration for the system. Processing begins anew, and the system searches for a production rule whose conditional statement is satisfied.

One feature of the current system is that it attempts to find rules that operate on the most recent entry in working memory and on the most recently established goal. Thus, if the goal list contains the subgoals of making a transformation from whole number to fractional form and also of subtracting the transformed number from a second, the transformation goal takes precedence. Once the transformation is complete, the goal is removed from the list and the most immediate goal to be satisfied becomes that of subtraction.

Unlike most production systems that model individual performance, the current system contains several strategies for solving problems. Thus, there exist decision points at which several production rules might be executed, each leading to different solutions. The choice of rule in this case depends upon the probability with which actual students used each strategy.

The current system is designed to model student performance on the pair of problems, Items 7 and 17 (see Table 1 of Appendix A). It also models behavior on other problems requiring fractions and with slight modification can model all but a few of the nineteen problems used in the present study.

Three Strategies

Three general strategies of problem solving are modeled. These correspond to those described in the previous section: (1) recognition of general structure, (2) key word identification, and (3) guessing patterns. These three strategies contain a surprisingly large number of common production rules. The strategies differ only in the inclusion or exclusion of a few important rules. Thus, the same system can model performance by a diverse set of individuals.

Figure 2 illustrates the complete production system as a network of arcs and nodes. Presence of a particular arc means a production rule exists that allows the system to move from one node via the arc to a second node. No individual would have the complete set of arcs and nodes shown in this figure. Rather, some arcs would be missing, indicating that the individual would not be able to employ a strategy requiring the associated productions. There are several points of entry into the system. These are denoted by the unlabeled small circles.

Execution of the system under each of the three strategies depends upon defining the initial state of the system. Under the recognition-of-general-structure strategy, initialization begins at the upper left start position of Figure 2. If the strategy is that of guessing, the initial goal becomes to identify two numbers. If the strategy depends upon determination of key words, clearly the initial goal must be to recognize key words in the problem. Many of the same production rules will then be called, but they will be executed in different order. For example, in the guessing strategy, first the numbers will be identified and relevant features about them noted. The choice of arithmetic operation will depend upon which features are recognized (e.g., large and small numbers or fraction and whole numbers). For the key word strategy, the key word determines the operation, and the numbers to be operated on must then be identified. Choice of operation determines the features of the numbers to be noted. That is, if subtraction is to be the operation, the system must identify both a large and a small number.

CONCLUSIONS

In both studies carried out here, girls were more likely to make errors leading to illogical responses and errors resulting from guessing strategies. However, both boys and girls were weak in solving story problems and could benefit a great deal from further instruction. A large number of children seemed to believe that there were firm rules to apply in solving story problems and that these rules could be applied just as one applies rules of subtraction or multiplication. Girls may be more likely than boys to make such assumptions. Very few students demonstrated ability to think about the problems and discern the underlying structure of the statements. (The manuscript in Appendix D discusses this issue in greater detail.)

Results of the production system simulation show that the three predominant strategies used by boys and girls are related. Many common nodes and arcs exist within the three approaches. The similarities between the strategies and the rules executed for each one suggest that the strategies are developmentally related. The strategy of guessing requires the fewest production rules; no parsing or understanding of the text is needed. The strategy is context-free. The key-word strategy is slightly more complex than the guessing one. The entire text is not processed, but isolated words or phrases are understood and the choice of arithmeic operation depends upon the words that are noticed. Finally, the strategy requiring recognition of the structure of the problem depends upon the entire text, and as such requires the largest number of production rules.

There is some evidence that this hierarchical or developmental relationship exists in sixth grade students. On the very easy problems (see Items 4 and 14 of Table 1), almost all students described the general structure of the problem. They discussed cookies in the box or apples in the basket with confidence, and their justifications of operation referred to the content of the item rather than to key words or numbers. In contrast, on the Items 7 and 17, many of these same students focused on the key words. Thus, the words 'have left' meant subtraction or the word 'of' meant multiply. Little or no reference was made to the situations described in the items or to the units being manipulated (cards or ribbons). Finally, on Items 3 and 13, many of the students could not find key words (or did not attempt to locate them) and they opted to divide or multiply with no justification other than 'it seems right' or 'maybe 3 goes into 24'.

The implication is that when all else fails, the students fall back on comparing the sizes of the numbers in an item. If the item is easy, almost all students can think about the context of the problem and can frame their responses in terms of the context. For more difficult items, there may be too much information for them to process and they ignore many of the contextual references. When the problem is not understood, all contextual clues are ignored and the numerical values alone are processed.

Girls did perform better than boys on computations and worse than boys on story problems. These results are clear in the large California data set. Girls also were slightly more proficient in reading than boys. At first glance, these seem to be contradictory results. If girls are better readers and better at solving computations, why are they weaker in solving story problems?

The results of the student interviews in Study 2 offer one explanation for this discrepancy. Girls who solve computations well are accustomed to selecting one or possibly two rules which they then use in performing their calculations. Proficient solvers access these rules in memory readily and can achieve the correct computation. Such students may attempt to transfer this approach to the solving of story problems. After all, firm reliance on rules has been consistently rewarded in computations. Teachers often introduce methods of problem-solving as if rules can be applied in story problems as well as in computations. Examples of such rules are (1) identify or isolate the numbers in the problems (many students explained carefully to us that their teachers always had them write down all numbers in a problem), or (2) focus on certain words in a problem, with the objective of identifying the necessary operation to perform ("My teacher says OF always means multiply" is a common response).

Since boys tend to score less well on tests of computations than do girls, one might tentatively postulate that boys are not as rule-governed as are girls and that, consequently, they are not as tempted to fit story problems into known rules or algorithms. Further study is necessary to determine if this is the case.

Reference Notes

1. Marshall, Sandra P. <u>Sex Differences in Children's Mathematics Achievement: Solving</u> Computations and Story <u>Problems</u>. Unpublished manuscript, submitted for publication. (Attached as Appendix C.)

2. Marshall, Sandra P. <u>Sex Differences in Mathematics Errors: An Analysis of Distractor Choices</u>, forthcoming in <u>Journal for Research in Mathematics Education</u>. (Attached as Appendix D.)

3. Sandra P. Marshall, Error Analysis and Cognitive Processes, forthcoming in Fennema, E. and Reyes, L. (Eds.), <u>Girls, Women, and Mathematics</u>, Lawrence Erlbaum Associates. (Attached as Appendix E.)

References

Bishop, Y. M., Fienberg, S. E., & Holland, P. W. <u>Discrete multivariate analysis: Theory and practice</u>. Cambridge, Massachusetts: The MIT Press, 1975.

California Assessment Program. <u>Student Achievement in California Schools, 1978-19 Annual Report</u>. Sacramento, California: State Department of Education, 1979.

Table 1

All Items Used In Study 2

Set 1

(1) A girl has a blank piece of paper. She writes eight words on the paper, erases five, and then writes nine more. How many words are on the paper?

(2) Janet spends 2/3 of her allowance on school lunches and 1/6 on entertainment. What part of her allowance is left?

(3) A case of soda contains 24 bottles. If one out of every three bottles is empty, how many full bottles of soda are in the case?

(4) Thirty apples fill a third of a basket. How many apples are in the basket when it is full?

(5) You've been asked to set the table for your family for a week. There are six members of your family, and you have 16 meals together each week. At every meal, you give each family member a knife, a fork, and a spoon. How many utensils do you set in place during the week?

(6) When it left the station, a bus was carrying 40 passengers. It made five stops. At the first two stops, no one got off the bus, but 5 and 7 passengers got on. Eight passengers got off at each of the last three stops, and no one got on. How many people were on the bus after the fifth stop?

(7) Mary has 15 hair ribbons. She gives 1/5 of them to Alice. How many ribbons does Mary have left?

(8) A parking area can accommodate 24 buses. If 5 cars can be parked in the space used by 3 buses, how many cars can be parked in the parking area?

(9) Pam counted 7 heads and 24 legs on her pet hampsters and parakeets. How many of Pam's pets are parakeets and how many are hampsters?

(10) Chris drove from Santa Barbara to Ojai to Ventura and to Santa Barbara without going back to Ojai. Altogether he drove 90 miles. The distance from Santa Barbara to is 26 miles, and the distance from Santa Barbara to Ojai is 33 miles. What is the distance between Ojai and Ventura?

Set 2

(11) A baby came to a staircase. She climbed up five steps, climbed down three steps, and then climbed up six steps and was at the top. How many steps were in the staircase?

(12) When Mike and Sally came home from school they found a chocolate cake. Mike ate half of the cake, and Sally ate a third of the cake that Mike had left. How much of the whole cake was left after Sally ate her piece?

(13) A bus has seats for 48 passengers. If one out of every six seats is empty, how many passengers are on the bus?

(14) There are ten cookies in half a box. How many cookies are in the box when it is full?

(15) You have been chosen as manager of your school band. There are eight members of the band, and they will have 11 practice sessions this year. If you have to put away an instrument, a music folder, and a music stand for each band member after every practice, how many objects will you put away this year?

(16) When it arrived at the elementary school, the bookmobile had 85 books. One student returned 8 books and checked out 11 others. Two more students brought back 4 books each, and another student checked out 6 books. How many books are now in the bookmobile?

(17) John has 12 baseball cards. He gives 1/3 of them to Jim. How many does John have left?

(18) In a grocery display, there are a dozen watermelons. Each melon takes up the same amount of space as 8 oranges. If the grocer decided to replace half the watermelons with oranges, how many oranges would he need?

(19) A group of bicycle riders took a three day trip. On the first day, they traveled 14 miles further than they had originally planned. On the second day, they went 8 miles less than they had planned, and on the third day, they traveled 16 miles more than they had planned. If the bike riders actually traveled a total of 180 miles in 3 days, how many miles had they originally planned to bike?

(20) How many tiles would you need to cover the floor of a room that is 8 feet wide and 12 feet long if every tile is 4 inches on each side?

Table 2

Matched Items for Pretest and Interview

(1) A girl has a blank piece of paper. She writes eight words on the paper, erases five, and then writes nine more. How many words are on the paper?

(11) A baby came to a staircase. She climbed up five steps, climbed down three steps, and then climbed up six steps and was at the top. How many steps were in the staircase?

(3) A case of soda contains 24 bottles. If one out of every three bottles is empty, how many full bottles of soda are in the case?

(13) A bus has seats for 48 passengers. If one out of every six seats is empty, how many passengers are on the bus?

(4) Thirty apples fill a third of a basket. How many apples are in the basket when it is full?

(14) There are ten cookies in half a box. How many cookies are in the box when it is full?

(5) You've been asked to set the table for your family for a week. There are six members of your family, and you have 16 meals together each week. At every meal, you give each family member a knife, a fork, and a spoon. How many utensils do you set in place during the week?

(15) You have been chosen as manager of your school band. There are eight members of the band, and they will have 11 practice sessions this year. If you have to put away an instrument, a music folder, and a music stand for each band member after every practice, how many objects will you put away this year?

(6) When it left the station, a bus was carrying 40 passengers. It made five stops. At the first two stops, no one got off the bus, but 5 and 7 passengers got on. Eight passengers got off at each of the last three stops, and no one got on. How many people were on the bus after the fifth stop?

(16) When it arrived at the elementary school, the bookmobile had 85 books. One student returned 8 books and checked out 11 others. Two more students brought back 4 books each, and another student checked out 6 books. How many books are now in the bookmobile?

(7) Mary has 15 hair ribbons. She gives 1/5 of them to Alice. How many ribbons does Mary have left?

(17) John has 12 baseball cards. He gives 1/3 of them to Jim. How many does John have left?

Table 3

Proportion Correct of Boys and Girls on Pretest and Interview Items

ITEM	PRETEST: BOYS	PRETEST: GIRLS	INTERVIEW: BOYS	INTERVIEW: GIRLS
(1) paper	.95	.96	1.00	.95
(2) allowance	.29	.15	.14	.05
(3) hampsters	.24	.12	.30	.15
(4) busstop	.62	.46	.68	.35
(5) parking	.48	.27	.62	.16
(6) soda	.33	.27	.18	.24
(7) ribbons	.33	.12	.13	.04
(8) driving	.38	.31	.30	.17
(9) apples	.81	.56	.74	.52
(10) meals	.38	.27	.17	.33
(11) baby	.43	.48	1.00	.88
(12) cake	.09	.22	.25	.09
(13) bikes	.13	.04	.50	.32
(14) bookmobile	.26	.30	.75	.43
(15) melons	.65	.39	.80	.75
(16) busseats	.13	.29	.45	.19
(17) baseball	.22	.22	.50	.30
(18) cookies	.96	1.00	1.00	1.00
(19) band	.30	.13	.59	.62

*This table contains only responses of individuals who had sufficient time to attempt the items. The items are given in full in Table 1 of Appendix A.

Table 4

A Portion of the Production System

RULE	CONDITION	ACTION
P1	If strategy is key-word dependent, and working memory contains neither key words nor numbers	Locate a key word in text.
P2	If key word is HAVE LEFT	Select operation subtraction.
P3	If operation is subtraction, and there are two numbers in working memory but they have not been classified,	Identify large and small values.
P4	If operation is subtraction, and working memory contains a large and a small number,	Subtract small value from the large one.
P5	If an operation has been carried out, and if the strategy is based on key words,	Return the result of the operation as the answer.

Figure 1

A general model of the underlying structure of Item 17.

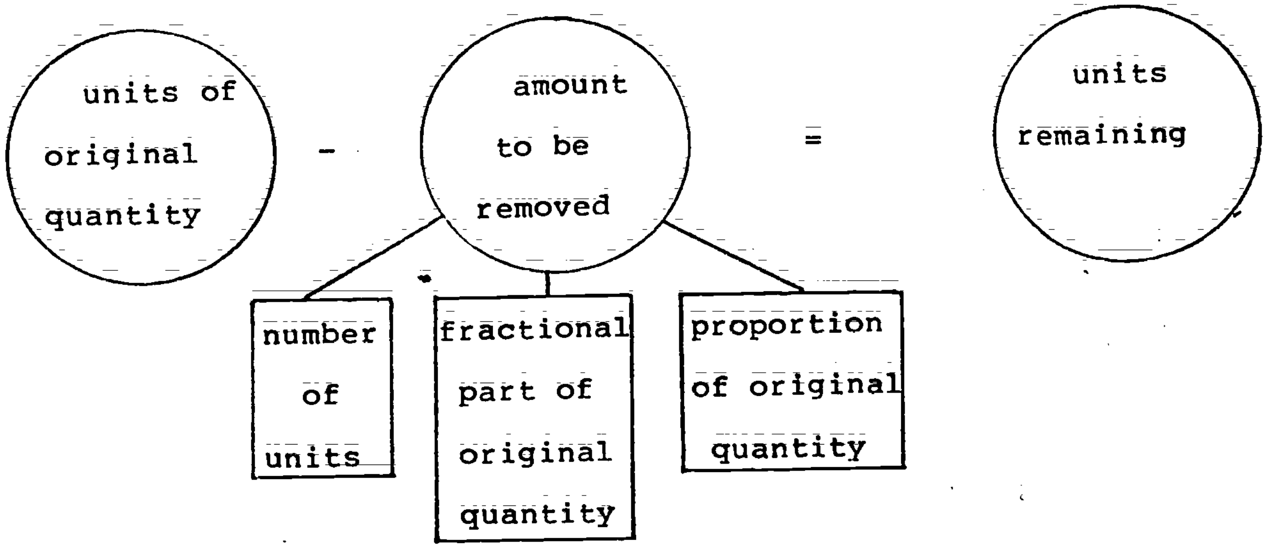
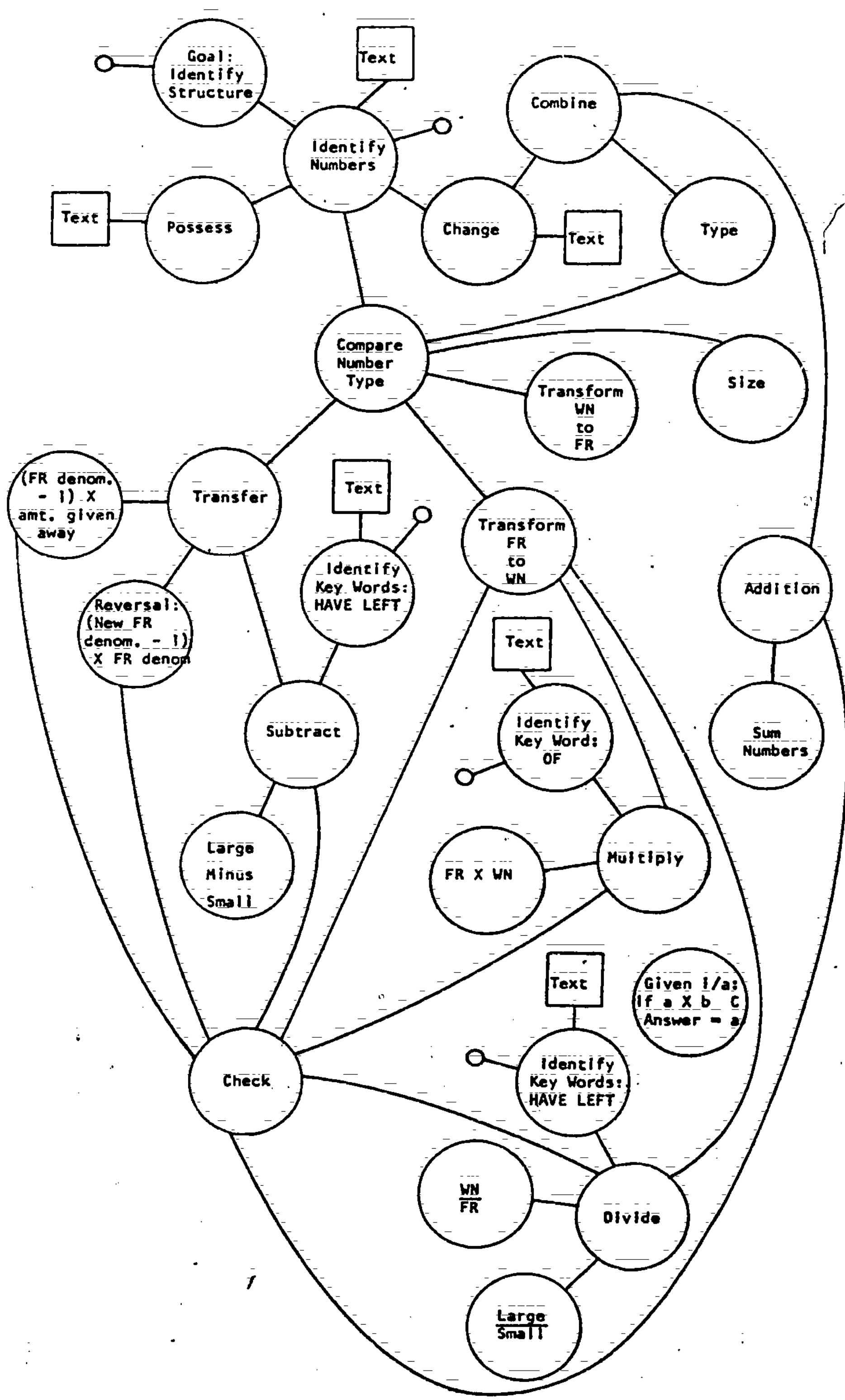


Figure 2

Graph network corresponding to the full production system.
production system.



Student Responses to Pretest and Interview Items

1. PROBLEM: A girl has a blank piece of paper. She writes eight words on the paper, erases five, and then writes nine more. How many words are on the paper?

CORRECT STRATEGIES:

(1) One strategy was used predominantly: 8 - 5 + 9 = 12.
INTERVIEW: 22 boys, 19 girls
PRETEST: 20 boys, 21 girls

(2) Drew 8 marks; crossed out 5; drew 9 more.
INTERVIEW: 0 students
PRETEST: 0 boys, 1 girl

(3) One student read 'eighty' for 'eight' and then solved the item correctly.
INTERVIEW: 0 students
PRETEST: 0 boys, 1 girl

(4) No work shown; correct answer.
INTERVIEW: 0 students
PRETEST: 0 boys, 2 girls

INCORRECT STRATEGIES:

(1) Some students added all numbers.
INTERVIEW: 0 boys, 1 girl
PRETEST: 1 boy, 1 girl

(2) Some students did not have time to solve the problem.
INTERVIEW: 1 boy, 3 girls
PRETEST: 0 students

2. PROBLEM: Janet spends 2/3 of her allowance on school lunches and 1/6 on entertainment. What part of her allowance is left?

CORRECT STRATEGIES:

(1) Only one strategy was used. Students first added the two fractions "to find out how much she spent". Then they subtracted from 6/6 ("the whole allowance").
INTERVIEW: 3 boys, 1 girl
PRETEST: 6 boys 4 girls

INCORRECT STRATEGIES:

(1) keyword: HOW MANY LEFT? Many students elected to subtract because of the keyword. All subtracted the smaller fraction from the larger.
INTERVIEW: 12 boys, 9 girls
PRETEST: 10 boys, 12 girls

(2) addition: first step of correct solution. Add the two fractions.
INTERVIEW: 3 boys, 2 girls
PRETEST: 2 boys, 1 girl

(3) Total not given: Four students said the problem could not be solved because the amount of allowance was not given. The two boys attempted to solve it anyway. One made up a quantity ($.90) and worked with it. The other reached the sum 5/6 and then said that "You would subtract that from the whole thing, if you knew what it was." The girls simply said it couldn't be done.
INTERVIEW: 3 boys, 4 girls
PRETEST: 0 boys, 1 girl

(4) Division: divide 1/6 / 2/3
INTERVIEW: 0 students
PRETEST: 0 boys, 1 girl

(5) Some students ran out of time before attempting this one.
INTERVIEW: 2 boys, 4 girls
PRETEST: 0 students

(6) On the pretest, two students left the item blank. Interview response was 'don't know'.
INTERVIEW: 0 boys, 3 girls
PRETEST: 2 boys, 6 girls

(7) Unique or unclear responses.
INTERVIEW: 0 students
PRETEST: 1 boy, 1 girl

3. PROBLEM: A case of soda contains 24 bottles. If one out of every three bottles is empty, now many full bottles of soda are in the case?

CORRECT STRATEGIES:

(1) Most students correct on the item solved it by dividing 24 by 3 and then subtracting the resulting 8 from 24 (answer 16). Two students on pretest did not show first step but did 24 - 8 = 16 only. One student on interview changed representation to multiplication rather than division for the first step and to addition rather than subtraction for the second step.
INTERVIEW: 3 boys, 5 girls
PRETEST: 6 boys, 5 girls

(2) Correct by drawing: 3 columns of 8 slashes, or 24 circles with every third one marked out.
INTERVIEW: 0 students
PRETEST: 0 boys, 2 girls

(3) One student wrote the series of numbers 3,6,9,12,...,24. Under each of these he wrote a '2'. He counted the number of '2's for the correct answer.
INTERVIEW: 0 students
PRETEST: 1 boy, 0 girls

(4) One student divided 24 by 3 and obtained 8. He then multiplied 8 x (3 - 1) = 16. This student showed evidence of transfer because he said he could now solve the bus seats problem.
INTERVIEW: 1 boy, 0 girls
PRETEST: 0 students

INCORRECT STRATEGIES:

(1) Most of the students made the same error—they divided 24 by 3 and considered that to be the answer. When asked why they divided, 5 said "because 1 out of every 3 was empty". (2 boys, 3 girls). 4 said they divided "to see how many are full" (3 boys, 1 girl). 2 (boys) said "to see how many are empty".
INTERVIEW: 10 boys, 8 girls
PRETEST: 8 boys, 13 girls

(2) 1 student multiplied 24 x 1/3 = 8. Could not say why multiplied.
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 students

(3) Several students subtracted 3 from 24.
One pretest response 21 with no work.
INTERVIEW: 3 boys, 3 girls
PRETEST: 1 boy, 0 girls

(4) Miscellaneous errors: each response is unique.
INTERVIEW: 1 boy, 1 girl
PRETEST: 0 boys, 2 girl

(5) Several students subtracted 3 - 1 = 2 (since took one away) and then divided 24 by 2 to get 12.
INTERVIEW: 3 boys, 1 girl
PRETEST: 0 boys, 1 girl

(6) Other responses of 8: one student transformed the problem into multiplication (3 x 8 = 24) and answer was 8.
Two students on pretest responded 8 with no work.
INTERVIEW: 0 students
PRETEST: 2 boys, 1 girl

(7) Several students could not solve the problem in the interview.
Pretest response was blank.
INTERVIEW: 1 boy, 2 girls
PRETEST: 3 boys, 2 girls

(8) Some students ran out of time on interview.
INTERVIEW: 1 boy, 2 girls
PRETEST: 0 students

4. PROBLEM: Thirty apples fill a third of a basket. How many apples are in the basket when it is full?

CORRECT STRATEGIES:

(1) Multiplication: 30 x 3 = 90. Recognize that 1/3 equals 30. State need to find what equals 3/3.
INTERVIEW: 8 boys, 5 girl
PRETEST: 12 boys, 7 girls

(2) Multiplication and addition: (2 x 30) + 30 = 90. Recognize that 1/3 of basked is accounted for and find what is needed for remaining 2/3. Look for explicit mention of 2/3 in this strategy.
INTERVIEW: 5 boys, 3 girls
PRETEST: 0 students

(3) Addition: 30 + 30 + 30. Frequent stress on keywords "how many".
INTERVIEW: 1 boy, 2 girls
PRETEST: 2 boys, 0 girls

(4) Division: 30 divided by 1/3. Students using this strategy could not explain why they did so.
INTERVIEW: 1 boy, 1 girl
PRETEST: 0 students

(5) Solving by analogy: 1/3 = 30; 2/3 = 60; 3/3 - 90.
INTERVIEW: 0 students
PRETEST: 1 boy, 2 girls

(6) Solution by drawing: drew basket divided into thirds and then multiplied or added for correct answer.
INTERVIEW: 1 boy, 1 girl
PRETEST: 0 students

(7) 1 student knew the answer was 90 but had no idea how he reached that response: "I can't tell you. I have it in my head." Pretest responses had correct answer but showed no work.
INTERVIEW: 1 boy, 0 girls
PRETEST: 2 boys, 4 girls

INCORRECT STRATEGIES:

(1) Several students answered 60. They either added 30 + 30 = 60 or responded 60 and showed no work.
INTERVIEW: 0 boys, 2 girls
PRETEST: 1 boy, 1 girl

(2) Combination of numbers: 30 and 1/3. Two girls added, because 30 and 1/3. Two girls added, because "it says how many are in "it says how many are in the basket when it's full." Two girls and one boy subtracted 1/3 from 30. Four multiplied. Two boys obtained 10; one girl answered 91; one girl responded 30/3.
INTERVIEW: 2 boys, 4 girls
PRETEST: 1 boy, 2 girls

(3) 2 (interviews) responded 120. One appeared to believe that 30 x 3 is 120. The other multiplied 4 times 30, implying that she interpreted thirds as quarters. One pretest response added 30 + 90.
INTERVIEW: 0 boys, 2 girls
PRETEST: 1 boy, 0 girls

(4) 2 students guessed the response of 50. One girl thought it had something to do with 1/2 and reached 50. One boy saw it as an addition problem and added 20 + 30.
INTERVIEW: 1 boy, 1 girl

(5) Miscellaneous errors: unique solutions.
INTERVIEW: 1 boy, 0 girls
PRETEST: 0 boys, 4 girls

(6) Some students had no idea and did not try to solve it. Corresponding response on pretest was to leave item blank.
INTERVIEW: 2 boys, 2 girls
PRETEST: 2 boys, 5 girls

(7) One student had no time to attempt solution.
INTERVIEW: 0 students
PRETEST: 0 boys, 1 girl

5. PROBLEM: You've been asked to set the table for your family for a week. There are six members of your family, and you have 16 meals together each week. At every meal, you give each family member a knife, a fork, and a spoon. How many utensils do you set in place during the week?

CORRECT STRATEGIES:

(1) 16 x 3 = 48 followed by 48 x 6 = 288. First found # utensils used by one person for the week and then multiplied by 6 because there are 6 members in the family. Students using this strategy could not say what the result of the first operation meant. One student (f05) multiplied 16 x 6; labeled 96 spoons, 96 forks and 96 knives.
INTERVIEW: 1 boy, 3 girls
PRETEST: 0 boys, 1 girl

(2) As above but with reversed order of operations: (16 x 6); (96 x 3)
INTERVIEW: 1 boy, 0 girls
PRETEST: 2 boys, 3 girls

(3) The second strategy begins with 3 x 6 = 18 (# utensils used at any meal). All students using this strategy specified this was # used at one meal. Then 18 x 16 = 288 for 16 meals.
INTERVIEW: 2 boys, 4 girls
PRETEST: 6 boys, 4 girls

(4) There were 4 partial solutions in the interview.

a. 3 x 6 = 18 (# utensils needed in 1 meal).
INTERVIEW: 3 boys, 0 girls
PRETEST: 0 boy, 1 girl

b. 6 x 16 = 96 (# meals all members attended)
One girl added step of 16 x 3, but did not know the relationship between the two answers; one boy divided 96 by 3.
INTERVIEW: 1 boy, 2 girls
PRETEST: 0 boys, 2 girls

c. 16 x 3 = 48 (# utensils for 1 person all week)
INTERVIEW: 5 boys, 2 girls
PRETEST: 4 boys, 6 girls

(5) Many students had interference from outside information. They attempted to use the number 7 (for seven days in a week) in their solutions. Most of the solutions were unique combinations of 3,7,6,16 or any 2 or 3 of these numbers.
INTERVIEW: 3 boys, 8 girls
PRETEST: 3 boys, 1 girl

(6) One student divided 16 by 3, then added that to 6, answer was 11 1/3.
INTERVIEW: 1 boy, 0 girls
PRETEST: 0 students

(7) Addition: One added (6 + 16) x 3 = 66 'because it's an addition problem'. The other
Another took (3 x 6) + 16. He recognized 18 utensils per meal and 16 meals. Therefore, he added. A third took 3 x 6 = 18 (the # utensils for one night). He then found 3 x 16 = # of utensils used in one week. Add these two products. Two other students added 16 + 3 or 10 + 20 .
INTERVIEW: 2 boys, 3 girls
PRETEST: 2 boys, 1 girl

(8) Final error was to omit spoon (set 2 things only, k and f) and to multiply 2 x 16 for the # utensils.
INTERVIEW: 1 boy, 0 girls
PRETEST: 0 students

(9) Miscellaneous errors: solutions are unique.
INTERVIEW: 1 boy, 0 girls
PRETEST: 0 boys, 1 girl

(11) Several students could not solve the problem or left it blank on pretest.
INTERVIEW: 0 boys, 1 girl
PRETEST: 4 boys, 7 girls

(12) Two students ran out of time and did not attempt the item.
INTERVIEW: 0 boys, 2 girls
PRETEST: 0 students

6. PROBLEM: When it left the station, a bus was carrying 40 passengers. It made five stops. At the first two stops, no one got off the bus, but 5 and 7 passengers got on. Eight passengers got off the bus at each of the last three stops, and no one got on. How many people were on the bus after the fifth stop?

CORRECT STRATEGIES:

General comments: The number of steps did not appear to trouble the students at all. The easy operations (addition and subtraction) were quickly identified and almost all students said specifically that getting on meant add and getting off meant subtract. Most of the students (even the incorrect ones) felt very confident about the item, and they found it relatively easy.

(1) Most students first added 5 + 7 = 12. Second step: 40 + 12 = 52. Third step: 3 x 8 = 24. Fourth step: 52 - 24 = 28.
INTERVIEW: 8 boys, 2 girls
PRETEST: 9 boys, 5 girls

(2) Some students bypassed the first step (mental arithmetic) and began by adding 12 and 40. The remaining steps are the same as above.
INTERVIEW: 2 boys, 4 girls
PRETEST: 3 boys, 6 girls

(3) Some students subtracted 8 three times rather than using multiplication (3 x 8 = 24) and subtracting the product.
INTERVIEW: 3 boys, 1 girl
PRETEST: 1 boy, 1 girl

INCORRECT STRATEGIES:

(1) The predominant error was subtraction of 8 rather than 24. The remaining steps were correctly executed.
INTERVIEW: 5 boys, 4 girls
PRETEST: 5 boys, 7 girls

(2) Some students did the first step as mental arithmetic. They then subtracted 8 rather than 24 from the total of 52.
INTERVIEW: 1 boy, 3 girls
PRETEST: 1 boy, 2 girls

(3) One student added the arabic numbers in the problem (40, 5, 7) because "they stand out".
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 students

(4) One approach was to subtract (12 − 8) and add + 40.
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 boys, 1 girl

(5) One student answered 44 with no work.
INTERVIEW: 0 students
PRETEST: 0 boys, 1 girl

(6) Miscellaneous errors. Each response was a unique solution.
INTERVIEW: 0 boys, 4 girls
PRETEST: 0 students

(7) Three students used addition only (answer = 52). Another added 52 + 8.
INTERVIEW: 0 students
PRETEST: 2 boys, 1 girl

(8) Miscellaneous errors in subtraction. Each solution is unique. Formulation of problem is incorrect in both cases.
INTERVIEW: 0 students
PRETEST: 0 boys, 2 girls

(9) Four students did not have enough time to attempt the item.
INTERVIEW: 4 boys, 3 girls
PRETEST: 0 students

7. PROBLEM: Mary has 15 hair ribbons. She gives 1/5 of them to Alice. How many ribbons does Mary have left?

CORRECT STRATEGIES:

(1) The most popular correct strategy transformed the problem into whole number arithmetic (3 x 5 = 15). Second step was subtracting 3 from 15 = 12.
INTERVIEW: 2 boys, 1 girl
PRETEST: 2 boys, 2 girls

(2) Same strategy as (1) but using fraction multiplication: 1/5 x 15 = 3 and 15 − 3 = 12.
INTERVIEW: 1 boy, 0 girls
PRETEST: 2 boys, 1 girl

(3) Mental arithmetic in part or in toto.
INTERVIEW: 0 boys, 0 girls
PRETEST: 3 boys, 0 girls

INCORRECT ARITHMETIC BUT CORRECT LOGIC:

(1) Several students had correct logic but responded with 10 rather than 12. The strategies used paralleled those described above. Error: 1/5 of 15 = 5.
INTERVIEW: 4 boys, 1 girl
PRETEST: 4 boys, 2 girls

(2) A second way of reaching 10 was through drawing but using erroneous arithmetic ("there are 3 5's in 15", so 1/5 is 5")
INTERVIEW: 2 boys, 0 girls
PRETEST: 0 boys, 1 girl

INCORRECT STRATEGIES

(1) 1 (interview) responded "its a times problem". He then multiplied 15 x 1/5. Another interview was similar. Four pretests also have this form. One pretest had 1/5 divided by 15 = 3. Sign confusion?
INTERVIEW: 1 boy, 2 girls
PRETEST: 1 boys, 6 girls

(2) 3 students responded (pretest) with "1/5 of 15 = 3" with no multiplication sign.
INTERVIEW: 0 boys, 1 girl
PRETEST: 2 boys, 0 girls

(3) Some students had difficulty deciding whether to divide 15 by 1/5 or to multiply 15 by 1/5. Several of these switched back and forth before selecting one operation.
INTERVIEW: 1 boy, 1 girl
PRETEST: 1 boy, 1 girl

(4) Some students bypassed the fraction arithmetic and divided 15 by 5 or 3. A common explanation was "to find how many she gave to Alice". A second explanation was "to find how many 5's in 15".
INTERVIEW: 2 boys, 2 girls
PRETEST: 0 students

(5) Several students were unable to solve the item. Corresponding pretests were blank.
INTERVIEW: 1 boy, 3 girls
PRETEST: 4 boys, 10 girls

(6) Strategy of subtracting 1/5 from 15. Includes arithmetic error in subtraction. All students in this category treated 1/5 as the number to be subtracted (not perceiving it represents a proportion of the total).
INTERVIEW: 7 boys, 9 girls
PRETEST: 1 boy, 2 girls

(7) Subtraction strategy: recognition that 1/5 represents some other quantity but unable to do computations.
INTERVIEW: 2 boys, 2 girls
PRETEST: 0 students

(8) Miscellaneous errors: each response is unique.
INTERVIEW: 0 boys, 1 girl
PRETEST: 2 boys, 1 girl

8. PROBLEM: A parking area can accommodate 24 buses. If 5 cars can be parked in the space used by 3 buses, how many cars can be parked in the parking area?

CORRECT STRATEGIES:

(1) The main feature of this item is that very few students know why they performed any of the operations. There is a single correct strategy used by most students: 24 / 3; 8 x 5 = 40.
INTERVIEW: 11 boys, 3 girls
PRETEST: 8 boys, 5 girls

(2) Multiplication rather than division: 3 x 8 = 24; 8 x 5 = 40
INTERVIEW: 1 boy, 0 girl
PRETEST: 1 boy, 1 girl

(3) Answer = 40; no work shown.
INTERVIEW: 0 boys, 0 girls
PRETEST: 1 boy, 0 girls

Students' Explanations of Correct Solutions:

(1) Only 1 student could explain precisely that one divides 24 by 3 to find out how many parking spaces are available for 3 buses. (girl)

(2) Three other students were close: one said "how many buses...", a second said "how many buses can fit", and the third said "how many 3 ..." (all boys)

(3) Two responded in terms of cars: "5 cars can fit in 3 spaces" and "want # of cars that can fit in bus space" (both boys)

(4) One apparently saw the problem as a multiplication problem that required an initial step of division because "you can't just say 24 x 5" (girl)

(5) One had no idea why he divided (boy)

(4) One alternate strategy was used. This student multiplied 24 x 5 = 120 and then specified division by 3.
INTERVIEW: 1 boy, 0 girls
PRETEST: 0 students

(5) Another strategy was 3/5 = 6/10 = 9/15 ...24/40
INTERVIEW: 0 students
PRETEST: 0 boys, 1 girl

INCORRECT STRATEGIES:
There was an obvious attempt in the interviews to classify this problem by operation. We got no keyword responses — no child said that any word prompted the choice of operation.

(1) addition: four students attempted to add in the interview. (Addition was never used in the pretest).
INTERVIEW: 1 boy, 4 girls
PRETEST: 0 students

(2) division: several students divided in the interview. Most specified that this problem would be solved by division, and they formulated the problem as 24/3 = 8. One student divided 24 by 5.
INTERVIEW: 0 boys, 4 girls
PRETEST: 1 boy, 2 girls

(3) multiplication: two girls in the interview attempted to multiply 5 x 3 = 15. Neither knew why. One then divided 15 into 24 for the final answer. One boy classified the problem as multiplication and multiplied 24 x 2 = 48 (2 comes from 5 - 3). Multiplication was the most frequent error on the pretest. Five students reached answers of 120 (24 x 5 or 5 x 3 x 8). One multiplied 24 x 5 x 3, one multiplied 24 x 10, one multiplied 24 x 3 = 72 and then divided 72 by 5.
INTERVIEW: 3 boys, 6 girls
PRETEST: 8 boys, 9 girls

(4) subtraction: 24 - 3 = 21.
INTERVIEW: 2 boys, 0 girls
PRETEST: 1 boy, 2 girls

(5) Miscellaneous errors--no work given in explanation.
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 boys, 1 girl

(6) Attempted to draw slots and mark out every third but erred.
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 students

(7) Two students in the interview had no idea how to solve. Similar pretest response was to leave blank.
INTERVIEW: 2 boys, 0 girls
PRETEST: 1 boy, 5 girls

(8) Some students did not have enough time in the interview to attempt the item.
INTERVIEW: 2 boys, 4 girls
PRETEST: 0 students

9. PROBLEM: Pam counted 7 heads and 24 legs on her pet hampters and parakeets. How many of Pam's pets are parakeets and how many are hampsters?

This item does not lend itself to strategies in the same way that other items do. However, there are some obvious similarities in the ways children attempted the item.

CORRECT STRATEGIES:

(1) Some students answering correctly first began with the largest number of hampsters possible (divide 4 into 24 = 6). Since needed 1 more head (or some parakeets) they then "dropped down one" hampster, that is considered 5 hampsters.
INTERVIEW: 1 boy, 1 girl
PRETEST: 0 boys, 1 girl

(2) Several students started by reasoning that 4 and 3 are a likely combination of 7. Most then tried 4 hampsters and 3 parakeets. Making adjustments, they reached correct answers.
INTERVIEW: 4 boys, 1 girl
PRETEST: 0 students

(3) Correct solution by drawing.
INTERVIEW: 0 students
PRETEST: 1 boy, 1 girl

(4) Correct response; no work given.
INTERVIEW: 0 students
PRETEST: 1 boy, 0 girls

(5) Analytic approach: 5 x 4= 20; 2 x 2 = 4; total = 24; therefore 5 hampsters and 2 parakeets.
INTERVIEW: 2 boys, 1 girl
PRETEST: 3 boys, 1 girl

INCORRECT STRATEGIES:

(1) Several students were close in their calculations and were concentrating on the content of the problem—trying to adjust heads and legs in order to come up with the required combination. These students had the right idea but got confused.
INTERVIEW: 2 boys, 0 girls
PRETEST: 0 boys, 1 girl

(2) One approach was to assign equal numbers of legs to parakeets and hampsters. This manifested itself in the division of 24 by 2. Numbers of each pet were not given.
INTERVIEW: 1 boy, 3 girls
PRETEST: 0 boys, 2 girls

(3) Most interview incorrect students attempted division. Most popular attempt was division of 24 by 7.
INTERVIEW: 2 boys, 3 girls
PRETEST: 3 boys, 3 girls

(4) A second division was dividing 24 by 4 and then dividing the resulting 6 by 2 for a response of 3.
INTERVIEW: 3 boys, 1 girls
PRETEST: 0 students

(5) Partial solution: response of either 5 hampsters or 2 parakeets but not both.
INTERVIEW: 0 boys, 1 girl
PRETEST: 1 boy, 0 girls

(6) 1 student added 24 + 7.
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 students

(7) Some students subtracted 24 - 7 = 17.
INTERVIEW: 0 boys, 2 girls
PRETEST: 0 boys, 1 girl

(8) One student multiplied 24 x 7 = 168.
Another student calculated 2 x 4 = 8; 4 x 4 = 16; 16 + 8 =24.
INTERVIEW: 0 students
PRETEST: 0 boys, 2 girls
f03,F39

(9) On the pretest, several students made incorrect guesses but gave no indication of how they achieved their answers.
INTERVIEW: 0 students
PRETEST: 3 boys, 4 girls

(10) Many students said they could not solve it and asked to go to next problem. Corresponding response on pretest was to leave blank.
INTERVIEW: 6 boys, 5 girls
PRETEST: 7 boys, 11 girls

(11) Several students ran out of time on the interview.
INTERVIEW: 3 boys, 3 girls
PRETEST: 0 students

(12) 1 student was solving (generally incorrectly) and got prompt from experimenter that led to correct response. Her solution is not considered here.

(13) Miscellaneous errors. Each response is unique.
INTERVIEW: 0 students
PRETEST: 3 boys, 0 girls

10. PROBLEM: Chris drove from Santa Barbara to Ojai to Ventura and returned to Santa Barbara without going back to Ojai. Altogether he drove 90 miles. The distance from Santa Barbara to Ventura is 26 miles, and the distance from Santa Barbara to Ojai is 33 miles. What is the distance between Ojai and Ventura?

CORRECT STRATEGIES:

(1) There was a predominant correct strategy. Students first added 26 + 33 = 59 ("because that was two of the three places he went"). Then they subtracted 59 from 90.
INTERVIEW: 5 boys, 4 girls
PRETEST: 7 boys, 8 girls

(2) A second strategy subtracted: 90 - 26 = 64. 64 - 31 = 31.
INTERVIEW: 0 students
PRETEST: 1 boy, 0 girls

(3) Two students had the correct logic but could not carry out the arithmetic successfully. One put the three numbers down the three numbers down vertically and indicated vertically and indicated subtraction but reached an answer of 43. He was very clear in his explanation of the item however. A second also knew the logical solution but wasn't quite sure how to carry out the second step. He said "you would have to keep adding 1 miles until you reached 90".
INTERVIEW: 2 boys, 0 girls
PRETEST: 0 students

INCORRECT STRATEGIES:

(1) addition: this was the most popular response. 26 + 33 = 59
INTERVIEW: 4 boys, 6 girls
PRETEST: 3 boys, 4 girls

(2) subtraction: 33 - 26 = 7. "Ventura is 26 miles away, and Ojai is 33 miles away, so take the distance between. That means subtract".
INTERVIEW: 8 boys, 3 girls
PRETEST: 6 boys, 6 girls

(3) Two students responded 7 with no work.
INTERVIEW: 0 students
PRETEST: 0 boys, 2 girls

(4) addition: 90 + 33 + 26 = 149. Add all the numbers because "it's an addition problem".
INTERVIEW: 1 boy, 3 girls
PRETEST: 0 students

(5) Other additions: 33 + 26 = 59. 59 + 31 = 90. Answer is 90. Or, 26 + 33 = 59; 59 + 26 = 85.
INTERVIEW: 0 students
PRETEST: 2 boys, 0 girls

(6) subtraction: 90 - 26 or 90 - 33. Each of these was used by one student. The student using the former then added 33 to the difference. The difference of the second strategy was left as the answer.
INTERVIEW: 0 boys, 2 girls
PRETEST: 1 boy, 1 girl

(7) addition and division: one student first added 33 + 26. She then divided 59 into 90 because "usually if you have a lower # and a higher # you divide into it."
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 students

(8) Several students could not solve the problem in the interview. Pretest response was blank.
INTERVIEW: 3 boys, 4 girls
PRETEST: 1 boy, 5 girls

11. PROBLEM: A baby came to a staircase. She climbed up five steps, climbed down three steps, and then climbed up six steps and was at the top. How many steps in the staircase?

CORRECT STRATEGIES:

(1) Predominant strategy: 5 - 3 = 2; 2 + 6 = 8. Almost every child could describe problem accurately and concretely. All students were able to tell why they added or subtracted.
INTERVIEW: 18 boys, 18 girls
PRETEST: 10 boys, 11 girls

(2) An alternative strategy was to add 5 + 6 and then subtract 3.
INTERVIEW: 0 students
PRETEST: 1 boy, 0 girls

(3) Several students drew their responses and reached the number 8.
INTERVIEW: 0 boys, 3 girls
PRETEST: 0 students

INCORRECT STRATEGIES:

(1) addition: sum all numbers in the problem. 5 + 3 + 6.
INTERVIEW: 0 boys, 1 girl
PRETEST: 7 boys, 6 girls

(2) addition: add only steps going up: 5 + 6 = 11
INTERVIEW: 0 boys, 2 girl
PRETEST: 3 boys, 3 girls

(3) addition: 5 + 3 = 8
INTERVIEW: 0 students
PRETEST: 1 boy, 0 girls

(4) addition: 5 - 2 = 3; 6 + 3 = 9.
INTERVIEW: 0 students
PRETEST: 2 boys, 2 girls

(5) Miscellaneous errors: no evidence of strategy.
INTERVIEW: 0 students
PRETEST: 0 boys, 1 girl

(6) Some students did not have time to attempt a solution.
INTERVIEW: 1 boys, 2 girls
PRETEST: 0 students

12. PROBLEM: When Mike and Sally came home from school, they found a chocolate cake. Mike ate half of the cake, and Sally ate a third of the cake that Mike had left. How much of the whole cake was left after Sally ate her piece?

CORRECT STRATEGIES:

(1) The students who answered correctly drew their responses. All first divided a circle in half and then divided each half in thirds.
INTERVIEW: 2 boys, 0 girls
PRETEST: 1 boy, 3 girls

SEMI - CORRECT STRATEGIES

(2) Some students were correct in determining (by drawing) the fraction portion that both children ate, but they could not say what part of the total cake had been eaten. For example, 'of the half that is left, Sally ate 1/3 so 2/3 are left.'
INTERVIEW: 3 boys, 1 girl
PRETEST: 1 boy, 2 girls

(3) One student worked the problem correctly but got confused about labelling the parts of the cake. She multipled 1/2 x 1/3 but referred to this amount as that which was left (rather than the amount eaten by Sally).
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 students

INCORRECT STRATEGIES:

(1) Subtraction: many children subtracted because the problem said 'how much is left'. They obtained: 1/2 - 1/3 = 1/6.
INTERVIEW: 3 boys, 5 girls
PRETEST: 5 boys, 5 girls

(2) Addition and subtraction: many students first added the amounts eaten (1/2 + 1/3 = 5/6) and subtracted 5/6 from 6/6. They also had final solutions of 1/6.
INTERVIEW: 3 boys, 5 girls
PRETEST: 0 boys, 1 girl

(3) Addition only. 1/2 + 1/3 = 5/6.
INTERVIEW: 1 boy, 1 girl
PRETEST: 3 boys, 1 girl

(4) Multiplication. 1/2 x 1/3.
INTERVIEW: 1 boy, 0 girls
PRETEST: 1 boy, 1 girl

(5) Some students were certain the answer was 1/4 and tried to find ways to show it. One imagined a square cake with corners, and he equated corners with quarters, and said Sally ate 1/4, Mike ate 1/2, so 1/4 was left. One subtracted 1/3 from 1/2 but was not satisfied with the result, and still said 1/4.
One student added 1/2 + 1/3 = 3/4 and 1/4 is left. One other also answered 1/4 - guessing at the size of the piece that must be left.
Several pretest responses were 1/4, (no work shown).
INTERVIEW: 3 boys, 1 girl
PRETEST: 3 boys, 2 girls

(6) Some students reached the correct number for their answer but did so by guessing (according to the students).
INTERVIEW: 0 boys, 1 girl
PRETEST: 1 boy, 0 girls

(7) Several students drew responses.
INTERVIEW: 2 boys, 7 girls
PRETEST: 4 boys, 2 girls

(8) Miscellaneous errors. Each solution is unique.
INTERVIEW: 1 boy, 1 girl
PRETEST: 0 boys, 3 girls

(9) Some students were unable to solve the problem.
INTERVIEW: 1 boy, 0 girl
PRETEST: 4 boys, 3 girls

(10) Some students did not have time to attempt the item.
INTERVIEW: 1 boy, 3 girls
PRETEST: 0 students

13. PROBLEM: A bus has seats for 48 passengers. If one out of every six seats is empty, how many passengers are on the bus?

CORRECT STRATEGIES:

(1) Most of the students divided 48 by 6 to obtain 8. The second step was subtraction of 8 from 48 for answer 40.
INTERVIEW: 7 boys, 2 girl
PRETEST: 2 boys, 4 girls

(2) Second strategy is to divide 48 by 6 and obtain 8. Then multiply 8 x (6 - 1) = 40.
INTERVIEW: 2 boys, 1 girl
PRETEST: 0 students

(3) Pretest: some students divided 48 by 6 but did not specify second step. Answer was 40.
INTERVIEW: 0 students
PRETEST: 1 boy, 2 girls

(4) One student was correct by suggesting a drawing and the way it should be interpreted (the work was not done, however).
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 students

INCORRECT STRATEGIES:

(1) Most of the incorrect responses were first step only: 48 / 6 = 8.
INTERVIEW: 7 boys, 8 girls
PRETEST: 14 boys, 9 girls

(2) Some students transformed the problem into multiplication: 6 x ? = 48.
INTERVIEW: 0 boys, 3 girls
PRETEST: 0 students

(3) Subtraction: some students subtracted 6 from 48: 42.
INTERVIEW: 3 boys, 2 girls
PRETEST: 1 boy, 2 girls

(4) Some students attempted to work with the numbers 5 (from 6 - 1) and 48. One multiplied and one divided.
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 boys, 1 girl

(5) Two students gave responses on pretest with no work.
INTERVIEW: 0 students
PRETEST: 1 boy, 1 girl

(6) One student was incorrect with drawing.
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 students

(7) One student tried to divide by 3 because "3 people can sit on 1 bus seat".
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 students

(8) Several students could not solve the problem.
INTERVIEW: 1 boy, 1 girl
PRETEST: 2 boys, 2 girls

(9) Several students did not have time to attempt the item.
INTERVIEW: 1 boys, 5 girls
PRETEST: 2 boys, 2 girls

14. PROBLEM: There are 10 cookies in half a box. How many cookies are in the box when it is full?

CORRECT STRATEGIES:

(1) Solution by multiplication. 2 x 10 = 20.
INTERVIEW: 7 boys, 6 girls
PRETEST: 4 boys, 7 girls

(2) Solution by addition. 10 + 10 = 20.
INTERVIEW: 7 boys, 8 girls
PRETEST: 7 boys, 8 girls

(3) Solution by doubling the # of cookies in half a box. This does not appear to be quite the same as multiplying by 2 in the students' descriptions of problem solving.
INTERVIEW: 1 boy, 2 girls
PRETEST: 0 students

(4) Solution by drawing. Eight drawings were given in interviews. One drew a circle and divided it into 2 parts, making 10 small circles in one of the parts. Then the student drew 10 other small circles in the other part. A second student drew lines representing cookies in a box—first 10 lines, and then 10 more. A third was more abstract. A box was divided into 2 parts, and the words 10 cookies were entered in the lower part.
INTERVIEW: 2 boys, 6 girls
PRETEST: 0 students

(5) Two students in the interview began by multiplying 10 x 1/2 = 5 but both realized they were incorrect and changed to multiplication of 10 x 2. Both had the concept of 'full box of cookies' in mind and realized the answer '5' was unreasonable.
INTERVIEW: 1 boy, 1 girl
PRETEST: 0 students

(6) Pretest only: many students responded with the number 20 and showed no work.
INTERVIEW: 0 students
PRETEST: 12 boys, 7 girls

(7) Some students ran out of time before attempting this problem.
INTERVIEW: 3 boys, 3 girls
PRETEST: 0 boys, 1 girl

INCORRECT STRATEGIES:

(1) Only one student did not answer the item correctly. He multiplied 10 x 5 = 50 on the pretest. This was questioned during the strategy session and he realized his error and solved the problem correctly at that time.
INTERVIEW: 0 students
PRETEST: 1 boy, 0 girls

15. PROBLEM: You have been chosen manager of your school band. There are eight members of the band, and they will have 11 practice sessions this year. If you have to put away an instrument, a music folder, and a music stand for each band member after every practice, how many objects will you put away this year?

CORRECT STRATEGIES:

Correct strategies differed by the order of multiplication.

(1) 8 x 3 = 24 (# objects to be put away each session); 24 x 11 = 264.
This is the most logical arrangement of the numbers because the first result is understandable.
INTERVIEW: 9 boys, 10 girls
PRETEST: 1 boys, 2 girl

(2) 8 x 11 = 88 (# of persons practicing all year); 88 x 3 = 264
This corresponds to direct translation of the item according to appearance of the numbers.
INTERVIEW: 4 boys, 3 girls
PRETEST: 3 boys, 1 girl

(3) 11 x 3 = 33 (# objects to be put away for 1 person for all sessions)
33 x 8 = 264.
INTERVIEW: 0 students
PRETEST: 2 boys, 0 girls

INCORRECT STRATEGIES:

(1) Most incorrect responses are the result of multiplying only two of the three numbers.

(a) 11 x 3 = 33 (# objects to be put away for 1 person for year)
INTERVIEW: 2 boys, 2 girls
PRETEST: 4 boys, 3 girls

(b) 11 x 8 = 88 (#person sessions per year)
INTERVIEW: 1 boy, 0 girls
PRETEST: 0 boys, 4 girls

(c) 3 x 8 = 24 (#objects to be put away after one session)
INTERVIEW: 0 boys, 4 girls
PRETEST: 1 boy, 0 girls

(2) Multiplication and addition: (8 x 3) + 11 = 35 .
INTERVIEW: 0 students
PRETEST: 0 boys, 2 girls

(3) subtraction: 11 - 3 = 8.
INTERVIEW: 0 boys, 1 girl
PRETEST: 2 boys, 1 girl

(4) Addition: 11 + 11 or 11 + 11 + 11
or 11 + 8
INTERVIEW: 1 boy, 1 girl
PRETEST: 2 boys, 0 girls

(5) No work shown. Miscellaneous errors.
INTEREVIEW: 0 students
PRETEST: 0 boys, 2 girls

(6) Miscellaneous errors: no identifiable strategy.
INTERVIEW: 0 students
PRETEST: 2 boys, 3 girls

(7) Some students were unable to solve the item. Corresponding pretest response was blank.
INTERVIEW: 1 boy, 1 girl
PRETEST: 7 boys, 6 girls

(8) Some students did not have time to attempt solu ion.
INTERVIEW: 1 boy, 2 girls
PRETEST: 1 boy, 1 girl

16. PROBLEM: When it arrived at the school, the bookmobile had 85 books. One student returned 8 books and checked out 11 others. Two more students brought back 4 books each, and another student checked out 6 books. How many books are now in the bookmobile?

CORRECT STRATEGIES:

(1) Most students solved the item phrase by phrase, emphasizing that one solved by looking at each return and check out.
INTERVIEW: 13 boys, 9 girls
PRETEST: 5 boys, 7 girls

(2) Alternative strategy: add # books returned. Get separate total for # books checked out. Add total returned to original total; subtract total checked out from result.
INTERVIEW: 2 boys, 1 girls
PRETEST: 1 boy, 0 girls

(3) Correct but no work shown.
INTERVIEW: 0 students
PRETEST: 1 boy, 0 girls

INCORRECT STRATEGIES:

(1) Primary error: add 4 rather than 8 (from 2 students brought back 4 books each).
INTERVIEW: 1 boy, 4 girls
PRETEST: 7 boys, 4 girls

(2) Complete reversal of addition and subtraction.
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 students

(3) Inconsistency: some students began by adding books returned and subtracting books checked out, but reversed operations later in the solutions. (The various answers are different.)
INTERVIEW: 3 boys, 9 girls
PRETEST: 3 boys, 2 girls

(4) Used correct strategy #2 but reversed operations.
INTERVIEW: 0 students
PRETEST: 1 boy, 0 girls

(5) Ignore 1 part of problem (some # of books returned or checked out).
INTERVIEW: 1 boy, 0 girls
PRETEST: 2 boys, 0 girls

(6) Miscellaneous addition errors: adding all or part of the numbers in the problem. Each student has unique solution.
INTERVIEW: 0 students
PRETEST: 3 boys, 7 girls

(7) Numerical answer reached but no work given.
INTERVIEW: 0 students
PRETEST: 1 boy, 2 girls

(8) Some students did not have time to attempt this item.
INTERVIEW: 1 boy, 3 girls
PRETEST: 0 students

17. PROBLEM: John has 12 baseball cards. He gives 1/3 of them to Jim. How many does John have left?

CORRECT STRATEGIES:

(1) The most common interview solution involves finding "1/3 of 12". For this strategy, the multiplication sign was not used as the student wrote the response. Rather, the students wrote 1/3 OF 12. The second step was 12 - 4 = 8.
INTERVIEW: 3 boys, 1 girls
PRETEST: 0 students

(2) An alternate strategy is to specify the multiplication sign: 1/3 x 12 = 4; 12 - 4 = 8.
INTERVIEW: 0 boys, 1 girl
PRETEST: 1 boy, 0 girls

(3) Another strategy is the same as the first except that students divide 12 by 3 to obtain 4.
INTERVIEW: 5 boys, 1 girl
PRETEST: 1 boy, 1 girl

(4) A third strategy involves recognizing that 4 is 1/3 of 12 and one seeks 2/3 of 12 which is 2 x 4 = 8.
INTERVIEW: 2 boys, 2 girls
PRETEST: 0 boys, 1 girl

(5) On the pretest, three students responded with the correct answer but showed no work. Two additional students were unclear in their responses. One reached the first step solution of 1/4 rather than 4. The other divided 12 by 1/3 and got 4. Both reached the correct solution of 8.
INTERVIEW: 0 students
PRETEST: 3 boys, 2 girls

(6) Drawing: several students either drew circle or made marks for the cards. Their solutions were correct.
INTERVIEW: 0 boys, 2 girls
PRETEST: 0 boys, 1 girl

INCORRECT STRATEGIES:

(1) The first strategy is 1/3 x 12 = 4.
INTERVIEW: 3 boys, 2 girls
PRETEST: 3 boys, 1 girl

(2) The second strategy divides 12 by 3 (no fraction computations).
INTERVIEW: 0 boys, 2 girls
PRETEST: 2 boys, 1 girl

(3) Many students found 1/3 of 12 without indicating any operation. Their responses had the form "12 1/3 = 4". No multiplication symbol is given. One explicitly said he was not multiplying.
INTERVIEW: 1 boy, 1 girl
PRETEST: 3 boys, 3 girls

(4) Some students reached the response of 4 without knowing how they found it. Two boys said "take 1/3 away from 12 and he would have 4 left." A similar response appears to be given on a pretest with the formulation of 12 - 1/3 = 4.
INTERVIEW: 2 boys, 0 girls
PRETEST: 1 boy, 0 girls

(5) A common error was to subtract 1/3 from 12. One interview response was 11 2/3, one was 35/3 (both girls), one took 12/1 - 1/3 = 11/2 (also girl). One (girl) answered 11 1/8. Two other interviews said that one is taking something away "subtract it somehow" but neither knew how to do subtraction. Four students on the pretest gave the response of 11 2/3. One more wrote 12 - 1/3 = 12 1/3.
INTERVIEW: 2 boys, 7 girls
PRETEST: 3 boys, 6 girls

(6) Two students knew that 1/3 represented some number of cards, but they did not know how to find the number. Both thought it would be subtracted from 12.
INTERVIEW: 0 boys, 2 girls
PRETEST: 0 students

(7) One error was the response of "1/4".
INTERVIEW: 1 boy, 0 girls
PRETEST: 1 boy, 0 girls

(8) Drawing : one drew circle and divided into 12ths, marked out 3 sections and answered 9 (f36). Two drew 4 rows of 3; responded 9.
INTERVIEW: 0 boys, 1 girl
PRETEST: 1 boy, 1 girl

(9) Response of 9: 12 - 3 = 9 or no work.
INTERVIEW: 1 boy, 0 girls
PRETEST: 2 boys, 3 girls

(10) Several students left the item blank on the pretest. One student could not solve it in the interview.
INTERVIEW: 0 boys, 1 girl
PRETEST: 3 boys, 3 girls

(11) Several students did not have time to attempt the problem.
INTERVIEW: 1 boy, 3 girls
PRETEST: 0 students

18. PROBLEM: In a grocery display, there are a dozen watermelons. Each melon takes up the same amount of space as 8 oranges. If the grocer decided to replace half the watermelons with oranges, how many oranges would he need?

CORRECT STRATEGIES:

(1) Most children first took half of 12 and then multiplied 6 x 8 = 48. Problem representation was very good--have 12 melons and then take half of them away so that's 6.
INTERVIEW: 12 boys, 13 girls
PRETEST: 11 boys, 8 girls

(2) Alternate strategy is to find the number of oranges that would be used to replace the entire display and then take half the total. 12 x 8 = 96; 96 / 2 = 48.
INTERVIEW: 2 boys, 2 girls
PRETEST: 0 students

(3) One student elected to draw the display. He drew 12 circles (for melons) then boxed off 6 of them ("that's half"). He then wrote the number 8 inside each of the circles in the box. He counted the number of 8's, and multiplied by the total. 6 x 8 = 48.
INTERVIEW: 1 boy, 0 girls
PRETEST: 0 students

(4) One student added 6 eight times (finding 1/2 of 12 first).
INTERVIEW: 1 boy, 0 girls
PRETEST: 0 students

(5) Some students were correct on pretest but gave no work.
INTERVIEW: 0 students
PRETEST: 4 boys, 1 girl

(6) One student misunderstood the problem as having 8 oranges and 8 melons. He took half the melons (4) times the number of oranges (8) for a total of 32. The strategy is correct.
INTERVIEW: 1 boy, 0 girls
PRETEST: 0 students

INCORRECT STRATEGIES:

(1) Predominant pretest error: 12 x 8 = 96. Ignore part of text specifying "one half".
INTERVIEW: 1 boy, 1 girl
PRETEST: 3 boys, 7 girls

(2) Errors with "one half": take 1/2 of 12 and 1/2 of 8 and multiply the resulting 4 x 6. Second approach is 1/2 of 12 = 6.
INTERVIEW: 0 students
PRETEST: 1 boy, 2 girls

(3) Errors relating 1/2 to both oranges and melons: since taking one half of 12 (which is 6 and students usually specify this), also take 1/2 of the oranges. 1/2 of 8 = 4. Therefore, answer is 8 - 4 = 4.
INTERVIEW: 1 boy, 1 girl
PRETEST: 3 boys, 3 girls

(4) Division: several students attempted to divide either 12 by 4 or 12 by 8. None could specify why.
INTERVIEW: 1 boy, 1 girl
PRETEST: 1 boy, 0 girls

(5) Multiplication: 8 x 8. No reason specified.
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 students

(6) Some students did not know how to solve the problem
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 boys, 2 girls

(7) Several students had no time to solve the problem.
INTERVIEW: 1 boys, 6 girls
PRETEST: 0 students

19. PROBLEM: A group of bicycle riders took a three day trip. On the first day, they traveled 14 miles further than they had originally planned. On the second day, they went 8 miles less than they had planned, and on the third day, they traveled 16 miles more than they had planned. If the bike riders actually traveled a total of 180 miles in 3 days, how many miles had they originally planned to bike?

CORRECT STRATEGIES:

(1) One strategy requires finding how many miles more than planned the group traveled and subtracting that amount from 180.
14 - 8 = 6; 6 + 16 = 22; 180 - 22 = 158
INTERVIEW: 6 boys, 4 girls
PRETEST: 2 boys, 1 girls

(2) One strategy first adds the amounts representing FURTHER: 14 + 16 = 30; then the amount traveled less than planned is subtracted: 30 - 8 = 22; finally, 180 - 22 = 158.
INTERVIEW: 1 boy, 2 girls
PRETEST: 0 students

(3) The amounts for each day can be subtracted from the total one at a time: 180 - 14 = 166; 166 + 8 = 174; 174 - 16 = 158.
INTERVIEW: 3 boys, 1 girl
PRETEST: 0 students

(4) First the daily average can be found (180/3 = 60). Then the difference per day is 60 - 14 = 46; 60 + 8 = 68; 60 - 16 = 44; 46 + 68 + 44 = 158.
INTERVIEW: 0 students
PRETEST: 1 boy, 0 girls

INCORRECT STRATEGIES:

(1) Two students used the same strategy described in CORRECT (4) but reversed the operations: 60 + 14, 60 - 8; 60 + 16.
INTERVIEW: 2 boys, 0 girls
PRETEST: 0 students

(2) Two students used the same strategy described in CORRECT (1) but reversed the operations: 180 + (14 - 8 + 16) = 202.
INTERVIEW: 0 boys, 1 girl
PRETEST: 0 boy, 1 girl

(3) addition: add the smaller numbers in the problem: 14 + 8 + 16 = 38. Now subtract from 180.
INTERVIEW: 0 boys, 1 girl
PRETEST: 1 boy, 0 girls

(4) addition: add the smaller numbers in the problem. 14 + 8 + 16 = 38.
INTERVIEW: 0 students
PRETEST: 1 boy, 4 girls

(5) first step correct: 14 - 8 + 16 = 22.
INTERVIEW: 2 boys, 1 girl
PRETEST: 0 boys, 3 girls

(6) First step correct; multiplied 22 x 3 or added 22 three times.
INTERVIEW: 0 students
PRETEST: 1 boy, 1 girl

(7) Find daily average if go 180 miles: 180 / 3 = 60. Ignore all the other information in the problem.
INTERVIEW: 2 boy, 1 girls
PRETEST: 6 boys, 5 girls

(8) Addition: add all numbers in the problem. 180 + 14 + 8 +16.
INTERVIEW: 0 students
PRETEST: 0 boys, 1 girl

(9) Subtraction: 14 - 8 = 6; one student then elected to multiply 6 x 16 and one added 6 to 180. A third subtracted 14 - 8 = 6; 16 - 6 = 10; 180 - 10 = 170.
INTERVIEW: 0 students
PRETEST: 2 boys, 1 girl

(10) Addition: 16 + 14 = 30; 180 - 30 = 150. Ignore amount traveled less than planned.
INTERVIEW: 0 students
PRETEST: 1 boy, 1 girl

(11) Miscellaneous errors or guesses with no work.
INTERVIEW: 1 boy, 0 girls
PRETEST: 1 boy, 1 girl

(12) Several students could not solve the problem.
INTERVIEW: 3 boys, 11 girls
PRETEST: 6 boys, 5 girls

(13) Some students did not have time to attempt the problem.
INTERVIEW: 1 boy, 4 girls
PRETEST: 0 students

Sex Differences in Children's Mathematics Achievement: Solving Computations and Story Problems[1]

Several recent studies compare males' and females' mathematics achievement (Benbow & Stanley, 1982, 1980; Swafford, 1980; Fennema & Sherman, 1978, 1977; Hilton & Berglund, 1974; Backman, 1972; Oleander & Ehmer, 1971). Results of these studies are not consistently in favor of one sex or the other. These mixed results suggest that we may be focusing on an inappropriate unit of measure. Some progress can be made in understanding this situation by comparing performance on well-defined types of items and determining where, if anywhere, the sexes differ.

Two item types that appear especially important in both elementary and secondary school are computations and story problems (Armstrong, Note 1; Corbitt, 1981). Computations are usually presented in equation form with the operation to be performed indicated by the appropriate arithmetic symbol. For story problems, the operations are not explicitly stated, and the student must determine which operation to perform as well as which pieces of information in the problem to use.

This paper addresses the issue of sex differences in responses of sixth-grade students to computations and story problems. For comparison, arithmetic content of both types of items is restricted to fraction, whole number, and decimal arithmetic. There are two major issues: (1) are there differences in boys' and girls' success in solving such items and (2) are there additional factors (e.g., reading achievement or socioeconomic status) that interact with gender to influence mathematics performance.

Sixth-grade students' responses are a valuable source of data. Several external factors previously found to be related to sex differences in mathematics can be controlled or eliminated from consideration. First, children have had similar instructional experiences. All students study mathematics in each year of elementary school. This fact eliminates one important source of sex differences in studies of high school students, for in these studies sex-related differences favoring boys appear to be artifacts of differential course-taking by boys and girls in high school (Fennema & Sherman, 1977). Second, children of this age have not yet reached puberty. Arguments of differences due to hormonal variation are not applicable (Petersen, 1979). Finally, perception of mathematics as a masculine domain is arguably weaker at the sixth grade than in high school (Fox, Tobin, & Brody, 1979). Girls are as likely as boys to rate mathematics as a favorite subject in sixth grade (California Assessment Program, 1980).

Patterns of Responses

The method of study employed here is detailed examination of student responses to standardized test items. Attention is given to a student's pattern of response to these items rather than to the number of correct responses. The objective is to define all possible patterns of response, to classify all boys and girls according to these patterns, and to compare the resulting distributions.

Assume two groups of individuals respond to a two-item test. Possible scores are 0, 1, and 2 (indicating number of correct responses). Scores of 0 and 2 are unambiguous; either both items were missed or both were solved correctly. A score of 1 cannot be as easily interpreted. One item was answered correctly, but it is impossible to determine which one. It is possible that individuals from one group obtain a score of 1 by missing the first item and that individuals from the second group obtain a score of 1 by erring on the other item. In such a situation, group differences would not be apparent from observation only of total correct responses. It is necessary to preserve and analyze item-by-item data.

Rather than using scores of 0, 1, and 2, one can categorize the responses as ordered pairs of 00, 01, 10, and 11, where 0 indicates an incorrect response and 1 is correct. The first element of the pair denotes response to the first item, and the second element denotes response to the second item. The pair 00 reflects errors on both items and 11 indicates both correct. The pair 01 is distinguished from 10; for the former an error occurred on the first item, and for the latter the error occurred on the second item. Each individual responding to the two items can be classified in one and only one of the four categories of responses.

Under such a classification, one can still examine total test performance. The ordered pairs can easily be extended to n-tuples of 0's and 1's, according to the number of test items to be studied. These n-tuples will be called response patterns.

Since one objective of the present study is to compare performance on computations and story problems, it is natural to construct the above classification so that there is one n-tuple or pattern for computations and a second n-tuple or pattern for story problems. Each individual will have one and only one response pattern for each item type. Each individual also will be classified according to gender. The result is a set of cross-classified categorical data over three factors:

response patterns for computations, response patterns for story problems, and sex.

If sex differences exist in the responses, they will be evident in the unequal proportions of boys and girls responding with any particular pattern. These differences can be evaluated statistically.

Other Factors Influencing Mathematics Performance

Reading. It has been suggested that a strong relationship exists between reading ability and successful solving of story problems. Aiken (1971) and Muscio (1962) report correlations between math and reading tests ranging from .549 to .816. The implication in these studies is that successful problem solvers are probably good readers and that poor readers are probably poor problem solvers.

Studies of sex differences on reading tests reveal that girls generally score higher than boys (Maccoby & Jacklin, 1974; Backman, 1972; Aiken, 1971; Muscio, 1962). One is tempted to conclude that girls should be superior to boys in solving story problems as well, under the assumption that successful solutions depend upon successful reading of the problems.

Several questions of interest can be addressed by looking at reading achievement and performance on mathematics items. One can ask whether the probability of success in solving story problems or computations varies with reading score. One expects the probabilities for both sexes to rise with reading score. It has not been known whether these probabilities rise at similar rates for boys and girls. The present study indicates that they do not.

Other Student Characteristics. Analyses of other student characteristics may contribute to our understanding of sex differences in mathematics achievement. Three additional variables are examined: socioeconomic status, primary language spoken at home, and chronological age.

Previous studies have demonstrated sizable test score differences among various socioeconomic classes (e.g., California Assessment Program, 1980). If there are differences in the achievement of girls and boys on computations and story problems, it is feasible to ask whether there is an interaction between SES and sex in arithmetic performance and whether there are identifiable groups of children of either sex who are more likely to solve problems successfully.

A second variable of interest is the primary language of the student. Classification by language provides an opportunity of examining results from three different cultural groups: Spanish, English, and Oriental.

Finally, one can question whether boys and girls of the same age are performing equally. Identified sex differences may be the result of comparing children of differing maturity.

To investigate the issues raised here, analyses are carried out on a large set of data. The first analysis determines whether there are differences in boys' and girls' responses to the two item types. The second analysis probes the relationship of additional student characteristics with mathematics performance to determine whether any subgroup of students more likely to success or fail can be identified.

Methods

Subjects

Test data from all sixth grade children in California enrolled in public schools in 1979 were analyzed. A total of 286,767 students responded to the _Survey of Basic Skills, Grade 6_, in May 1979: 144,462 boys and 142,305 girls.

Instrument

The _Survey of Basic Skills, Grade 6_, is a standardized test developed by the California Assessment Program to estimate the average achievement of children at school, school district, and state levels.[2] The test was constructed on a matrix sampling basis. Under this sampling procedure, each child answers only a small sample of all test items. There are 16 forms of the _Survey_, each containing unique sets of items. Each test form has 30 items: 10 problems of mathematics, 8 items each of reading and written expression, and 4 items of spelling. A child answers only a single form of the test. Thus, each child responds only to 10 of the 160 mathematics items.

The tests were administered in random order so that no classroom was tested with a single set of items. Approximately 18,000 (9,000 boys and 9,000 girls) answered each test form.

Of the 160 mathematics questions, 41 are problems of computation with whole numbers, fractions, or decimals, and 27 are story problems of the same arithmetic areas.

Either 4 or 6 of the 10 items on each test form are problems from these three arithmetic areas. (The remaining items are problems of geometry, measurement, number concepts, and probability. They are neither computations nor story problems, and they will not be discussed here.) discussed here.)

In addition to the children's responses to the 30 items of the Survey, the California Assessment Program gathers information about each child's socioeconomic status, language fluency, primary language, age, and sex. Estimates of socioeconomic status of students are made by teachers at the time of testing. Classification is made on the basis of occupational level of the principal breadwinner in the student's family. The four possible categories of classification are professional, semi-professional, skilled worker, and semi-skilled worker.

Two language classifications are also made by teachers. First, the teachers indicate whether students are fluent, limited, or restricted in the use of the English language. Teachers specify the primary language (if not English) of each student. For the analyses here, only students fluent in English were included.

Several language choices were available to the teachers. Some of these have very few occurrences and are not used in the analyses. The three groups used in the comparisons are English, Spanish, and Oriental. Members of the first group speak English only, of the second speak Spanish in addition to English, and of the third group speak Chinese or Japanese in addition to English.

Each student's age at the time of testing was recorded. The ages for analyses here are defined in three-month increments. The tests were administered in the spring; the ages, therefore, are those of children completing the sixth grade of elementary school.

Results

Log-Linear Analysis

Fifteen of the sixteen forms of the Survey were analyzed to determine the underlying model that best fits the data of computation and story problem responses. One form contains no story problems and was excluded. Thirteen forms contain four relevant items, and two forms contain six items each. The responses to each form were cross-classified according to pattern of response to computations, pattern of response to story problems, and sex.

The data from a randomly selected test form were compared with the expected frequencies that occur under all possible hierarchical log-linear models (i.e., all combinations of main, second, and higher-order effects).[3] Main effects are sex, computation response pattern, and story problem response pattern. Likelihood ratio χ^2 values were obtained from the comparisons.[4] All but a single model failed to predict the data; these models were rejected with $p < .001$. For the remaining model, a χ^2 of 9.06 was computed, with 9 degrees of freedom. The model that fits contains all main effects and all two-factor interactions. The three factor interaction is not a significant component of the model.

To test the hypothesis that this general model also predicts adequately the data from other test forms containing different sets of items, log-linear analyses were carried out on the remaining fourteen test forms. For all but a single test form, the model of main effects and all two-factor interactions was the only satisfactory model. In each case, all other models were rejected with p .001. Probability levels for the models of best fit ranged from .074 to .846, with a mean value of .332. Given the large number of individuals responding to each test form (roughly 8,000 of each sex), one can be fairly confident that the low χ^2 values are an indication of good fit and not the result of small sample size.

Inspection of the full table of cross-classified data shows the direction of the interactions present in the model of best fit. For all test forms, successful solving of computations is positively associated with successful solving of story problems. Interaction of sex with patterns of solving computations results because girls are more successful than boys. Similarly, interaction of sex and patterns of solving story problems occurs because boys are more successful than girls. Girls and boys differ in their ability to solve these two types of items.

These analyses were replicated on similar data from 1977-78 and 1976-77. For each set of test scores, the identified model of fit holds for fourteen of the fifteen test forms.

<u>Additional Factors of Reading, SES, Language, and Age</u>

To analyze the influence of additional factors on computation and story problem responses, attention was restricted to successful solvers of each type of item. Successful computation solving is defined by all computations correct. Similarly, successful story problem solving is defined by all story problems correct. For consistency, only those test forms having two items each of computation and story problem were considered. Eight forms meet this requirement, yielding 32 items.

Reading. Table 2 gives the percentages of boys and girls scoring zero to eight on the eight-item reading test of the Survey. Boys are more likely than girls to score 0, 1, 2, 3, 4, and 5. Girls are more likely than boys to score 6, 7, and 8. Mean scores are similar: 5.23 and 5.38 for boys and girls, respectively.

Figure 1 shows the probabilities of answering correctly both computations and/or both story problems at every obtainable reading score. Data for boys and girls is plotted separately. The upper two lines of the figure represent success on computations. Success on story problems is given by the lower two lines. Without exception, sixth-grade children at every reading score are more successful in solving computations than in solving story problems.

Excluding scores of 0 and 1, boys consistently have higher probabilities of success on story problems than do girls, and girls consistently have higher probabilities of success on computations than do boys. These differences are not large, but they are evident at every reading score.

The question of interest is whether the discrepancy between probability of computation success and probability of story problem success remains constant or varies at each reading score. Discrepancy is computed by subtracting the probability of success on story problems from the probability of success on computations. A related question is whether there is more variation in this discrepancy for one sex than the other. This issue translates to a simple question of regression: Do the slopes of the lines of best fit predicting discrepancy from reading score for boys and for girls differ. Figure 2 gives the regression lines and also the discrepancies for both sexes. The discrepancies range from .092 to .160 for boys and from .068 to .233 for girls. The slopes of the two lines differ significantly, $t\ (14) = 3.67$, $p < .01$. For both sexes, the discrepancy between computation ability and story problem ability grows larger as reading score increases. However, the discrepancy for girls increases at a significantly greater rate than that for boys.

Socioeconomic Status. Table 3a provides percentages of boys and girls classified by teachers in each of the four SES groups. Similar numbers of boys and girls are in each category.

As in the previous analysis with reading achievement, success on computations is again defined by correct responses on both computations. Success on story problems is similarly defined. Figure 3 illustrates probabilities of girls' and boys' success on the two types of items for each SES category.

As expected, socioeconomic level is a major factor in successful problem solving, and the average probabilities for all children ascend as SES classification increases. It is clear from this figure that no single SES group of either boys or girls has a particular advantage or disadvantage. A comparison of the discrepancies in probability of success at computation and probability of success at story problems indicates that the slopes of regression lines for discrepancies of boys and girls are almost identical. Neither slope differs significantly from zero.

Again, the discrepancy for girls is higher than that for boys. The discrepancy values are approximately .14 for boys and .19 for girls. A test of differences yields $\underline{t}$ (3) = 4.193, $\underline{p} < .05$.[5]

Primary Language. Table 3b provides percentages of students classified by primary language spoken at home. Not all students responding to the Survey are tabled; only the three groups described previously are represented.

Figure 4 compares the probabilities of success for boys and girls on the two types of items. All groups have higher probability of success on computations than on story problems. Again, the pattern of girls' relative excellence in computations and boys' relative excellence in story problems emerges.

The discrepancies between probabilities of success on computations and story problems varies widely. For boys, the size of the mean discrepancy of Spanish and English students is approximately .135. For Oriental boys, it jumps to .183. Among the girls, Spanish-speaking children have the lowest value of .177 while the other two groups are approximately .200.

Age. Table 3c provides the percentages of each sex in eleven age categories. The average age for sixth-grade boys is 12 years, 2 months (146 months) and that for girls is 12 years, 1 month (145 months).

Figure 5 illustrates the probabilites of success of boys and girls for the eleven age categories. The general trend noted in the three previous analyses is again evident. However, the computation superiority of girls is not consistent over every age. For both boys and girls, performance declines with increasing age on all items.

There are differences in the discrepancies of boys and girls. The range of discrepancies for boys is .110 to .145. The range of values for girls is .109 to .205. As with reading scores, there is more fluctuation in the discrepancy values for girls than for boys. The slope of predicted discrepancy values from age does not differ significantly from zero for boys. However, the slope of predicted values for girls is negative (t (9) = -2.327, $p < .05$).

Discussion

There are differences in sixth grade boys' and girls' responses to computations and story problems. Boys tend to perform better than girls on story problems and girls tend to perform better than boys on computations. Absolute values of the differences are small, but the direction of difference remains constant. Adjustment of probabilities of success according to various student characteristics does not alter the direction of difference.

Similar differences have been found at the high school level (Fennema, 1977; Armstrong, Note 1). Two large assessments, the National Assessment of Educational Progress and the California Assessment Program, also note the distinction (Fennema & Carpenter, 1981; California Assessment Program, 1979). It appears that these differences do not arise at the high school level but exist much earlier, at least as early as sixth grade.

Log-Linear Analyses

Sex is a significant factor in the statistical model that describes the students' patterns of responses to all items. Log-linear analyses of fifteen test forms indicates that the three factors of sex, response pattern to computations, and response pattern to story problems are necessary components of a model that predicts performance. Moreover, there is significant interaction of sex with successful solving of both computations and story problems.

A possible criticism of the present study is that the number of relevant items answered by any individual child is small (i.e., four or six items). However, over 285,000 students responded to fifteen unique sets of items, providing about 18,000 observations for each of 68 distinct items. Even with the limited number of items per individual, clear differences in boys' and girls' patterns of responses were found. Expanding the number of items answered by each individual would no doubt yield additional detail about these patterns.

Additional Factors

The comparisons of boys' and girls' probabilities of correct responses conditional upon reading achievement, SES, primary language, and age are remarkably consistent. For any level of each of these classifications, girls generally have higher probabilities than boys of answering both computations correctly and boys generally have higher probabilities than girls of answering the story problems correctly.

Reading. The reading results are particularly surprising. On the CAP Survey, girls have higher measured reading achievement than boys. High achieving girls in reading are nonetheless relatively weaker in solving story problems than other girls or than boys. This is evident in the increasing size of the difference between probability correct for computations and probability correct for story problems--the discrepancy values. These discrepancies are highest for girls with superior reading achievement. As reading score increases, the discrepancy values increase as well.

One interprets these results cautiously, of course, because the measure of reading achievement is an eight-item test. The items were developed in accordance with the curriculum of the California public school systems and thus are expected to have content validity. On each of the fifteen forms of the Survey there are eight items, for a total of 120 items evaluating reading achievement. Since the same results emerge on each of the fifteen test forms, we may have some confidence that the trend does indeed exist.

It is more difficult to explain why it occurs. The present study cannot provide a definitive explanation, but it does suggest some ideas. Successful solving of computations implies successful use of arithmetic laws and rules for combining numbers. It also implies fairly automatic use of these rules. Girls appear to have mastered the rules to a higher degree than boys have and thus are more rewarded for rule-governed behavior. Attempts to solve simple word problems analogously by a set of rules are occasionally fruitful, but as problem complexity increases, the rule-governed approach becomes less unsuccessful. (Support for this argument can be found in Marshall, Note 2.)

If boys are less rule-governed (and more inaccurate) than girls on computations, they may be less rule-governed and more flexible in approaching story problems, and they may consequently be more successful in solving the latter. Further study is necessary to determine whether this is the case.

Language, SES, and Age. Two of the factors that might affect the observed sex differences in performance are socioeconomic status and primary language. Certain groups of boys or girls may have particular advantages or disadvantages. The present results suggest this is not the case with SES or language. While there are large differences between the levels of SES and language, there are no similarly large differences between the sexes at any level of these variables. Thus, these factors do not appear to explain sex differences in mathematics performance.

The age results are somewhat misleading. According to the age requirements in California public schools, the average sixth grade child should be between 138 and 150 months at the time the Survey is administered.5 Thus, the individuals in the first two categories of Table 3c and Figure 5 have either skipped a grade or have had other extenuating circumstances. Similarly, the individuals in the last five categories are not in the expected age range. One suspects that the majority of these children have repeated one or more grades. The abrupt drop in all lines of Figure 5 at the category of 150-152 months may indicate the presence of a second and different population of students, those who have not progressed at the normal rate through elementary school. Since data on grade repetition was not gathered at the time of testing, one can only speculate about this distinction.

If the first two and last five categories are omitted from Figure 5, the remaining categories represent the probabilities for the expected average ages of sixth grade students. All four graphs in this range are relative flat, suggesting that age is not a contributing factor to sex differences in mathematics performance for individuals progressing normally through elementary school.

The analyses presented here may help interpret results of studies in which total test scores of boys are compared with those of girls. Such studies have mixed results. Boys excel in some cases (Benbow & Stanley, 1982; Hilton & Berglund, 1974), girls excel in others (Olander & Ehmer, 1971; Wozencraft, 1963) , and some comparisons reveal no differences (Swafford, 1980; Fennema & Sherman, 1977). Comparisons of total test scores implicitly or explicitly assume that boys and girls have the same type of ability but in differing quantities. It is clear from the present results that ability type plays an important role in the success of a student on a particular test. If the test contains more computations than story problems, girls would be expected to surpass boys. If the opposite weighting of items occurs, boys should excel.

Support for the theory of differential ability over item type can be found in studies that have tests of different skills. Fennema and Sherman (1978, Table 1) report that sixth grade girls consistently score higher than boys on tests of computations and that boys are superior on tasks of problem-solving. Although the differences are not statistically significant, the trend is evident in each group of students tested. The Second National Assessment of Educational Progress and California Assessment Program report similar findings (without statistical analysis).

The present analyses support the theory of differential abilities over two specific item types. It may well be the case that differences exist in other item types as well. Spatial ability is a much-studied area, for example, and boys appear to do better than girls on many (but not all) types of spatial items (see Maccoby & Jacklin (1974) in general and Vandenberg & Kuse, 1979, or Petersen, 1979, in particular). Additional studies are necessary to identify and evaluate other item types and sex-differences that may exist for them.

Footnotes

1. This research was supported by the National Institute of Education under Grant No. NIE-G-80-0095. The findings and conclusions presented here do not necessarily reflect the views of the National Institute of Education.

2. The author wishes to thank Tej N. Pandey of the California Assessment Program for his assistance in conducting this study.

3. The reader is referred to standard texts such as Bishop, Fienberg, and Holland (1975) or Haberman (1978) for details of log-linear analysis.

4. As Bishop, Fienberg, and Holland (1975) point out, the likelihood ratio statistic is minimized by maximum likelihood estimates and can be partitioned conditionally and structurally (pp. 125-127).

5. Given the few degrees of freedom, the reader may wish to view this result cautiously.

6. To enter kindegarten, a child must have reached his/her fifth birthday by December 1.

Reference Notes

1. Armstrong, Jane M. Achievement and Participation of Women in Mathematics: An Overview. Report of a Two-Year Study Funded by the National Institute of Education, Report 10-MA-00. Denver, Colorado: Education Commission of the States, 1980.

2. Marshall, S. P. Sex differences in mathematics errors: An analysis of distractor choices, forthcoming in Journal for Research in Mathematics Education.

References

Aiken, L. R. Intellective variables and mathematics achievement: Directions for research. Journal of School Psychology, 1971, 9, 201-212.

Backman, M. E. Patterns of mental abilities: Ethnic, Socioeconomic, and sex differences. American Educational Research Journal, 1972, 9, 1-12.

Benbow, C. & Stanley, J. Consequences in high school and college of sex differences in mathematical reasoning ability: A longitudinal perspective. American Educational Research Journal, 1982, 19, 598-622.

Benbow, C. & Stanley, J. Sex differences in mathematical ability: Fact or artifact? Science, 1980, 210, 1262-1264.

Bishop, Y. M., Fienberg, S. E., & Holland, P. W. Discrete multivariate analysis: Theory and practice. Cambridge, Massachusetts: The MIT Press, 1975.

California Assessment Program. Student Achievement in California Schools, 1978-79 Annual Report. Sacramento, California: State Department of Education, 1979.

Fennema, E. L. Influences of selected cognitive, affective, and educational variables on sex-related differences in mathematics learning and studying. In Women and mathematics: Research perspectives for change (National Institute of Education Papers in Education and Work: No. 8). Washington, D.C.: U.S. Government Printing Office, 1977.

Fennema, E. Mathematics learning and the sexes: A review. Journal for Research in Mathematics Education, 1974, 5, 126-139.

Fennema, E. L. & Carpenter, T. P. Sex-related Differences in Mathematics. In Corbitt, M. K. (Ed.), Results from the Second Mathematics Assessment of the National Assessment Educational Progress. Reston, Virginia: The National Council of Teachers of Mathematics, 1981.

Fennema, E. & Sherman, J. Sex-related differences in mathematics achievement, spatial visualization and affective factors. American Educational Research Journal, 1977, 14, 51-71.

Fennema, E. & Sherman, J. Sex-related differences in mathematics achievement and related factors: A further study. Journal for Research in Mathematics Education, 1978, 9, 189-203.

Fox, L. H., Tobin, D., & Brody, L. Sex-role socialization and achievement in mathematics. In Wittig, M. A. & Petersen, A. C. (Eds.), Sex-related differences in cognitive functioning. New York: Academic Press, 1979.

Haberman, S. J. Analysis of qualitative data (volume 1). New York: Academic Press, 1978.

Hilton, T. L. & Berglund, G. W. Sex differences in mathematics achievement, a longitudinal study. Journal of Educational Research, 1974, 67, 231-237.

Maccoby, E., & Jacklin, C. Psychology of sex differences. Palo Alto, California: Stanford University Press, 1974.

Muscio, R. D. Factors related to quantitative understanding in the sixth grade. Arithmetic Teacher, 1962, 9, 258-262.

Olander, T., & Ehmer, C. L. What pupils know about vocabulary in mathematics. Elementary School Journal, 1971, 71, 361-367.

Petersen, A. C. Hormones and cognitive functioning in normal development. In Wittig, M. A. & Petersen, A. C. (Eds.), Sex-related differences in cognitive functioning. New York: Academic Press, 1979.

Swafford, Jane O. Sex differences in first year algebra. Journal for Research in Mathematics Education, 1980, 11, 335-345.

Vandenberg, S. G. & Kuse, A. R. Spatial ability: A critical review of the sex-linked major gene hypothesis. In Wittig, M. A. & Petersen, A. C. (Eds.), Sex-related differences in cognitive functioning. New York: Academic Press, 1979.

Wozencraft, M. Are boys better than girls in arithmetic? Arithmetic Teacher, 1963, 10, 486-490.

Table 1

Sample Items*

	Computations	Story Problems
Whole Numbers:	744 + 578	When it left the station a bus was carrying 40 passengers. At the first stop, 3 passengers got off and 7 passengers got on. How many passengers were then riding in the bus?
Fractions:	1/2 - 1/6 =	Mr. Witt worked 6 3/4 hours on Saturday building the table and finished it in 2 1/2 hours one evening. How many hours did it take to build the table?
Decimals	.5 x .5 =	Tom earned $4.20 in 2 days. The second day he earned $2.15. How much did he earn the first day?

*The items are reprinted here by permission of the California Assessment Program.

Table 2

Distribution of Boys and Girls According to Reading Score

Reading Score	Percent of Each Sex: Boys	Girls
0	1.25	1.05
1	3.71	3.07
2	7.08	5.99
3	9.91	9.46
4	12.52	12.10
5	15.18	14.88
6	17.40	18.13
7	18.74	19.58
8	14.22	15.75
Total	100.00	100.00

a Based upon a total of 71,202 boys.
b Based upon a total of 70,262 girls.

Table 3

Distribution of Boys and Girls According to Student Characteristics

A. Distribution by Socioeconomic Level

Occupation of Primary Breadwinner in Family	Percent of Each Sex: Boys	Girls
Unskilled Worker	19.41	13.59
Semi-Skilled Worker	41.47	[illegible].13
Semi-Professional	21.37	21.70
Professional	17.75	17.58
Total	100.00	100.00

Based upon a total of 56,606 boys and 55,692 girls.

B. Distribution by Primary Language Spoken at Home

Primary Language	Percent of Each Sex: Boys	Girls
Oriental	1.18	1.13
Spanish	12.84	13.47
English	85.98	85.40
Total	100.00	100.00

a Chinese or Japanese only

Based upon a total of 61,701 boys and 61,142 girls.

C. Distribution by Age:

Age in Months	Percent of Each Sex: Boys	Girls
Under 135	2.96	2.43
135-137	1.63	2.40
138-140	14.59	18.28
141-143	18.33	20.80
144-146	19.27	21.22
147-149	19.41	20.11
150-152	9.58	6.10
153-155	5.87	3.73
156-158	4.29	2.43
159-161	2.66	1.62
Over 161	1.41	0.88
Total	100.00	100.00

Based upon a total of 72,130 boys and 71,185 girls.

Figure 1

Average probabilities for boys and girls of answering all computations or all story problems correctly given reading score.

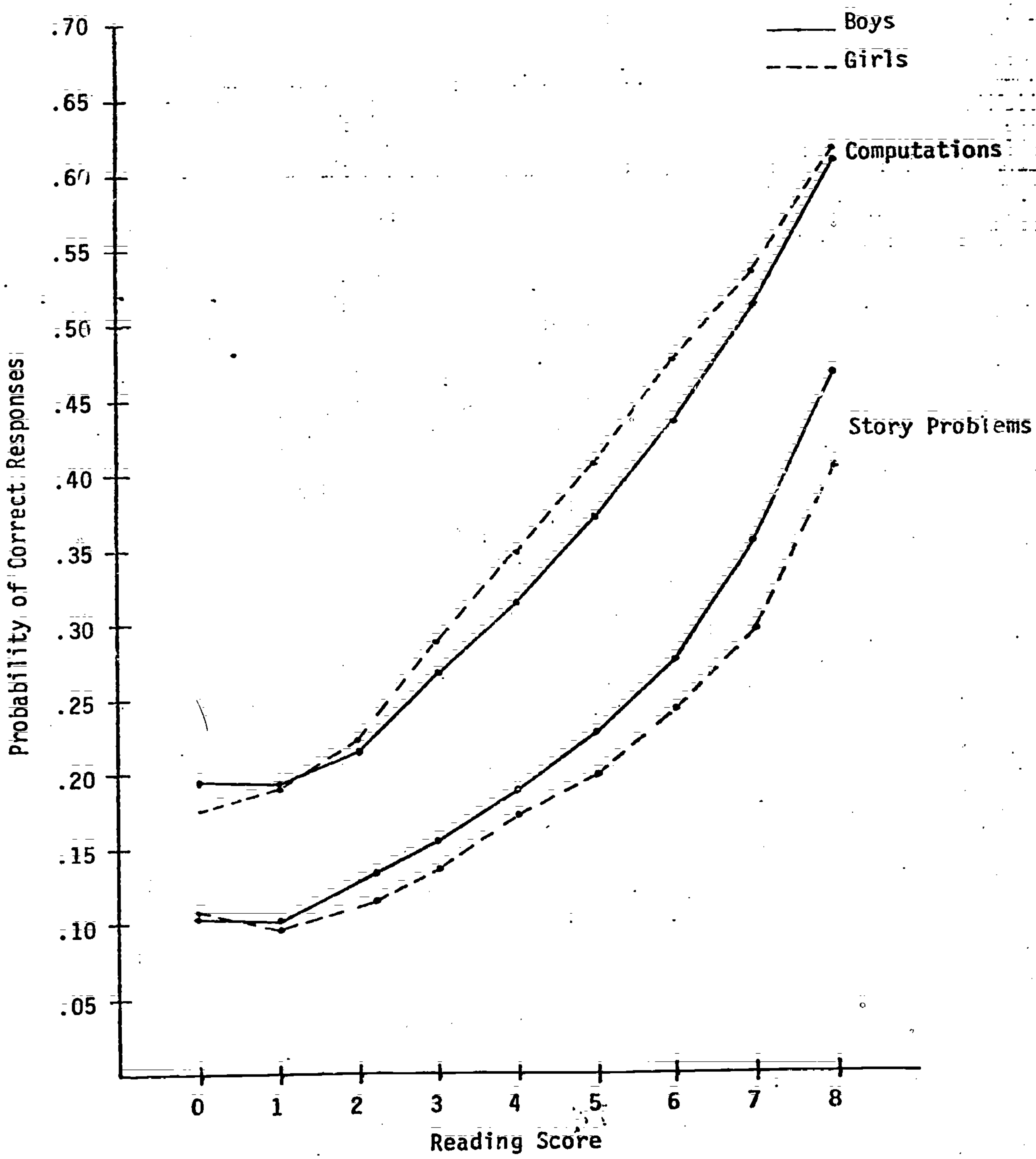


Figure 2
Average discrepancies for each sex between probability of answering all computations correctly given reading score and probability of answering all story problems correctly given reading score. Regression lines predict discrepancies from reading score for each sex.

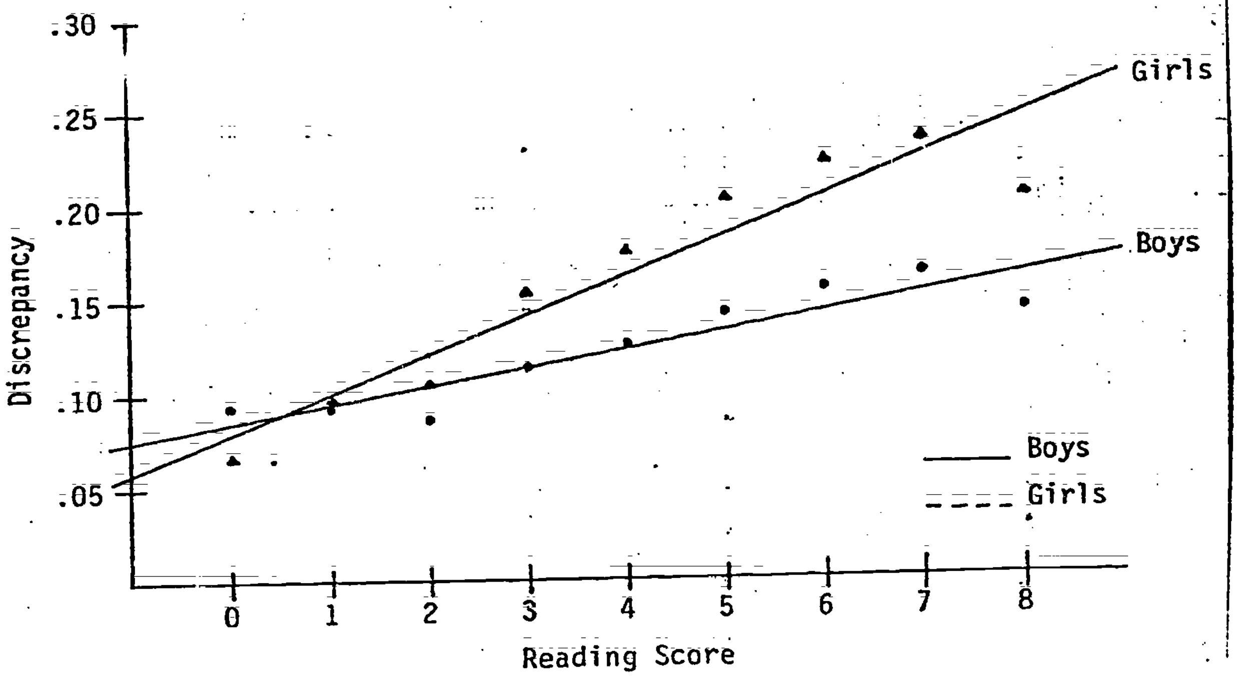


Figure 3
Average probabilities for boys and girls of answering all computations or all story problems correctly given socioeconomic level.

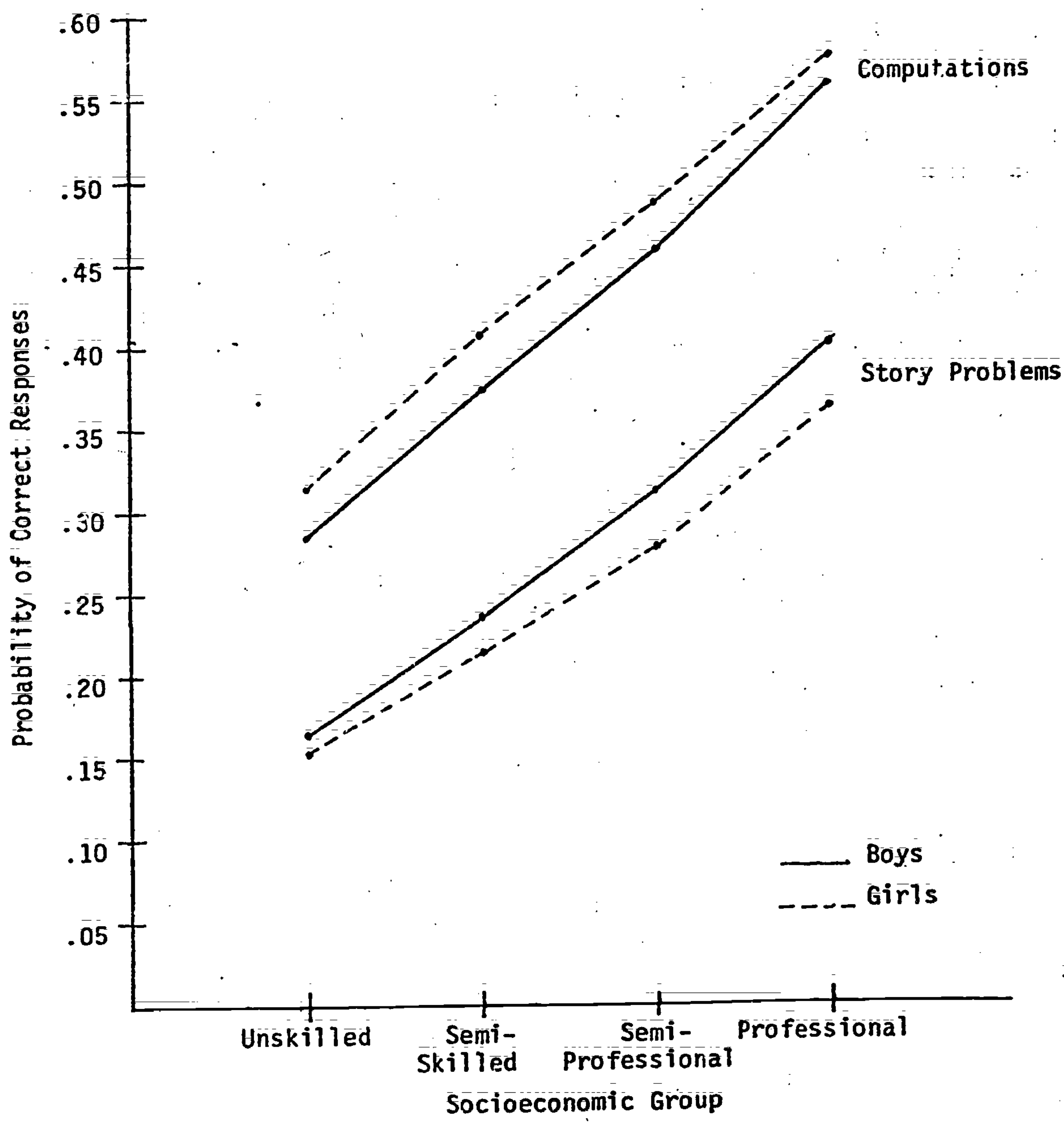
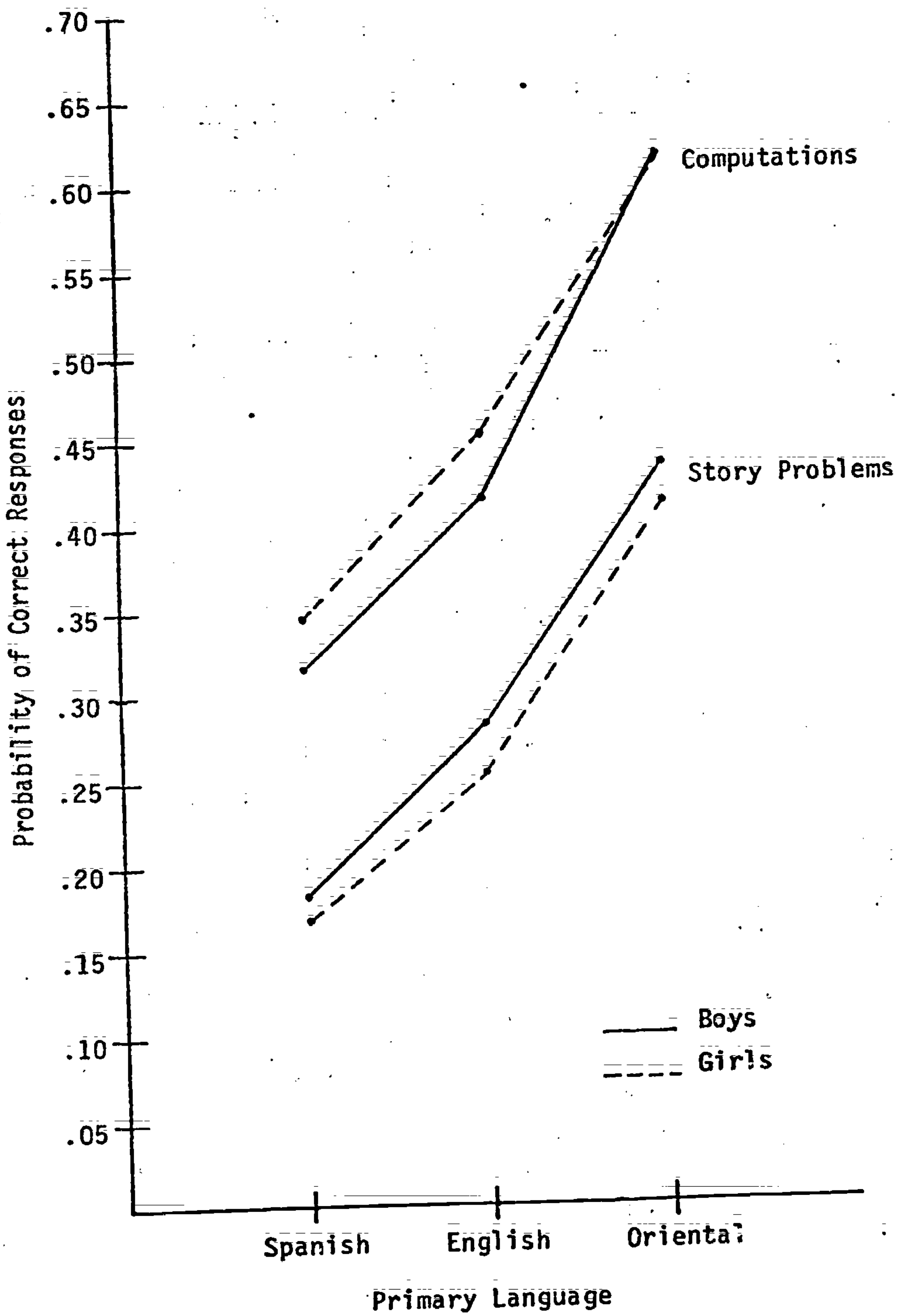


Figure 4
Average probabilities for boys and girls of answering all computations or all story problems correctly given primary language.



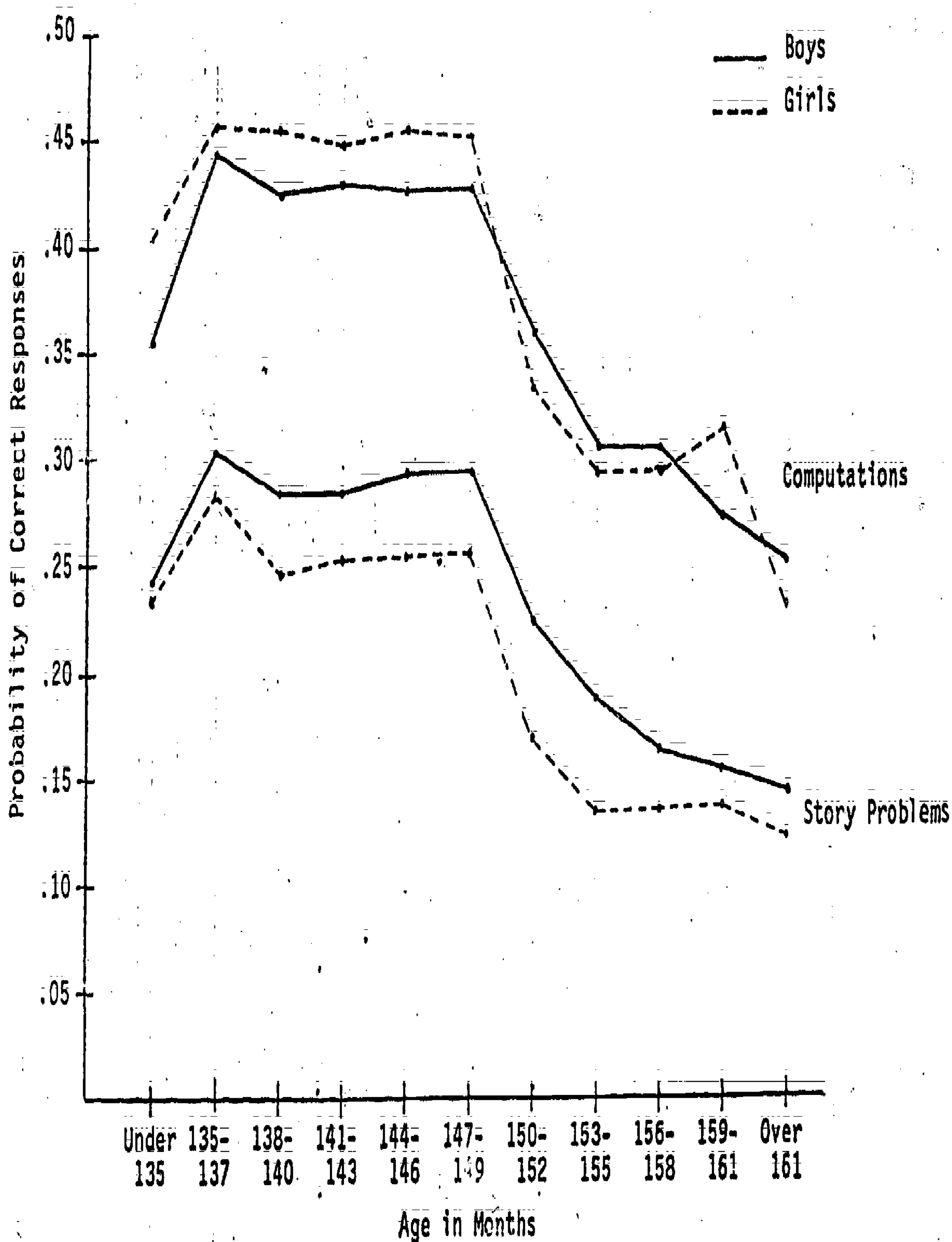


Figure 5

Average probabilities for boys and girls of answering all computations or all story problems correctly given chronological age.

Sex Differences in Mathematics Errors: An Analysis of Distractor Choices

Studies of sex differences in mathematics performance typically focus on differences in total test scores or on differential rates of success on particular types of items. Studies of total test scores have mixed results. Some suggest that differences do exist (Fennema & Carpenter, 1981; Benbow & Stanley, 1980; Maccoby & Jacklin, 1974; Hilton & Berglund, 1974; Backman, 1972). Others reveal little or no variation between boys' and girls' performance (Swafford, 1980; Fennema & Sherman, 1978; Fennema & Sherman, 1977). The second approach, studying the interaction of item type and sex, has yielded more consistent findings. It has produced evidence that girls are more successful than boys in solving computation items and that boys are more successful than girls in solving story problems (Marshall, 1981; California Assessment Program, 1979; Fennema, 1976; Armstrong, Note 1; Marshall, Note 2). Most recently, the National Assessment of Educational Progress (NAEP) found similar results (Fennema & Carpenter, 1981).

Both approaches have value in determining whether girls and boys are equally successful in solving mathematics items, but neither provides information about how students attempt to solve problems or about the errors they commit. There is instructional merit in knowing whether boys and girls differ in their perceptions of individual problems and consequently in their attempts to solve items. To this end, a new approach to the study of sex differences is pursued here: investigation of the errors made by each sex on individual items. Two questions are addressed: (1) Do girls and boys make the same or different errors in solving mathematics problems? (2) Do they use the same strategies as they solve problems?

Both questions are investigated by means of analysis of children's errors on a variety of mathematics items. There are two parts to the study. The first describes a statistical analysis of the stability of boys' and girls' errors. Data from four groups of sixth grade children are examined to determine whether there are characteristic errors made consistently by either sex. The second part of the study is an identification and classification of errors made by each sex. Six general categories of mistakes are identified, and the errors for each sex from a sample of eight mathematics items are analyzed according to these categories.

Test Instrument and Data

The items examined in this paper are taken from the Survey of Basic Skills: Grade 6. The Survey is a 30 item achievement test given annually to all sixth grade children in California through the California Assessment Program. Approximately 300,000 students respond to the test each year. There are sixteen forms of the test; roughly 9,000 girls and 9,000 boys respond to each form.

Each test form contains thirty unique items: eight of reading, eight of written expression, four of spelling, and ten of mathematics. The four areas are scored separately. Of the 160 mathematics items, there are 32 items on measurement and graphing, 28 items on number concepts, 28 items on whole number arithmetic, 20 items on fraction arithmetic, 20 items on decimal arithmetic, 20 items on geometry, and 12 items on probability and statistics. Additional details about the Survey are available elsewhere (California Assessment Program, 1979).

Table 1 about here

The data presented here were gathered in 1976, 1977, 1978, and 1979. Table 1 gives details of sample size. Mean mathematics scores for boys and girls in 1979 were 5.92 and 5.88, respectively, and standard deviations were 2.17 and 2.08. Results from the three previous years were similar. Clearly, there is no large difference in total mathematics scores, although a statistical test is highly significant ($p < .001$) because of the very large sample of 144,462 boys and 142,305 girls.

Distractor Analysis: A Statistical Model of Responses

The method of distractor analysis developed here involves three factors: sex, response, and year. Only incorrect answers are evaluated. The concern is with the proportion of incorrect boys and incorrect girls selecting each distractor. The year of response is included because it is valuable to know whether different groups of girls and boys having different instruction make similar errors. In analyzing the types of errors made by boys and girls, it is valuable to know whether the error patterns for each sex are stable or fluctuate each year.

The 160 mathematics items of the Survey are analyzed individually. For each item, there is a three-way contingency table of sex by distractor by year. For an item having five response alternatives (four distractors plus the correct answer), the corresponding contingency table would be 2 x 5 x 4. There are two values for sex (male and female), five values for distractors (each of the four incorrect choices plus the alternative of leaving the item blank), and four values for year (1976, 1977, 1978, 1979). Each of the 40 cells of the contingency table contains frequency counts. For data such as these, loglinear models are appropriate. Multiway loglinear models are similar to extensions of the usual χ^2 tests of association of two factors.[2]

Loglinear models contain main effects and interaction effects. The main effects capture unequal proportions of the variables. For example, a main effect for sex indicates unequal numbers of boys and girls responding. The main effect for distractor indicates that all alternatives were not equally selected. For the situation described here, main effects alone are of little interest since one already knows that slightly more boys than girls responded and that more children responded in 1976 than in 1979.

Main effects must be part of the model. The issue here is which interaction terms to include. With a three-factor model, there are four possible interactions: distractor by sex, distractor by year, sex by year, and the three way interaction of distractor by sex by year. The two-factor interactions have relatively simple interpretations. Inclusion of the sex-by-year term implies differing proportions of boys and girls for the various years. Like the main effects, the sex-by-year term is not especially interesting because it reflects only the differences in numbers of boys and girls responding in the four years.

The remaining two-factor terms have substantive interpretations. Inclusion of the distractor-by-sex term implies that there are indeed sex differences in the alternatives selected by the children for each item. The distribution of girls' choices for an item differs from the distribution of boys' choices. Similarly, inclusion of the distractor-by-year term implies that the distributions of both sexes over all distractors vary in the same way from year to year.

If there are stable sex differences in children's errors, the statistical model predicting children's responses must include main effects plus the interaction term of distractor-by-sex. The other two-factor interactions may or may not be part of the model. If the differences are stable, the three-factor interaction cannot be part of the model. Its presence indicates that responses for each sex are shifting each year.

Each of the 160 mathematics items was analyzed individually to determine the model that best predicted the frequency data. The models of best fit were identified by chi-square likelihood ratio tests. The results are given in Table 2. The model requiring only the two-factor interactions of distractor-by-sex and distractor-by-year was the most appropriate model for the majority of items (model 7 in Table 2). This model accounts for eighty percent of all items. Models 4, 6, and 8 also support the hypothesis of stable sex differences.

Only the ten items requiring the three-factor interaction (model 9) show inconsistency of responses by girls and boys over the years. Neither the single item requiring only main effects, the six items fit by the model with the single interaction of distractor-by-year, nor the two items having sex-by-year and distractor-by-year interactions show sex differences (models 1, 2, and 5). Thus, there are consistent sex differences in responses for all but 19 of the 160 items.

Table 2 about here

The interaction between sex and distractor choice affirms the main hypothesis of the present study. Boys and girls are selecting different answers. The differences are not simply between right and wrong choices. Rather, differences also exist in the wrong alternatives selected.

A Classification of Errors

A simplified model of solving a mathematics problem has three stages: recognition of problem structure or type from features of the item itself; identification of an appropriate strategy for use with the identified structure; and implementation of the selected strategy. Errors may occur in any of these phases. For example, an individual may attend to an irrelevant or secondary feature of an item and may ignore important information. The remaining stages might be correctly executed, but an error would result nonetheless. Similarly, an error in the second stage may arise if an inappropriate strategy is selected. The first stage may or may not have been successfully carried out. Finally, it is possible to recognize important features of an item and to identify an appropriate problem-solving strategy, and nevertheless to fail in executing the strategy. Obviously, these errors can occur singly or intermixed.

This section describes a classification of errors based on this simple model. Errors may be made in the following processes:

1. Attention to spatial and visual cues
2. Attention to verbal cues
3. Selection of inaccurate or inappropriate algorithm
4. Use of a pattern of response preference
5. Translation from words to arithmetic expressions
6. Perseverance in use of a chosen algorithm

Errors 1 and 2 are errors of recognition and reflect an individual's inability to identify necessary salient features of an item. Errors 3 and 4 indicate difficulties in selecting a problem-solving strategy. Either an inaccurate or inappropriate strategy is selected, or the student guesses according to a pattern of response preference. Finally, errors 5 and 6 are errors in implementation or execution.

The first five categories are related to research of sex differences. Spatial ability, verbal ability, computation and word problem solving, and persistence have all been investigated with sex as a variable. Boys tend to excel at spatial tasks and in solving word problems, and they tend to persevere longer than girls when performing a task (Fennema & Carpenter, 1981; Vandenberg & Kuse, 1979; Maccoby & Jacklin, 1974; Backman, 1972; Blum & Broverman, 1967; Marshall, Note 2). These studies also show that girls are likely to excel in verbal ability and computations. It is reasonable to hypothesize that boys' and girls' performance are related to these factors and, further, that boys' and girls' errors are also directly related.

The only category not investigated in the sex-related differences literature is that of response preference. There is evidence that individuals do have preferences for multiple-choice alternatives. College students, for example, favor response B given the choice of A, B, C, or D (van Heerden & Hoogstraten, 1979). One questions whether sixth grade girls and boys have response preferences and whether these preferences differ.

This classification is a useful basis for interpreting the sex differences that exist in the data. In many cases, the categories have obvious applications to items from the Survey. Groups of similar errors emerge when the responses are analyzed, and the classification proposed here captures most of the distractor alternatives.

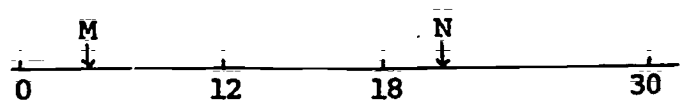
Examples of Errors in Each Category

To illustrate how the present classification distinguishes boys' and girls' errors, several examples are offered. Table 3 contains eight items from the Survey of Basic Skills. Two i[illegible]s each of four types were selected: word problems, geometry and scaling items, computations, and mathematics concepts. The distractors for these items are discussed below according to each category in the error classification.

Table 3 about here

Spatial and Visual Cues. Cues of this type are present in four of the eight items of Table 3 (items 2,4,5, and 6). Items 5 and 6 are the most obviously influenced by spatial features. Item 5 requires the student to recognize similar triangles. Shape and size of the five figures in the item are important visual cues. Shape is a relevant feature, and size is irrelevant. Attention to the former leads to the correct solution, while attention to the latter leads to an incorrect response of D (the smallest of the figures in the response choices).

In Item 6, representation of the intervals as equal on the number line is important in understanding the item. The concept of 'number line' is not necessary; a student need only recognize the equal spacing in the diagram and interpolate the missing values. Were the figure drawn with the letters M and N shifted to the left, as



the student could not assume equal intervals and the response '3 and 21' could be correct. All distractors in this item reflect failure to attend to relevant visual information.

The visual information of Item 2 (apples) is of a different nature and applies only to the choice of '4' as the answer. In this case, students may realize that '6' would be too great a number of apples and that anything fewer would be satisfactory. Scanning the list of distractors results in a response of '4', the number closest to but not exceeding '6'. The last option 'None of these' is not considered, perhaps because it is not a numerical value.

Item 4 (decimal addition) is an example of horizontal arithmetic. Since the majority of incorrect responses to this item were 'None of these', it is impossible to infer which arithmetic solutions were obtained. One possibility is that students failed to align the numbers correctly or lost track of decimal positions. A response of 13.104 (distractor A) may result from the spatial orientation of the item. Students reaching this solution may perceive the problem as (7 + 6) . (58 + 46) = 13 . 104, in effect creating a distributive property for the decimal point.

Verbal Cues. Verbal cues are individual words or phrases in an item. Key words may determine a student's choice of operation. In Item 1 (map scale), the words 'If ... is ..., then ... represents [unknown]' lead the student to formulate the problem as:

<u>X</u> is to <u>Y</u> as <u>Z</u> is to [<u>unknown</u>]

where <u>X</u> and <u>Z</u> are inches and <u>Y</u> and the <u>unknown</u> are miles Setting up the problem in this way yields an equation of

<u>X/Y</u> = <u>Z/unknown</u>

and leads to the correct solution.

Items 2,5, and 7 also contain key words that may guide an individual's choice of strategy. In Items 2 and 7 the word 'greatest' appears. If one attends only to that word, one is led to the choice of '47' in Item 2 (apples) and to '8' in Item 7 (GCD). Similarly, the word(s) 'common' or 'common divisor' in Item 7 suggest the response '2'. A student focusing on this cue would have no motivation to examine other response alternatives since the first one is satisfactory.

Spatial information in Item 5 (similar triangles) has already been mentioned. However, the word 'similar' may be misleading for children who have not learned the geometric concept. They may focus on the word and interpret it to mean "closely resembling but not identical". In that case, the correct figure of alternative C might be eliminated from consideration because its shape is identical to the figure in the stem. Students may then look for a figure that only approximates the original one in size or shape or both.

Inaccurate Algorithms. An algorithm is a procedure for making an arithmetic calculation. Errors from inaccurate algorithms include more than careless arithmetic errors. Systematic misplacement of decimal points and incorrect rules in 'carrying' or 'borrowing' are also examples.

Several items of Table 3 have distractors that correspond to inaccurate algorithms. All distractors of Item 1 (map scale) coincide with incorrect decimal placement. In this item, the incorrect algorithms co-exist with other errors, such as verbal or translation errors.

Inaccurate algorithms are also probable in Items 3, 4, and 7. All distractors for Items 3 and 4 represent computational error; however, the nature of the errors differs. In the vertical addition of Item 3, response A corresponds to a procedure that ignores 'carrying' altogether.[3] (A less likely alternative leading to this distractor is to ignore the middle number and combine only the leftmost and rightmost numbers for a response of 1,212.) Response C also can result from one of two procedures. In the first, the student adds 4 + 8 = 12, records '2' and carries '1' to the leftmost column. The student then adds 4 + 7 = 11, records '1', and again attempts to carry '1' to the leftmost column. However, 1 has already been added to this column (from the first carry), so the second carry is disregarded. The final step adds 1 + 5 + 7 = 13, and the result is the value 1,312.

A second alternative yielding the same final solution corresponds to inconsistent use of the carrying procedure. The process is ignored in the first step and is then employed correctly in the addition of the middle pair of numbers.

In the horizontal addition of Item 4, the predominant response is 'None of these'. As pointed out above, the incorrect solutions obtained by most students cannot be observed. Response A is also a popular choice and illustrates a particular incorrect algorithm. For this answer, children correctly perceive the problem but formulate the response as

7	.	58
+ 6	.	46
13	.	104

adding separately the numbers to the left and to the right of the decimal. This approach was discussed above in reference to the horizontal appearance of the item.

Response Preferences. One cannot ignore the possibility that children select certain responses on a multiple-choice test as a result of the arrangement of the distractors or as part of a general guessing strategy. For example, a student may select consistently the first or last alternative or may choose the smallest or largest value. Since distractors are frequently listed in ascending or descending order, it is difficult to tell whether a student following this procedure is selecting the first or the smallest value (since they are frequently identical). Predominant errors on the items of Table 3 are more frequently first or last choices rather than middle ones.

A second response preference may involve the selection of arithmetic operation on the basis of relative size of the numbers in the stem. This response is almost wholly context-free; the student picks out numbers and combines them according to rules of [illegible]. For example, in Item 1 (map scale) a student using this approach would focus only on the numbers 350 and 14. Since one is much larger than the other, the student elects to divide the smaller into the larger value rather than add or subtract or multiply the two. A student using this approach would get the correct solution.

This strategy results in error for Item 2 (apples). The response preference in this item appears to be add or subtract two amounts of money. Given 58¢ and 11¢, one reaches either 69¢ or 47¢. The solution 69 is not an available option, and thus 47 is selected.

Translation. In solving word or story problems, students generally must translate the problem from the verbal statement into a mathematical expression. Errors in translation result in incorrect solutions even though problem-solving techniques used subsequently are employed correctly. Items 1 and 8 illustrate the possibilities of such errors.

Consider Item 1 (map scale). One correct representation of this item was given above. Should a student take the numbers in the order presented in the stem and insert them into the equation above,

X is to Y as Z is to [unknown],

the result is:

350 is to 14 as 1 is to [unknown],

or

350/14 = 1/ unknown.

To solve this equation, the student must divide 14 by 350. By misplacing the decimal, he or she reaches '4' or '40' as an answer.

A similar translation error can be made on Item 8 (numeral). All errors on this item may be translation errors. The first response B) is simply the reversal of the digits 3 and 4. A literal translation of seven thousand six hundred thirty-four might be 7,000,600,34. Several variations are possible. One might expect most children to translate correctly six hundred-thirty four. (In any case, incorrect translations other than 643 are not options for this item.) Therefore, a student might translate the phrase to 7,000.634 (again, not a distractor choice) or the nearest value of 700,634 (response D).

Perseverance. It is possible to explain some of the errors of Table 3 in terms of student persistence in solving the items. Items 3, 6, 7, and 8 have distractors that correspond to lack of perseverance in repeating a correct process when necessary.

In the addition problem of Item 3, response B may be explained in the following way. The first step in carrying is executed correctly. However, the second carrying procedure is carried out incorrectly. The final result is then 1,222. One explanation of the error is that students failed to persist in performing the carrying operation.

A second example of failure to persevere is given by distractor B of item 6 (number line). In this selection, the student has correctly identified the value corresponding to M as 6 but then does not use the same process again to identify the value corresponding to N.

A third example is found in item 7 (GCD). The response '8' would be correct if the values in the stem were only 8 and 16. However, 8 is not a factor of 12. Choice of the response D (8) may result from failure to determine that 8 is indeed a factor of all three numbers and not just the first and last ones.

Failure to persevere may be related to lack of attention. The error of translating 'seven thousand six hundred thirty-four' into 7,643 (response B of item 8) may be an error of this type rather than a problem in translation. Having correctly interpreted the first two values (seven thousand and six hundred), the student may lessen his or her attention to detail.

Sex Differences In Errors

Table 4 summarizes the errors by sex to the items of Table 3. A distractor may appear in more than a single category; it is sometimes possible to reach the same incorrect solution because of different errors. Interaction among categories is also possible. For example, verbal and spatial features may combine to occasion a particular error.

Both sexes make errors in all six categories. For the items of Table 3 girls are more likely to make errors in recognizing relevant features of an item, whether spatial or verbal. More of their errors can be classified in the first two catetories than in those associated with selection or implementation of correct strategies. Boys have more difficulty in implementing a chosen strategy.

Boys appear to make more errors than girls from spatial and visual cues, from translations, from inaccurate algorithms, and from lack of persistence. These results occur in part because boys frequently show preference for two distractors while girls generally have a single preference. Items 2, 4, and 7 illustrate this point.

Within individual categories, distinctions are evident. Boys have a greater tendency than girls to make errors in items with unusual spatial orientation. Both sexes err on items with diagrams but in different ways. Both sexes also err in recognizing different verbal cues, although they attend to different ones.

Boys make more mistakes corresponding to errors in translation and from inaccurate algorithms than do girls. The former is somewhat surprising since boys generally perform better than girls in solving word problems (Fennema & Carpenter, 1981; Marshall, 1981; Marshall, Note 2). The present results suggest that girls' weakness is in something other than the translation step. This weakness appears at an earlier stage of the process, before the translations and algorithms are needed.

Very few items of the <u>Survey</u> had uniformly distributed errors. It is possible that many students have response preferences that are used when the students realize that they are unable to solve specific problems. The errors in the category of response preferences indicate that boys and girls hve different response preferences. Girls tend to select the first alternative (usually the smallest value) while boys tend to select the last one (usually the largest).

Finally, the responses to the sample items suggest that perseverance may be a factor in the particular errors made by children and especially by boys. It can be argued that this result is occasioned by the nature of the data: children may be less motivated to perform maximally on assessment tests than on other tests of achievement. However, as long as girls and boys respond to the same situation, one may assume their motivation is similar. Assessment data reveal typical errors, and it is these that are of interest here. The perseverance findings do not support the body of research that suggests boys pers[illegible] to a greater degree than do girls.

The difference between this finding and those of other studies may be due to the narrower definition of perseverance used here. All errors could in principle be defined as errors in perseverance: failure to detect all key words could be interpreted as lack of persistence in reading the entire problem, failure to select an appropriate strategy could be interpreted as lack of persistence in scanning all available strategies, and failure to apply an algorithm correctly could be interpreted as lack of persistence in repeatedly following the appropriate steps of the algorithm. Only the latter is taken to be a problem in perseverance in this paper. The other instances require inferences of omission; one assumes that all words were not read or that all strategies were not evaluted. As yet, there is no evidence to support these assumptions.

<u>Summary</u>

There are substantial differences in the errors made by sixth grade boys and girls on multiple-choice mathematics items. The statistical analysis reveals that boys and girls select different distractors to a variety of items and that the choices for each sex are consistent for different groups tested in separate years.

A classification of errors has been presented that emphasizes areas in which the sexes are thought to differ. Neither sex makes only one type of error, but within each classification, boys and girls nevertheless make different mistakes. Boys appear more likely to make errors in selecting and implementing problem-solving strategies than girls. Girls are more likely to err in determining the relevant features of problems. More of their errors occur in the first two categories than in the middle or last two ones.

It should be stressed that the above discussion does not capture all possible ways children err on the items of Table 3 nor is it argued here that only the errors described lead to selection of particular distractors. What is argued is that we can learn much about student errors from multiple-choice responses. Any study of sex differences in problem solving might benefit from examination of already-existing test data. Variables to be studied could be selected on theoretical and empirical grounds. Such a procedure would be time-saving in the long run because it can identify general areas and variables in which the sexes differ and can generate testable hypotheses about them. More detailed analyses with specially constructed items and small groups of individuals naturally follow.

Finally, it should be pointed out that the items studied here were not constructed to facilitate distractor analysis. Nevertheless, much information has been gained from them. More information would be available from test items if attention were given to details such as including distractors corresponding to more distinct strategies or phrasing the questions in such a way as to eliminate overlapping explanations for a single distractor choice. Any theory of problem solving could be better evaluated by administration of tests with appropriately chosen distractors.

Reference Notes

1. Armstrong, Jane M., Achievement and Participation of Women in Mathematics: An Overview. Report of a Two-Year Study Funded by the National Institute of Education, Report 10-MA-00. Denver, Colorado: Education Commission of the States, 1980.

2. Marshall, S. P. Sex Differences in Children's Mathematics Achievement: Solving Computations and Story Problems. Manuscript submitted for publication, 1982.

References

Backman, M. E. Patterns of mental abilities: Ethnic, socioeconomic, and sex differences. American Educational Research Journal, 1972, 9, 1-12.

Benbow, C. & Stanley, J. Sex differences in mathematical ability: Fact or artifact? Science, 1980, 210, 1262-1264.

Bishop, Y. M., Fienberg, S. E., & Holland, P. W. Discrete Multivariate Analysis: Theory and Practice. Cambridge, Massachusetts: The MIT Press, 1975.

Blum, A. H. & Broverman, D. M. Children's cognitive style and response modification. Journal of genetic psychology, 1967, 110, 95-103.

California Assessment Program. Student Achievement in California Schools: 1978-1979 Annual Report. Sacramento, California: State Department of Education ,1979.

Fennema, E. L. Influences of selected cognitive, affective, and educational variables on sex-related differences in mathematics learning and studying. In Women and mathematics: Research perspectives for change (National Institute of Education Papers in Education and Work: No. 8). Washington, D.C.: U.S. Government Printing Office, 1977.

Fennema, E. L. & Carpenter, T. P. Sex-related differences in mathematics. In Corbitt, M. K. (Ed.), Results from the Second Mathematics Assessment of the National Assessment of Educational Progress. Reston, Virginia: The National Council of Teachers of Mathematics, 1981.

Fennema, E. L., and Sherman, J. Sex-related differences in mathematics achievement, spatial visualization and affective factors. American Educational Research Journal, 1977, 14, 51-71.

Fennema, E. L. & Sherman, J. Sex-related differences in mathematics achievement and related factors: A further study. Journal for Research in Mathematics Education, 1978, 9, 189-203.

Hilton, T. L., & Berglund, G. W. Sex differences in mathematics achievement, a longitudinal study. Journal of Educational Research, 1974, 67, 231-237.

Maccoby, E. & Jacklin, C. The Psychology of Sex Differences. Stanford, California: Stanford University Press, 1974.

Marshall, S. P. Sex differences in sixth grade children's problem solving. Santa Barbara, Ca.: University of California, Santa Barbara, 1981. (ERIC Document Reproduction Service No. ED 200 649) Report ED-200-649, ERIC Reports, April, 1981.

Suppes, P. & Morningstar, M. Computer-assisted instruction at Stanford, 1966-68: Data, models and evaluation of the arithmetic programs. New York: Academic Press, 1972.

Swafford, J. Sex differences in first-year algebra. Journal for Research in Mathematics Education, 1980, 11, 335-345.

Vandenberg, S. G. & Kuse, A. R. Spatial ability: A critical review of the sex-linked major gene hypothesis. In Wittig, M. A. & Petersen, A. C. (Eds.), Sex-related differences in cognitive functioning. New York: Academic Press, 1979.

van [illegible], J. & Hoogstraten, J. Response tendency in a questionnaire without questions. Applied Psychological Measurement, 1979, 3, 117-121.

Footnotes

1. This research was supported by the National Institute of Education under Grant No. NIE-G-80-0095. The findings and conclusions presented here do not necessarily reflect the views of the National Institute of Education.
2. Details of loglinear analysis may be found in Bishop, Fienberg, and Holland, 1975.
3. The errors in this problem are similar to those described for younger children by Suppes and Morningstar (1972). However, they do not separate results for boys and girls.

Table 1

Number of Children Responding to the
Survey of Basic Skills:
1976-1979

	Boys	Girls
1976	166,204	163,868
1977	157,013	153,710
1978	148,801	145,647
1979	144,462	142,305

Table 2

Models of Best Fit for 160
Mathematics Items

Model	Number of Items for Which This Model is the Best Fit
(1) D + Y + S	1
(2) D + Y + S + DY	6
(3) D + Y + S + SY	0
(4) D + Y + S + DS	4
(5) D + Y + S + SY + DY	2
(6) D + Y + S + DS + SY	2
(7) D + Y + S + DS + DY	128
(8) D + Y + S + DS + DY + SY	7
(9) D + Y + S + DS + DY + SY + DSY	10

Key: D = distractor choice
Y = year of response
S = sex of student

Single letters indicate main effects.
Double letters indicate interactions.

Table 3

Mathematics Items Answered by Sixth Grade Students

Word Problems:

(1) If a distance of 350 miles is represented by a segment of 14 inches on a map [illegible] on the map 1 inch represents:
[a] 4 miles
* [b] 25 miles
[c] 40 miles
[d] 250 miles

Percentages Selecting Each Distractor:

	[a]	[c]	[d]	Total Incorrect
boys	31	37	32	13,689
girls	32	30	38	14,674

(2) Sue has 58¢. If apples cost 11¢ each, what is the greatest number of whole apples that Sue can buy?
[a] 47
[b] 6
[c] 4
[d] 3
* [e] None of these

Percentages Selecting Each Distractor:

	[a]	[b]	[c]	[d]	Total Incorrect
boys	40	20	34	6	13,466
girls	51	16	27	6	13,758

Computations:

(3) 744
+578

[a] 1,212
[b] 1,222
[c] 1,312
* [d] 1,322

Percentages Selecting Each Distractor:

	[a]	[b]	[c]	Total Incorrect
boys	12	42	46	2,982
girls	8	49	43	1,873

Table 3 continued

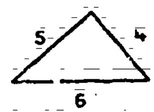
(4) $7.58 + 6.46 =$

[a] 13.104
* [b] 13.04
[c] 13.4
[d] None of these

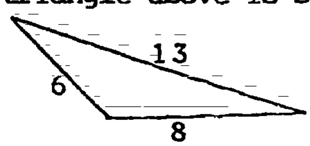
Percentages Selecting Each Distractor:

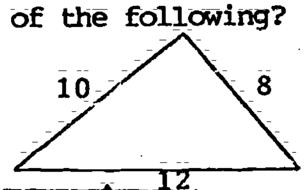
	[a]	[c]	[d]	Total Incorrect
boys	30	10	60	6,703
girls	14	7	79	4,770

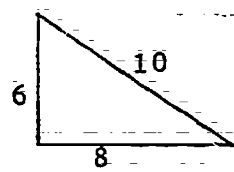
Geometry and Scaling:

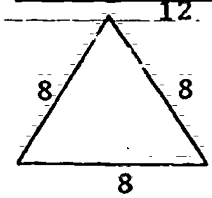


(5) The triangle above is similar to which of the following?

[a]


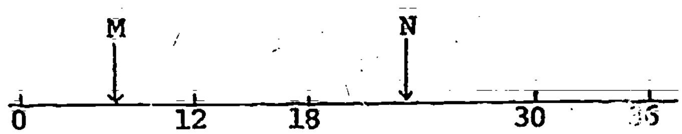
* [c]


[b]


[d]


Percentages Selecting Each Distractor:

	[a]	[b]	[d]	Total Incorrect
boys	5	45	[illegible]	7,680
girls	4	53	[illegible]	8,078

(6)


On the number line above, lettered arrows M and N point to the missing numbers:

[a] 3 and 21
[b] 5 and 20
* [c] [illegible] and 24
[d] 9 and 24

Percentages Selecting Each Distractor:

	[a]	[b]	[d]	Total Incorrect
boys	27	53	20	6,022
girls	32	46	22	6,550

Table 3 continued

Mathematical Concepts:

(7) What is the greatest common divisor of 8, 12, and 16?
[a] 2
[b] 3
* [c] 4
[d] 8

Percentages Selecting Each Distractor:

	[a]	[b]	[d]	Total Incorrect
boys	49	7	44	10,300
girls	58	5	37	10,537

(8) Which is a numeral for seven thousand six hundred thirty-fo[illegible]?
* [a] 7,634
[b] 7,643
[c] 70,634
[d] 700,634

Percentages Selecting Each Distractor:

	[b]	[c]	[d]	Total Incorrect
boys	37	28	35	3,362
girls	30	23	47	3,178

* correct response

The items are reproduced here by permission of the California Assessment Program of the California Department of Education.

Table 4

A Summary of Student Errors on Eight Mathematics Items According to Six Categories of Errors

Category	Item	Response	Sex Selecting Response
(1) Spatial-Visual	2	C	boys
	4	A	boys
	5	B	girls
	5	D	boys
	6	A	girls
(2) Verbal	1	D	girls
	2	A	girls,boys
	5	B	girls
	5	D	boys
	7	A	girls,boys
	7	D	boys
(3) Inaccurate Algorithms	1	D	girls
	1	C	boys
	3	B	girls
	3	C	boys
	4	A	boys
	7	D	boys
(4) Response Styles	2	A	girls
	6	A	girls
	7	A	girls
	4	D	boys
	5	D	boys
	7	D	boys
	2	A	boys,girls
(5) Translation	1	C	boys
	8	B	boys
	8	D	girls
(6) Perseverance	3	B	girls
	6	B	girls,boys
	7	D	boys
	8	B	boys

Error Analysis and Cognitive Processes

In this chapter, I describe cognitive strategies used by sixth-grade children in solving mathematics problems. There are two parts to the study. First, a large number of test responses by sixth-grade children are examined. Error analysis in a large set of data allows determination of characteristic mistakes made by boys and by girls. In this study, the basic question becomes whether boys and girls make different or identical errors on items.

The second study builds upon the results of the first. Ninety-three sixth-grade students were interviewed individually and their problem-solving styles and strategies were observed. The first study indicates the types of problems on which boys and girls differ. The second is a detailed examination of steps the children take in solving these problems.

Other Studies of Errors

Error analysis is a relatively new field in mathematics research. An interesting line of research is offered by several Australian researchers using a classification of errors that includes: reading deficiencies, comprehension problems, transformation errors, encoding or writing errors, and carelessness (Clements, 1980; Watson, 1980; Newman, 1977). A second approach (Radatz, 1979) uses a classification grounded in information-processing theory. Possible errors are those of language difficulties (semantics), spatial representations, lack of math knowledge, rigidity of thinking, and application of irrelevant rules.

More recently, I have proposed a classification based upon a three-stage model of problem solving: recognition of features of an item, selection of a strategy to be used in solving the item, and implementation of the selected strategy (Marshall, Note 1). In each stage, there are two possible types of errors. In the first stage of recognition, errors are those related to spatial/visual cues or verbal features of the item. In the second stage, errors correspond to selection of inappropriate algorithms or to use of guessing patterns for solving the item. The latter are context-free and depend only upon the relative size of the numbers in an item or in the list of response alternatives on a multiple-choice item. The student bases his or her choice of arithmetic operation upon the perceived relationship of the numbers. Finally, there are also two types of errors occurring in the third stage of the problem-solving model: those of translation from words to equations and those of persistence. Discussion of errors in this chapter will refer frequently to the three-stage model and the associated classification of errors.

Error classification offers a framework for examining cognitive processes in problem solving. By focusing on particular types of errors, we may discover the point at which a child's processing makes a wrong turn. We need no longer be concerned only with whether the child answers correctly or incorrectly, and we can devise instruction aimed at specific elements of cognitive strategies.

Study 1: Errors on Multiple-Choice Tests

The first study is an examination of a large number of responses given by children to a standardized assessment test. The responses are those of all sixth-grade children enrolled in California public schools during the period 1976-1979. Approximately 300,000 students responded in each year.

The test instrument is the Survey of Basic Skills: Grade 6, a 30-item assessment test covering reading, written expression, spelling, and mathematics. Details of the test are given by the California Assessment Program (1979). Only the mathematics items are of interest here.

The Survey is administered annually by the California Assessment Program of the California Department of Education. Each child responds to one of 16 different forms of the test, and each form contains unique items. There are 10 mathematics items on each form, yielding a total of 160 mathematics items for the entire Survey.

Study of the 160 mathematics items reveals substantial sex differences in the errors made by boys and girls. Distributions of boys' and girls' errors are significantly different for 88 percent of the items. Two examples of the items analyzed and the differences in boys' and girls' responses are given in Table 1. During the four years studied, a total of 13,689 boys and 14,674 girls erred on the word problem, and 7,680 boys and 8,078 girls erred on the geometry item. Proportions of each sex selecting each incorrect distractor are given in Table 1. These proportions are essentially the same for each of the years studied. That is, girls were more likely to select [d] on the map item in all four years, and boys were more likely to select [c]. Similarly, girls tended to prefer [b] for the geometry item, while boys were more likely to respond [d].

Consider the ways of deriving answers to the word problem (Table 1). There are two methods that yield the correct solution. First, a student can ignore everything in the problem except the numbers 350 and 14. Since one is large and the other is relatively small, the student elects to divide and reaches the correct solution of 25 without ever attempting to encode or translate the problem. Such a strategy of problem solving is an example of a guessing pattern. The student selects the arithmetic operation on the basis of the relative size of the numbers in the item. The student's preference of operation is based upon rules such as: add all the numbers together if more than two are present; subtract the smaller from the larger if two numbers have the same number of digits; or divide the smaller into the larger if one is much larger than the other. Direct evidence of such guesswork is described in the second part of this chapter. It is is important in this item. There are five geometric figures; the student is asked to identify which of four is similar to the fifth. There is also verbal information in the item. The word similar usually means "closely resembling." Here similar means "having sides with the same proportions." If it is interpreted as "closely resembling but not identical", the meaning attached by the student to the word similar may direct the choice of response.

There are two spatial features of the item to which the student can attend: shape and size of the figures. In this case, shape is relevant and size is not. Students searching for a similar shape will elect response [c]. Students seeking a similar size will select the smallest of the alternatives, [d]. Students interpreting similar as "close but not identical" may select either [b] or [d], depending upon whether size or shape was the more salient visual feature.

Table 1 shows the differences in boys' and girls' responses. Boys are more likely to respond [d] while girls select [b]. Thus, the boys appear to be responding to the inappropriate spatial feature of size, while girls may attend to erroneous verbal information and the relevant spatial feature of shape. This suggests that when spatial and verbal information occur simultaneously in a problem, girls and boys have selective attention for different features.

Analysis of students' choice of distractor provides useful information about the types of errors students make and about possible cognitive strategies used in solving problems. Further information is gained from interviews with students as they solve individual problems and describe the problem-solving process. The advantage of studying multiple-choice responses from standardized tests is that responses from many students to many types of items can be examined. The advantage of studying interview responses is that the responses of a few students to a few items can be explored in detail.

Study 2: Student Interviews

For the second study, 93 sixth-grade students participated individually in one interview session. Each student first responded to a ten-item paper-and- pencil test and then answered questions about ten additional items posed by the interviewer. Each set of ten items was used as pretest for approximately fifty percent of the children and as interview for the remaining students.

During the interview, a student was encouraged to describe his or her problem-solving strategies and was asked frequently to justify the use of particular operations and to interpret the results. Students responded orally in the interview but were allowed to write their computations as they talked. Each interview lasted approximately 45 minutes. Most of the children responded to the full set of twenty items, although a few did not an important source of errors and also, as here, of correct answers.

A second way to reach a correct solution is by following the simple model of problem solving mentioned previously. A student first recognizes the salient features of the item. In this case, the words

"If ... is represented by ... , then ... represents ... "

lead to selection of the strategy that sets up the proper proportion:

A is to B as C is to D.

Other salient features that must be noted in the item are the units of measurement: 14 and 1 are inches, 350 and the unknown quantity are miles. Finally, correct solution requires combining the salient features and the strategy selected:

14 is to 350 as 1 is to unknown,

or

14/350 = 1/ X,

where X is the unknown number of miles.

The cognitive process described above may fail in several ways. An individual may fail to recognize all the necessary verbal cues present in the item. For example, the structure of the problem may be correctly identified (leading to the appropriate selection of a strategy) but the units of measurement may be ignored. Were this to happen, the individual might easily make errors in setting up the appropriate equation. If the numbers in the problem are inserted into the equation in the order they occur in the item, the result is

350 is to 14 as 1 is to unknown .

Such a formulation leads to the incorrect response of 0.04. This answer is not an option in the list of distractors, but the values 4 and 40 are. Arbitrarily shifting the decimal allows one to select either of these. A similar error in decimal placement results in the value of 250 (choice [d]) rather than the correct response of 25.

From Table 1, it is evident that boys are more likely to respond '40' when they are incorrect and girls are more likely to respond '250'. Both distractors contain errors of decimal arithmetic, but they correspond to different approaches in cognitive processing. Although the girls' choice of 250 is in some sense 'closer' to the proper formulation (since the proportion was correctly set up), it may be the more serious error because it is more illogical. Given an initial value of 350 miles for 14 inches, it should be obvious to students thinking about relative size and distances that a response of 250 is much too large for a single inch. Selection of the value 250 indicates a serious lack of attention to the logical or common sense value of an answer. Certainly, a response of 40 is more plausible. It is possible that boys are attending more to the logical structure of their answers than girls.

The errors on the geometry item illustrate additional cognitive strategies. Undoubtedly, spatial and visual information removed. For the baseball cards and the hair ribbons item, the portion removed is a fraction of the original quantity. For the soda and bus items, the portion removed is expressed as a ratio: one in three or one in six. In all cases, the student must first determine the number of units to be removed and second compute the number of remaining units. As shown in Figure 1, the difficulty in this item is in determining the number of units to remove. When this value is known, the item has a simple structure and can be solved easily by subtraction. When the number of units to be removed is not known, the student must decide how to find it.

Table 4 provides detail of responses to the individual items. Four categories of errors are defined. First are those errors that yield only the amount to be removed from the original quantity. Individuals perform the first step only on an item that has a multi-step solution. Second are errors that occur when an individual attempts to subtract a fraction or ratio from the original quantity without first finding a common unit of measure. For example, in the baseball card problem, such an error is the subtraction of 1/3 from 12, yielding an answer of 11 2/3. The third category is a miscellaneous one; all other errors are included here, except the response of 'I don't know', which is the fourth category. For these items, miscellaneous errors comprise only five to eight percent of total responses. Several children gave no answer to the items, and their problem-solving strategies remain unknown. However, most of the errors come from failure to carry out the second operation of subtraction or from failure to recognize the appropriate number to be subtracted. These errors reflect differences in the processing of information in the items.

Table 4 also shows the relative difficulty of these items for boys and girls. The ribbons item is the easiest of the four for boys and the most difficult one for girls. A large number of girls responded "I don't know" to this item.

Given the structure common to all four items, we questioned whether students were consistent in their cognitive processing. There are two issues of consistency: first, do students approach the pairs of items in the same way? The matched items are nearly identical. One expects similar behaviors for similar items. Different responses (e.g., use of different arithmetic operations) suggest that a student does not recognize the underlying semantic structure of the pair of items. In effect, the student is guessing or using a strategy that is context-free.

The second issue of consistency concerns whether students tend to use one strategy for all four items. In practice, very few students were able to solve all four items. Therefore, we modified this issue to include students consistent on at least three of the four items. A comparison of the numbers of students consistent within pairs of items with those consistent over pairs provides an estimate of how many students perceive the generic structure of the two types of items. Word similarities alone may account for recognition that the cards and ribbons items are alike. An understanding or encoding of the problem into a have sufficient time to do so. Failure to complete the set of items appears to be more highly related to the amount of detail the children provided on earlier problems than to ability in solving problems. Individuals providing a great deal of information or those making many irrelevant statements tended to run out of time on the ninth and tenth interview items.

The problems solved by the students are word problems requiring fraction or whole number arithmetic. Most of the items are complex and require more than a single computation for correct solution. Boys were slightly more successful in solving the problems than were girls, but a comparison of mean scores on the paper and pencil test was statistically insignificant. The mean score for boys is 3.86 and that for girls is 3.12. Standard deviations are 2.04 and 2.08 respectively. (A comparison of number of items answered correctly in the interviews was not carried out because not all children had the opportunity to respond to all items. Frequently, the items not reached were the easiest ones in the set and would presumably have been answered correctly.)

General Results

An item-by-item comparison of boys' and girls' success on all twenty items (pretest responses only) shows that boys were more successful than girls on eighteen of the items and girls were more successful than boys on only two of them. In most cases, however, the difference in proportions correct for the sexes is not statistically significant. On only three items were boys significantly better than girls.

The items of Table 2 are four of the items used in the study. Seventy-nine students (39 boys and 40 girls) responded to all four of the items, two on the paper-and-pencil pretest and two in the interview. Only the responses from these students are discussed here. Notice that the items are matched: the baseball cards and hair ribbons items are almost identical in wording, the soda and bus items are also very similar. One from each pair appears on a pretest and the other is a corresponding interview item. These items were selected for discussion in this chapter because they allow comparisons of cognitive processes for a common form of question.

Student performance on these four items is given in Table 3. The distributions of boys' and girls' correct responses are significantly different ($\chi^2 = 7.86$, $p < .05$). The most striking difference is between the numbers of boys and girls erring on all four items. Many more girls than boys have difficulty with these items. This is reflected as well in the larger proportion of boys answering at least three items correctly.

All four of these items have the same basic structure. This structure is illustrated in Figure 1. Each of the four items begins with some original quantity, expressed in particular units of measurement (cards, ribbons, bottles, or seats). The unknown in each item is the number of units remaining after a portion is

There does not appear to be a sex-related difference in ability to recognize structure within a pair of items. Large percentages of both sexes solved both items in a pair in the same way. Girls may be slightly better than boys in recognizing semantic structure: 20 percent of the girls and 15 percent of the boys did so on both pairs of items. Boys are better than girls in discerning the common structure of all four items: 38 versus 25 percent. However, this is not a statistically significant difference in this sample.

Student Descriptions of Problem-Solving Process

What is going on in the problem-solving process itself? To determine the answer as fully as possible, we questioned the children while they were solving the interview items. Among other questions, we asked them to identify the important features of each item. Their responses can be evaluated under the error classification described earlier in this paper.

Recall the simple model of problem solving described above. The first stage is recognition of salient features of an item. In word problems, one expects the salient features to be verbal cues directing the student's attention. The second stage is selection of an appropriate strategy for solving the problem. For the problems studied here, this becomes a choice of operation or operations. In many instances, the decision is whether to use one or two successive computations. Finally, the third stage is implementation. Errors in this stage are typically those of arithmetic computation.

Several commonalities are evident in the girls' and boys' responses. One is the direct translation of certain words into operations. Rather than encode the problem according to a structure such as that of Figure 1, many students search for key words in a problem to identify the operation required. Table 5 gives the most common key words and associated operations described by students as they solved the four items. Each of these was specified by students as justification for an associated operation. Frequently, the student volunteered statements such as "My teacher says of always means multiply."

There are distinct differences in the key words noticed by each sex. In the bus/soda items, the most frequent phrases mentioned by the students as they solved the items were "one out of every six (three)", depending upon the item being solved. Roughly one-third of all students gave this phrase as justification for the arithmetic operation performed. More boys than girls did so: 41% of all boys versus 25% of all girls. Most of the students were led to the operation of division by the phrase. However, some interpreted the words to mean subtraction and thought they should subtract the value 1 from the other number in the phrase. These students reached either (3 - 1) or (6 - 1) as solutions.

A second response to the bus/soda items is to focus on the representation such as that of Figure 1 is necessary to recognize that the ribbons and soda items are also alike.

Consistency in the matched items takes the form of answering both items correctly, of reaching the quantity to be removed as the answer on both, or of ignoring the step requiring conversion to the proper unit of measurement and subtracting the fraction or proportion directly from the original unit. Roughly thirty-five percent of boys and girls were consistent in their approach to the ribbons/cards items. Fourteen of each sex demonstrated consistency. Similar results were obtained for the second pair of items. Forty-three percent of all students were consistent: 16 boys and 18 girls.

On all items, most of the students who erred consistently found the amounts to be subtracted; 60 percent of the consistent respondents reached such solutions. Neither sex is more likely to be consistent than the other. Both boys and girls perceived the matched items as having similar structure and attempted to solve the items within a pair using one cognitive strategy.

A related issue of consistency is whether students use a single strategy for each of the matched pairs. Fifteen percent of the boys were consistent within both pairs of matched items. That is, they took a consistent approach to the pair of ribbons/cards and also took a consistent approach to the pair of bus/soda. However, these approaches were not in general the same for the two pairs of items. Similarly, 20 percent of the girls were consistent within both pairs of items. Again, they did not necessarily use the same strategies for the two pairs. Only two boys and three girls used a single strategy for all four items. Both boys and one girl solved the items correctly and consonant with Figure 1. The remaining girls found the number of units to be removed in each case.

About one-third of the students were consistent on at least three of the four items. Nine of these students were correct in their responses (3 girls and 6 boys). Most of the remaining students consistently found the amount to be removed from the original quantity: five girls and nine boys. Two additional girls consistently made inappropriate subtractions. A larger proportion of boys than girls used a single strategy for at least three of the four item: 38 percent of the boys did so in contrast with 25 percent of the girls.

A surprisingly large number of students were inconsistent in their responses. Several of the seeming inconsistencies are simply differences in finding the quantity to be removed for one item and in carrying out the necessary second step for a correct solution to the second item. However, many of the inconsistencies correspond to choice of different operations. That is, a student might subtract on the ribbons item and divide on the soda item. Such responses indicate failure to recognize the common structure found in both items.

are also more likely than boys to resort to solutions based upon guessing patterns. Both of these errors reflect context-free or context-independent methods of problem solving.

Summary

From the studies described here we have new evidence of differences and similarities in boys' and girls' mathematics performance. From the first study, it is evident that boys and girls tend to make different errors on a variety of items. Their responses suggest attention to different features of the items. The second study provides details of how students differ in their understanding of word problems and in their use of cues in the text.

One important finding in the second study is that girls and boys appear equally likely to discern (or fail to discern) the underlying structure of the problem. The tests of consistency in response over the four items discussed here indicate that boys and girls are similar in their tendency to answer similar problems with repeated use of one strategy. On the other hand, this is primarily a negative finding, since the majority of both sexes are inconsistent. Both sexes would probably benefit from instruction on the underlying structure of word problems.

Some general results are shared by the two studies discussed here. Girls make errors that lead to illogical responses more often than boys do. This is evident from the word problem (map item) on the Survey as well as from the cards/ribbons pair of items in the interview study. Just as 250 miles is an illogical response to Item 1 of Table 1, responses of 11 2/3 baseball cards or 14 4/5 hair ribbons to Items 1 and 2 of Table 2 are also not reasonable. The operations and strategies used by girls are not wholly inappropriate in either instance. However, one expects a good problem solver to recognize solutions that are not feasible and to modify strategies accordingly.

There is also evidence in both studies that students resort to guessing patterns rather than employ strategies of problem solving. Both boys and girls make responses that are based upon the relationship or relative size of the numbers in the problem rather than the content of the item. Both studies suggest that girls are more likely to rely upon this strategy than are boys.

The present discussion of error analysis (or response analysis) shows that individual differences in cognitive strategies can be identified by examining students' responses to mathematics items. Either technique presented here can be modified for use in the classroom. On the one hand, it is relatively easy to categorize responses to a set of selected problems that have been answered by students in the usual paper-and-pencil format. Similarly, it is possible to have children explain in an oral question-and-answer session why particular operations are or are not appropriate. Results of such analyses may direct remediation or may lead to alternate forms of value 3 or 6 and to ignore the words "one out of every". Students using this approach then subtracted 24 − 3 = 21 for the soda item or 48 − 6 = 42 for the bus item. Justification for subtraction was a transformation of the problem to "Take away the number of empty seats (bottles)", and "Take away means subtract". Only girls made this error. Ten percent of the girls responding to the items made this interpretation.

The most noticeable response made by boys and girls to this pair of items was the operation of division (either 24/3 or 48/6) coupled with the inability of the student to explain why the operation was carried out. "I don't know—just instinct" and "It's the only number that goes into 24 (or 48)" and "It just seems right" are some of the responses given by the students. Twenty-two percent of all students divided but had no idea why. Most of their responses suggested that they did not understand the problem and simply combined the available numbers (either 3 and 24 or 6 and 48) in the most feasible way. A common response was "I don't really know how to solve it. Maybe 3 goes into 24?" For these items, division seemed most appropriate to them. This is another example of a guessing pattern. Many more girls than boys made responses of this type. While 5 of 39 boys (or 13 percent) could not explain their work, 12 of 40 girls (or 30 percent) could not do so. This difference is marginally significant ($.05 < p < .10$).

A somewhat different pattern emerges for the second pair of items. Forty-one percent of all students used key words for the cards/ribbons items: 20 girls and 14 boys. Thus, fifty percent of the girls attended to key words or phrases as directions for arithmetic operations. Most of the students focused on the words "gave away" or "have left", and these phrases led predominantly to the operation of subtraction. In these instances, the students subtracted the fractions directly from the original quantities given in the items. Only five students were misdirected by these words and interpreted them to mean that division was required. Of the students focusing on these phrases, 11 of 39 boys did os (28 percent) and 17 of 40 girls did so (43 percent). thus girls are more likely to attend erroneously to these phrases.

A second key word in these items is the word 'of'. A common interpretation of this is that "of means multiply". Several students volunteered the statement that "of is the same as times." This interpretation makes sense in fraction problems (e.g., give 1/3 of 12 away) but is incorrect in items using only whole numbers (e.g., give 3 of 12 away). Boys and girls were equally likely to attend to this cue.

On the basis of these four items, there are differences in the problem-solving strategies of boys and girls. Attention to key words is a major factor. Girls appear more likely than boys to combine partial recognition of cues with failure to identify appropriate units of measure. This is evident in the greater tendency of girls to subtract without first determining the number of units (cards, ribbons, bottles, or seats) to be removed. Girls instruction. In either case, instruction is better tailored to fit individual needs.

Reference Notes

1. Marshall, S. P. Sex Differences in Mathematics Errors: An Analysis of Distractor Choices. Manuscript submitted for publication, 1982.

References

California Assessment Program. Student achievement in California schools: 1980-81 Annual Report. Sacramento, California: State Department of Education, 1981.

Clements, M. A. Analyzing children's errors on written mathematical tasks. Educational studies in mathematics, 1980, 11, 1-21.

Newman, M. A. An analysis of sixth-grade pupils' errors on written mathematical tasks. In M. A. Clements & J. Foyster, (Eds.) Research in mathematics education in Australia. Melbourne, Australia: 1977.

Radatz, H. Error analysis in mathematics education. Journal for research in mathematics education, 1979, 10, 163-172.

Watson, I. Investigating errors of beginning mathematicians. Educational studies in mathematics, 1980, 11, 319-329.

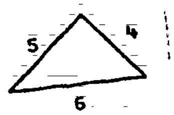
Table 1

Sample Items from the
Survey of Basic Skills: Grade 6

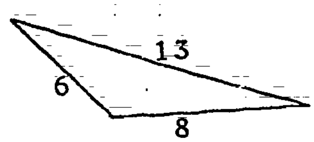
(1) If a distance of 350 miles is represented by a segment of 14 inches on a map, then on the map 1 inch represents:
[a] 4 miles
* [b] 25 miles
[c] 40 miles
[d] 250 miles

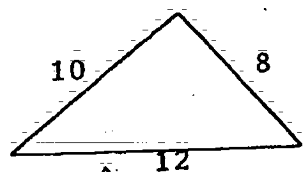
Percentages Selecting Each Distractor:

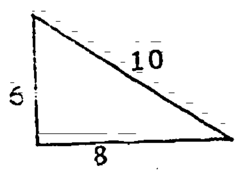
	[a]	[c]	[d]	Total Incorrect
boys	31	37	32	13,689
girls	32	30	38	14,674

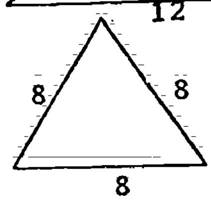


(2) The triangle above is similar to which of the following?

[a]


* [c]


[b]


[d]


Percentages Selecting Each Distractor:

	[a]	[b]	[d]	Total Incorrect
boys	5	45	50	7,680
girls	4	53	43	8,078

* correct solution

These items are reproduced here with the permission of the California Assessment Program, California Department of Education.

Table 2
Word Problems Solved by Sixth Grade Students in Individual Interviews

Pair 1:

(1) John has 12 baseball cards. He gives 1/3 of them to Jim. How many does John have left?

(2) Mary has 15 hair ribbons. She gives 1/5 of them to Alice. How many ribbons does Mary have left?

Pair 2:

(1) A bus has seats for 48 passengers. If one out of every six seats is empty, how many passengers are on the bus?

(2) A case of soda contains 24 bottles. If one out of every three bottles is empty, how many full bottles of soda are in the case?

Table 3

Frequency Distributions of Boys' and Girls' Correct Responses to the Items of Table 2

	Number of Items Answered Correctly				
Sex	0	1	2	3	4
Boys	9	12	9	7	2
Girls	15	13	9	2	1

Table 4

Frequency of Correct and Incorrect Responses
to Each Item of Table 2

Item:		Correct	First Step	Student Response Subtraction	No Answer	Other
CARDS	Boys	16	15	5	0	3
	Girls	16	9	10	4	1
RIBBONS	Boys	21	8	7	2	1
	Girls	7	10	8	12	3
BUS	Boys	11	20	4	3	1
	Girls	9	20	5	3	3
SODA	Boys	11	17	4	3	4
	Girls	9	20	4	4	3

Table 5

Key Words and Their Associated Operations

KEY WORD	ASSOCIATED OPERATION	# OF STUDENTS USING: BOYS	GIRLS
(1) OF	multiplication	3	3
(2) GAVE AWAY/TAKE AWAY	subtraction	4	8
(3) HAVE LEFT/ARE LEFT	subtraction	5	9
(4) HAVE LEFT/ARE LEFT	division	3	2
(5) GAVE AWAY	division	0	3
(6) 1 IN 6 / 1 IN 3	division	10	7
(7) 1 IN 6 / 1 IN 3	subtraction	6	3

Figure 1

A general model of the underlying structure of the items given in Table 2.