In general, studies investigating sex differences in mathematics performance have empirically analyzed students' responses from standardized tests. This paper reports a study to extend the investigation of sex differences related to mathematics performance by examining cognitive and motivational factors as well as standardized test scores. Boys (n=16) and girls (n=16) from third-, fifth-, and eighth-grade urban classes were examined on three criteria: (1) the cognitive factors of strategies used in solving computational problems, use of manipulatives to solve and represent a solution, and error types made during the solution process; (2) the motivational factors of confidence, willingness to participate in mathematics activities, and causal attributions; and (3) mathematics performance as measured by the Metropolitan Achievement Test. Student interviews were conducted to examine motivational and cognitive factors after three tasks: a grade appropriate computation, a motivational interview, and a projective interview. Analysis of variance was performed on the interview responses. Results indicated that motivational factors were influential in explaining the emergence of sex differences in mathematics performance. While no differences were found in any of the computation measures, sex differences were found in the motivational measures. Males were more confident with respect to mathematics, and eighth-grade males were more willing to attempt mathematics. Appendixes include drawings used for the projection interview and the questions included in the confidence scale and attribution measures. (Contains 45 references.) (MDH)
SEX DIFFERENCES IN COGNITIVE AND MOTIVATIONAL FACTORS UNDERLYING CHILDREN'S AND ADOLESCENTS' MATHEMATICS PERFORMANCE

Gavrielle Levine
C.W. Post Campus, Long Island University
Brookville, New York 11548

Presented at the Annual Meeting of the American Educational Research Association
San Francisco, 1992

In recent years, a considerable amount of research has used standardized test results to investigate sex differences in mathematics performance. The generally accepted findings are that there is no difference between the overall mathematics achievement scores of boys and girls in elementary school, and that sex differences in overall mathematics performance favoring males seem to emerge by the beginning of secondary school (Friedman, 1989; Leder, 1990). The appearance of sex differences in mathematics performance is surprising since they emerge from an instructional context putatively designed to provide equal opportunities for boys and girls. Why do overall sex differences in performance appear in the later grades?
Current models of academic performance suggest that a possible explanation for differential performance lies in relevant cognitive and motivational factors (Helmke, 1989; Pintrich, 1990; Zimmerman, 1990). For example, one conceptualization (Corno & Snow, 1986) combines motivational factors (the student's task goal) with cognitive factors (the strategy used to accomplish the task). Models of self-regulated learning (Corno & Mandinach, 1983; Corno & Rohrkemper, 1985; Weinstein & Mayer, 1986) describe academic performance as involving three components: cognitive features, motivational factors, and meta-cognitive or self-regulatory control processes. Thus, in major models recently proposed, interactions among cognitive and non-cognitive factors are hypothesized to predict differences in academic performance. The present study examines sex differences in mathematics performance through the filter of cognitive and motivational factors.

Consider first cognitive factors. Most studies of sex differences in mathematics performance have focused on response accuracy (Fennema & Carpenter, 1981; Friedman, 1989; Halpern, 1986; Leder, 1990) and have ignored underlying cognitive factors like strategies used to solve computations (Lesh, Behr, & Post, 1987), use of materials to solve a computation or represent a solution (Kaplan, Yamamoto, & Ginsburg, 1989), and types of errors made during the solution process (Ashlock, 1990; Brown & Burton, 1978; Marshall & Smith, 1987; Maurer, 1987). The present study examines response accuracy as well as these additional
cognitive factors related to mathematics performance to explain sex differences in mathematics performance.

Several motivational factors associated with sex differences in mathematics performance in older students may have origins in younger students. In particular, confidence in relation to engaging in mathematics tasks (Dweck & Goetz, 1978; Eccles, 1983; Fennema & Sherman, 1978) and causal attributional reasoning (Fennema, 1985; Meyer & Koehler, 1990) have been related to sex differences in mathematics performance in high school and college students. Furthermore, confidence and causal attributions have been considered the two motivational factors "most important" for mathematics performance (Becker, 1991). In addition to these motivational factors, sex differences in participation in mathematics-related activities, starting with course selection and developing into differential opportunities for future education and employment, have been described and studied starting with high school students (Chipman, Brush, & Wilson, 1985; Eccles, 1983; Eccles [Parsons], Adler, Futterman, Goff, Kaczala, Meece & Midgley, 1985; Fennema & Sherman, 1977; Leder, 1990; Sells, 1973). This study uses a measure of students' willingness to participate in mathematics activities prior to high school as a third indicator of motivation for mathematics. These three motivational factors, confidence in one's ability to do mathematics, causal attributions related to mathematics performance, and willingness to attempt mathematics problems, have been included in the present study.
In summary, the literature reveals a gap between current theoretical models of academic performance applied to the study of sex differences, and research examining these differences in mathematics performance. Empirical investigations have focused on accuracy scores from standardized tests to describe performance differences. Current theoretical models propose that academic performance is best described by examination of related cognitive and motivational factors. The present study extends the investigation of sex differences in mathematics performance by examining in detail several cognitive and motivational factors related to sex differences in mathematics performance.

**METHOD**

This study examined differences between boys and girls in several cognitive and motivational factors related to solving computation problems at three grade levels. The design features were selected with the goal of observing as natural an interaction as possible from as representative a sample as possible while still maintaining comparability across grades.

**Sample**

Participating in the study were 16 female and 16 male students in each of the third, fifth, and eighth grades of urban
public schools. Their standardized mathematics test scores from the Metropolitan Achievement Test (MAT) were analyzed by grade for sex differences using t-tests, with probabilities corrected for the number of comparisons using Dunn's multiple comparison procedure (Dunn, 1961). Significant sex differences were found only in the eighth grade (p < .01). This result is consistent with those reported for similar grades as summarized by Halpern (1986).

Procedure

Each student was interviewed individually as he or she completed three tasks: (a) a grade-appropriate computation, (b) a motivational interview, and (c) a projective interview. Every participant responded to each of the three tasks. All questions were presented orally.

Grade-Appropriate Computation Task

The structure of this task was the presentation of a grade-appropriate computation example which the student was encouraged to solve in as many ways as possible while thinking aloud. After hearing the problem but before solving it, each student was asked whether or not he or she thought he or she could solve it. The student was then given the opportunity to solve the problem. A brief clinical interview, designed to clarify the solution processes (Ginsburg, Kossan, Schwartz, &
Swanson, 1983), was administered to each student following each different solution. A second grade-appropriate computation example, again requesting multiple solution approaches and including brief interviews, was then administered, to establish test-retest reliability.

The computation task for third-grade students was whole number addition (e.g., $38 + 14 =$), for fifth-grade students was whole number multiplication (e.g., $15 \times 3 =$), and for eighth-grade students was multiplication of a mixed number with a whole number (e.g., $2 \frac{1}{2} \times 4 =$). Materials (e.g., counting chips, Cuisiniere Rods, play money) were available throughout the task, and each student was encouraged to use them to solve the problem.

**Motivational Interview**

The Motivational Interview investigated aspects of students' motivation for engaging in specific mathematics tasks, and for their overall mathematics performance. It consisted of three types of orally presented questions: open-ended questions; multiple-choice questions based on the Mathematics Attribution Scales (Fennema, Wolleat, & Pedro, 1979); and Likert-scale-type questions adapted from the Fennema-Sherman Mathematics Attitudes Scales (Fennema & Sherman, 1976) tapping motivation for mathematics, which were presented in a format similar to that used by Nicholls (1990) and by Fennema and Peterson (1984). Since these scales were originally developed for use with high school students, questions were selected and language was adapted for
the lower grade levels of the students included in this study.

This task was constructed so that motivational questions were asked immediately following engagement in a set of mathematics tasks. Questions were worded to refer to both overall and specific motivation in the mathematics tasks students completed during earlier parts of the study. As mentioned previously, focal variables were confidence and causal attribution reasoning in relation to mathematics learning and performance.

**Projective Interview**

The Projective Interview consisted of a structured set of focused questions regarding motivation for mathematics, using two stimulus cards from a preliminary version of the *Mueller Ginsburg Projective Test* (Mueller & Ginsburg, in press). The *Mueller Ginsburg Projective Test* is an open-ended, projective instrument used to assess attitudes and feelings in a school/learning environment. In the present study, this experimental instrument examined motivational factors related to mathematics performance. Appendix A presents the drawings used in this study. Questions concerned the same motivational issues addressed in the Motivational Interview Task described above. Responses to these projective questions may be interpreted as reflecting relatively more unconscious or preconscious aspects of motivation.

The two stimulus cards used in the present study presented
two drawings, one featuring a student studying in a bedroom, and the other in a mathematics classroom. Male students viewed drawings with a male main character, and female students viewed drawings with a female main character. Students were asked to respond to questions from the point of view of the character in the drawing.

RESULTS

For data analysis purposes, the cognitive and motivational measures obtained were reduced to provide a meaningful summary of the data. Student behavior and responses to questions posed during the cognitive, motivational and projective tasks were collapsed into seven descriptive variables: solution accuracy, use of materials to solve computations, use of strategies, types of errors, confidence in one's ability to do mathematics, causal attributional reasoning in relation to mathematics, and willingness to attempt mathematics.

It must be recalled that grade-level differences are confounded by task differences for the computation tasks. Therefore, each grade is examined separately for sex differences on computation-based measures.
Computation Variables

A total of four computation variables were developed from the computation task: solution accuracy, use of materials to solve computations, use of strategies to solve computations, and types of errors. These variables were chosen to present a picture of the subtle aspects of the solution process. Most studies reported in the literature present results comparing accuracy scores derived from standardized test score data. In this study, additional cognitive variables related to the computation process were examined.

Reliability Data for the Computation Variables

Students' behaviors during and responses to the first two trials of each problem during the cognitive task were recorded on videotape during the interview, and were later coded for each of four variables by two independent raters. An estimation of inter-rater reliability on the coding of computation tasks from videotapes was derived from a randomly-selected sample of 48 of the 96 students interviewed (50% of the sample) along four cognitive variables (solution accuracy, use of material, use of strategy, and types of errors). It was calculated as percent agreement [total agreements/ (total agreements + total disagreements) X 100] between two judges. Average overall agreement for the questions in the first two trials of Problem One and Problem Two 95%.
Solution Accuracy

Response accuracy for each trial of both computation problems was coded as correct or not correct depending on the final answer attained. Trials that were not complete were coded as not correct.

Overall, students were more accurate on the first trials of each problem than on the second trials. Approximately 84% of the students in this study answered the first trial of the computation problems correctly (Problem One: 84%; Problem Two: 85%). When solving the second trial of each problem, success dropped to approximately 57% (Problem One: 58.33%; Problem Two: 56.25%). At first glance this is surprising. However, the second trial was the unconventional request that the student try the problem again another way. In light of the difference in task demands, the drop in accuracy on the second trial is less surprising.

Performance on the first trial of the two problem was similar over the sample, as was performance on the second trials. Therefore, for analysis purposes, performance on the first trials of the two problems was examined together, as was performance on the second trials. Chi-square tests or Fisher's exact tests were used to compare the frequency of male and female responses at each grade for correct responses to Trial One and Trial Two. No significant sex differences were found at any grade for either Trial One or Trial Two. Male and female accuracy at all three grades was similar on this computation task.
Use of Materials to Solve Computations

For purposes of analysis, the materials were grouped into one of five categories: mental, written, counting, base ten, and combination (using more than one of the materials to solve the problem during a given trial). The pattern of the use of materials to solve the problems was similar for males and females. In all three grades, no sex differences were found in the use of materials.

Use of Strategies to Solve Computations

For data analysis purposes, the strategies students used were grouped into one of four categories: counting-based, algorithmic, other, no solution. No significant sex differences were found in the use of strategies to solve problems at any of the grade levels.

Types of Errors

For data analysis purposes, errors observed during the computation process were categorized into one of three categories: computation, representation, and interpretation. These categories were developed to reflect errors in skill as well as concept. No significant sex differences were found in the types of errors made while solving Trial One or Trial Two of both computations.
Motivation Variables

A total of three motivation variables were developed from the computation task, the motivation task, and the projective task: confidence in one's ability to do mathematics, willingness to attempt mathematics, and causal attributions for success and failure related to doing mathematics.

Confidence in One's Ability to do Mathematics

Confidence related to mathematics is usually studied independently of a student's participation in mathematics activities. In the present study, a confidence variable was constructed from responses to questions obtained during the experimental computation task, so confidence is in part specific to this task. Because other measures were also obtained, from the follow-up motivation interview and the projective test, this variable can be extended beyond the tasks used here, to represent overall confidence in mathematics as well. Appendix B presents the questions included in the confidence variable.

Internal consistency of the combined 20-question scale, computed using coefficient alpha, was found to be .61. A confidence score was computed for each student. The value 0 was assigned to all "no" and missing responses, .5 to all "maybe" responses, and 1 to all "yes" responses. The confidence score was computed as the sum of responses for each student, and could range from 0 to 20. Responses to open-ended questions included in
the confidence variable were assigned one of the three values by two independent raters. An estimation of inter-rater agreement on this coding was derived from a randomly-selected sample of 48 of the 96 students who participated in this study (50% of the sample). It was calculated as percent agreement \[ \frac{\text{total agreements}}{\text{total agreements} + \text{total disagreements}} \times 100 \] between two judges. Average inter-rater agreement for open-ended questions was 93%. Responses to Likert type questions were recoded so that all "yes" responses (Y, y) were coded as "yes" (1) and all "no" responses (N, n) were coded as "no." In this way, scores from the two types of responses could be combined in one confidence variable.

Analysis of variance was used to examine sex and grade mean differences in confidence scores related to mathematics tasks. A higher score reflects more confidence. Results are presented in Table 1. There was no main effect for grade, and a significant main effect for sex. Figure 1, which presents confidence score means by sex and grade, shows the significantly higher mean confidence scores for males than for females. For students in this study, males were more confident than were females about their math ability.

Willingness to Attempt Mathematics

A measure of willingness to attempt mathematics was operationally defined as students' willingness to attempt the second trial of a computation problem that they had just solved.
### TABLE 1

ANALYSIS OF VARIANCE OF THE CONFIDENCE SCALE
BY GRADE AND SEX

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>DF</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRADE</td>
<td>.57</td>
<td>2</td>
<td>.04</td>
</tr>
<tr>
<td>SEX</td>
<td>51.77</td>
<td>1</td>
<td>7.85**</td>
</tr>
<tr>
<td>SEX-GRADE INTERACTION</td>
<td>15.52</td>
<td>2</td>
<td>1.18</td>
</tr>
<tr>
<td>ERROR</td>
<td>593.52</td>
<td>90</td>
<td></td>
</tr>
</tbody>
</table>

**p<.01
FIGURE I
MEAN CONFIDENCE SCORES

GRADE 3
GRADE 5
GRADE 8

FEMALE
MALE
FEMALE
MALE
FEMALE
MALE

18 16 14 12 10 8 6 4 2 0
MEAN SCORES
This measure of willingness to attempt mathematics combines several motivational factors that cannot be separated. It includes willingness to put in more effort on the task, willingness to make more choices (of material and/or strategy to use), and willingness to risk being less correct.

Two independent raters coded students' willingness to attempt mathematics from videotapes of the interview that they viewed separately. An estimation of inter-rater agreement was derived from a randomly selected sample of 48 of the 96 students who participated in this study (50% of the sample). It was calculated as percent agreement \[\frac{\text{total agreements}}{\text{total agreements} + \text{total disagreements}} \times 100\] between two judges. Overall agreement for the first trials of Problems One and Two were 96% and 98% respectively, and for the second trials were 98% and 100% respectively.

All students were willing to attempt the first trial of each of the computation problems. However, not all students were willing to attempt a second trial of each problem. Table 2 presents the distribution of students who were willing to attempt a second trial. Fisher's exact tests were used to examine differences between males and females in frequency of willingness to make a second attempt on each of the computation problems. No significant sex differences were found for third and fifth grade students on the frequency of willingness to make a second attempt on either Problem One or Problem Two. The number of eighth grade females unwilling to make a second attempt on Problem One, even
TABLE 2

DISTRIBUTION OF STUDENTS WILLING TO ATTEMPT A SECOND TRIAL ON COMPUTATION PROBLEMS

<table>
<thead>
<tr>
<th></th>
<th>FEMALE</th>
<th></th>
<th>MALE</th>
<th></th>
<th>FISHER'S EXACT PROBABILITY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WILLING</td>
<td>REFUSED</td>
<td>WILLING</td>
<td>REFUSED</td>
<td></td>
</tr>
<tr>
<td>PROBLEM 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 3</td>
<td>16</td>
<td>0</td>
<td>14</td>
<td>2</td>
<td>.48</td>
</tr>
<tr>
<td>(N=32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5</td>
<td>16</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>a</td>
</tr>
<tr>
<td>(N=32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8</td>
<td>10</td>
<td>6</td>
<td>16</td>
<td>0</td>
<td>.02*</td>
</tr>
<tr>
<td>(N=32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROBLEM 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 3</td>
<td>13</td>
<td>3</td>
<td>14</td>
<td>2</td>
<td>1.00</td>
</tr>
<tr>
<td>(N=32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5</td>
<td>15</td>
<td>1</td>
<td>15</td>
<td>1</td>
<td>1.00</td>
</tr>
<tr>
<td>(N=32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8</td>
<td>12</td>
<td>4</td>
<td>16</td>
<td>0</td>
<td>.10</td>
</tr>
<tr>
<td>(N=32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05
*a Unable to compute due to lack of variability.

with encouragement by the interviewer, was significantly lower
than the number of eighth grade males unwilling to make a second attempt on Problem One (p < .05). Although for eighth grade students there were no significant sex differences in frequency of willingness to attempt a second trial of Problem Two, the pattern of willingness to attempt a second trial was similar to that of Problem One; namely, boys were more willing than girls to attempt a second trial.

Causal Attributions

The causal attribution variable was comprised of responses to interview questions coded into one of five categories using Weiner's (1974) model as modified by Nicholls (1975). The category reflecting unstable external factors was divided into two parts called "luck" and "teacher qualities" based on research showing differential patterns for male attributional responses for failure (Nicholls, 1975). Thus, the five component attribution categories used in this study were: effort, ability, task, teacher, and luck. Appendix C presents questions included in the causal attribution variable.

Responses to questions examining success attributions were examined separately from those measuring failure attributions, since the literature (Fennema, 1985; Nicholls, 1975) suggests that there will be a differential pattern of sex differences in those attributional responses. The causal attribution variable for success consisted of 14 questions and for failure consisted of six questions.
Responses to open-ended questions included in the causal attribution variable were coded by two independent raters into one of the five attribution categories. An estimation of inter-rater agreement on this coding was derived from a randomly-selected sample of 48 of the 96 students who participated in this study (50% of the sample). It was calculated as percent agreement \[
\frac{\text{total agreements}}{\text{total agreements + total disagreements}} \times 100
\]
between two judges. For the causal attribution variable, the average inter-rater agreement for the open-ended attributional responses was 76%.

Two types of analyses were used to examine grade and sex differences in patterns of attributions. Frequencies of responses in attribution categories were used to rank order the categories for each individual for success and failure separately. The Friedman Test, a two-way nonparametric analysis of variance, was used to examine the patterns of the ranks. Within attribution categories, individuals' responses were ranked by grade and sex. Differences in ranks for sex and grade within attribution categories were examined using the Kruskal-Wallis Test, a one-way nonparametric analysis of variance. Probabilities were corrected for the number of comparisons using Dunn's multiple comparison procedure (Dunn, 1961).

**Success Attributions.** Across the five attribution categories, results of the Friedman Tests showed significant grade differences for both males and females. For third grade
males ($X^2(15) = 43.26, p<.01$) and females ($X^2(15) = 26.68, p<.01$), and fifth grade males ($X^2(15) = 43.38, p<.01$) and females ($X^2(15) = 41.88, p<.01$), effort was the most highly ranked success attribution category. For eighth grade males ($X^2(15) = 39.73, p<.01$) and females ($X^2(15) = 30.01, p<.01$), ability was the most highly ranked success attribution category.

Within each of the attribution categories, results of the Kruskal-Wallis tests showed significant differences among the grades for the effort ($X^2(95) = 15.18, p<.01$), ability ($X^2(95) = 18.32, p<.01$), and teacher ($X^2(95) = 10.13, p<.01$) attribution categories. While eighth grade students cited effort less frequently than did either third or fifth grade students, they cited ability more frequently than did either of the other grades.

While no significant sex differences for ranks within an attribution category were found, there were several significant sex-grade interactions. Table 3 shows average ranks of success attribution responses for male and females students for each grade. The results of Kruskal-Wallis tests indicate that for females there were no significant grade differences in any of the success attribution categories. For males, however, effort was cited less frequently ($X^2(47) = 19.09, p<.01$) and ability was cited more frequently ($X^2(47) = 16.68, p<.01$) by eighth grade students when compared to third and fifth grade students. Males, therefore, may have contributed more heavily to the overall
TABLE 3

KRUSKAL-WALLIS TEST:
MEAN RANKS BY SEX AND GRADE FOR
SUCCESS ATTRIBUTION CATEGORIES

<table>
<thead>
<tr>
<th>Attribution Category</th>
<th>Grade 3</th>
<th>Grade 5</th>
<th>Grade 8</th>
<th>X²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Females</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effort</td>
<td>22.28</td>
<td>29.38</td>
<td>21.84</td>
<td>2.98</td>
</tr>
<tr>
<td>Ability</td>
<td>16.63</td>
<td>25.97</td>
<td>30.91</td>
<td>8.80</td>
</tr>
<tr>
<td>Task</td>
<td>26.47</td>
<td>21.16</td>
<td>25.88</td>
<td>1.50</td>
</tr>
<tr>
<td>Teacher</td>
<td>19.03</td>
<td>27.22</td>
<td>27.25</td>
<td>3.90</td>
</tr>
<tr>
<td>Luck</td>
<td>28.38</td>
<td>21.63</td>
<td>23.50</td>
<td>2.44</td>
</tr>
<tr>
<td><strong>Males</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effort</td>
<td>38.22</td>
<td>31.13</td>
<td>12.16</td>
<td>19.09**</td>
</tr>
<tr>
<td>Ability</td>
<td>23.41</td>
<td>15.13</td>
<td>34.97</td>
<td>16.63**</td>
</tr>
<tr>
<td>Task</td>
<td>22.94</td>
<td>25.63</td>
<td>24.94</td>
<td>.36</td>
</tr>
<tr>
<td>Teacher</td>
<td>17.47</td>
<td>27.59</td>
<td>28.44</td>
<td>6.43</td>
</tr>
<tr>
<td>Luck</td>
<td>24.41</td>
<td>21.06</td>
<td>28.03</td>
<td>2.63</td>
</tr>
</tbody>
</table>

** p < .01

Note. Dunn's correction (based on the number of tests) was used to establish significance levels. Grade differences in success attribution patterns identified.
above.

**Failure Attributions.** As shown on Table 4, Friedman Tests found significant differences for the pattern of failure attributions for both males and females in all three grades. For all groups, lack of effort was always the most frequently chosen category to describe reasons for failure. Differences among grade, sex, and sex and grade were examined using the Kruskal-Wallis tests. No significant differences were found.

**Correlations Among Cognitive and Motivation Variables and Prediction of Standardized Mathematics Achievement Test Scores**

As discussed above, recent models of academic performance suggest relationships between cognitive and motivational factors in predicting mathematics performance (Pintrich, 1990; Zimmerman, 1990). To explore these relationships in the present data, accuracy on the computation problems and total scores on the mathematics subscale of the Metropolitan Achievement Test (MAT) were correlated with all of the motivational variables. Based on the bivariate correlation patterns, variables were selected for inclusion in regression analyses predicting total scores on the mathematics subscale of the Metropolitan Achievement Test.

To compute the correlation matrix and predict mathematics achievement, several variables were constructed. An accuracy score was constructed for each student that was the sum of correct (1) and incorrect (0) answers on the first two trials of
### TABLE 4

**FRIEDMAN TEST:**
**MEAN RANKS FOR FAILURE ATTRIBUTION CATEGORIES**
**BY SEX AND GRADE**

<table>
<thead>
<tr>
<th>Attribution Category</th>
<th>Effort</th>
<th>Ability</th>
<th>Task</th>
<th>Teacher</th>
<th>Luck</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Females</td>
<td>4.61</td>
<td>2.75</td>
<td>2.45</td>
<td>2.78</td>
<td>2.41</td>
<td>64.79**</td>
</tr>
<tr>
<td>Grade 3</td>
<td>4.72</td>
<td>2.38</td>
<td>2.34</td>
<td>2.75</td>
<td>2.81</td>
<td>24.79**</td>
</tr>
<tr>
<td>Grade 5</td>
<td>4.88</td>
<td>2.66</td>
<td>2.22</td>
<td>2.72</td>
<td>2.53</td>
<td>29.08**</td>
</tr>
<tr>
<td>Grade 8</td>
<td>4.25</td>
<td>3.22</td>
<td>2.78</td>
<td>2.88</td>
<td>1.88</td>
<td>18.81**</td>
</tr>
<tr>
<td>Males</td>
<td>4.61</td>
<td>2.80</td>
<td>2.63</td>
<td>2.33</td>
<td>2.63</td>
<td>64.74**</td>
</tr>
<tr>
<td>Grade 3</td>
<td>4.47</td>
<td>2.66</td>
<td>2.76</td>
<td>2.34</td>
<td>2.75</td>
<td>18.03**</td>
</tr>
<tr>
<td>Grade 5</td>
<td>4.66</td>
<td>2.78</td>
<td>2.84</td>
<td>1.94</td>
<td>2.78</td>
<td>25.55**</td>
</tr>
<tr>
<td>Grade 8</td>
<td>4.72</td>
<td>2.97</td>
<td>2.25</td>
<td>2.72</td>
<td>2.34</td>
<td>25.78**</td>
</tr>
<tr>
<td>Overall</td>
<td>4.61</td>
<td>2.78</td>
<td>2.54</td>
<td>2.56</td>
<td>2.52</td>
<td>126.82**</td>
</tr>
</tbody>
</table>

**p < .01**

**Note.** Dunn's correction (based on the number of tests) was used to establish significance level.
the two computation problems. A higher score reflected more correct responses. Accuracy was the only cognitive factor included in correlation analyses since the other cognitive factors were categorical variables.

A high confidence score was defined as one that was greater than or equal to the median confidence score (16.5) for the entire sample. A low confidence score was defined as one that was less than the median confidence score for the entire sample. Willingness to attempt a second trial of the first computation problem was also included in the correlation matrix. For each of the success and failure attribution categories, percent of total responses was computed. Percent of total responses rather than frequencies were used to create a relative set of responses for each student since some students did not answer all questions.

Since cognitive tasks differed for the three grades, all variables were normalized prior to inclusion in the correlation matrices and regression analyses. In this way, the entire sample could be examined together.

**Relationships Between Cognitive and Motivation Variables.** Table 5 shows correlations between computation accuracy, total scores on the mathematics subscale of the Metropolitan Achievement Test, and the various motivation variables. High or low confidence score was positively correlated with mathematics MAT score, and percent of attributions for failure to luck was negatively correlated.
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.MAT</td>
<td>1.00</td>
<td>.13</td>
<td>.09</td>
<td>.24</td>
<td>-.07</td>
<td>-.06</td>
<td>-.05</td>
<td>-.05</td>
<td>.18</td>
<td>-.17</td>
<td>.09</td>
<td>.15</td>
<td>-.00</td>
<td>.03</td>
<td>-.38</td>
</tr>
<tr>
<td>2.ACU</td>
<td>1.00</td>
<td>.03</td>
<td>.04</td>
<td>.42</td>
<td>-.11</td>
<td>.17</td>
<td>.17</td>
<td>-.10</td>
<td>-.04</td>
<td>.03</td>
<td>.01</td>
<td>-.17</td>
<td>-.00</td>
<td>.12</td>
<td></td>
</tr>
<tr>
<td>3.S X</td>
<td>1.00</td>
<td>-.25</td>
<td>.15</td>
<td>.05</td>
<td>.13</td>
<td>.13</td>
<td>-.03</td>
<td>-.17</td>
<td>.05</td>
<td>-.06</td>
<td>.12</td>
<td>-.16</td>
<td>.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.CONF</td>
<td>1.00</td>
<td>.09</td>
<td>.16</td>
<td>-.09</td>
<td>-.09</td>
<td>.08</td>
<td>-.16</td>
<td>-.01</td>
<td>-.07</td>
<td>.03</td>
<td>.03</td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.WILL</td>
<td>1.00</td>
<td>-.09</td>
<td>.14</td>
<td>.14</td>
<td>-.16</td>
<td>-.02</td>
<td>-.19</td>
<td>.21</td>
<td>-.20</td>
<td>-.04</td>
<td>.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.ABILS</td>
<td>1.00</td>
<td>-.71</td>
<td>-.58</td>
<td>-.10</td>
<td>-.07</td>
<td>.22</td>
<td>-.10</td>
<td>-.02</td>
<td>-.01</td>
<td>-.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.EFFS</td>
<td>1.00</td>
<td>-.71</td>
<td>-.25</td>
<td>-.20</td>
<td>-.18</td>
<td>.22</td>
<td>.02</td>
<td>-.16</td>
<td>-.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.TASKS</td>
<td>1.00</td>
<td>-.26</td>
<td>-.03</td>
<td>.04</td>
<td>-.15</td>
<td>.12</td>
<td>.03</td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.TEACHS</td>
<td>1.00</td>
<td>-.32</td>
<td>.06</td>
<td>-.01</td>
<td>-.06</td>
<td>.20</td>
<td>-.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.LUCKS</td>
<td>1.00</td>
<td>-.11</td>
<td>-.09</td>
<td>-.07</td>
<td>.05</td>
<td>.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.ABILF</td>
<td>1.00</td>
<td>-.58</td>
<td>-.02</td>
<td>-.06</td>
<td>-.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.EFFF</td>
<td>1.00</td>
<td>-.38</td>
<td>-.35</td>
<td>-.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.TASKF</td>
<td>1.00</td>
<td>-.11</td>
<td>-.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.TEACHF</td>
<td>1.00</td>
<td>-.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.LUCKF</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* A correlation of .21 or above, p<.05; .26 or above, p<.01.
Prediction of Standardized Mathematics Achievement Test Scores. Generalized regression analysis was used to estimate the predictive value of computational and motivational variables for total scores on the Metropolitan Achievement Test (MAT) of mathematics. Included in this analysis, in order of entry, were the following variables: sex of student, accuracy on the first two trials of each of the computation problems, a variable labeled Confal which was constructed to include the difference between high/low confidence scores and percent of attributions for failure to the luck category, and grade. These two motivational variables correlated highest with MAT scores. High/low confidence scores were positively correlated with MAT scores, while percent of attributions for failure to the luck category were negatively correlated. Constructing Confal from the difference between the two variables insured that the overall variable would be positive.

Results of the regression analysis are shown on Table 6. Confal was observed to account for 21% and grade for 5% of the variability among MAT scores, once effects of sex and computation accuracy were removed.
TABLE 6

PROMINENT COGNITIVE AND MOTIVATIONAL EFFECTS FOR MAT© SCORES FROM GENERALIZED REGRESSION ANALYSIS

<table>
<thead>
<tr>
<th>Predictor</th>
<th>df</th>
<th>F</th>
<th>b/SE</th>
<th>$R^2_{\text{change}}^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>1</td>
<td>.85</td>
<td>.08/.18</td>
<td>.01</td>
</tr>
<tr>
<td>Accuracy</td>
<td>2</td>
<td>1.14</td>
<td>.23/.10</td>
<td>.02</td>
</tr>
<tr>
<td>Confal^c</td>
<td>3</td>
<td>9.11*</td>
<td>.29/.07</td>
<td>.21</td>
</tr>
<tr>
<td>Grade</td>
<td>4</td>
<td>8.65*</td>
<td>.11/.05</td>
<td>.05</td>
</tr>
</tbody>
</table>

Note. Degrees of freedom vary according to the number of terms entering the equation. The $b$ represents the unstandardized regression coefficient from the last step in the full model; the SE is the standard error associated with the $b$.

- Metropolitan Achievement Test, total mathematics scores.
- Lines in this column indicate that the numbers following are summed over a group.
- Confal is the difference between high/low confidence score and percent of responses for failure attributions to the luck category.
- Constant.
- This number represents the total cumulative $R$ for the full model.
- *$p<.01$
DISCUSSION

The results revealed the importance of motivational factors for explaining the emergence of sex differences in mathematics performance. It was notable that while no sex differences were found in any of the computation measures, sex differences were found in motivational measures.

The absence of cognitive sex differences agrees with conclusions of recent meta-analyses (Friedman, 1989; Hyde, Fennema, & Lamon, 1990) that find a trend toward decreasing appearance and size of sex differences in mathematics performance. However, these cognitive results do not explain the emergence of sex differences in adolescent and adult mathematics performance. Current theory (Helmke, 1989; Pintrich, 1990; and Zimmerman, 1990) suggests that the independent and combined effects of several cognitive and motivational factors that characterize young children may contribute to the later emergence of sex differences in mathematics performance. In particular, these factors may contribute to the emergence of sex differences in mathematics test scores favoring males that begin to appear in secondary school.

In this study, sex differences were found in both student confidence with respect to mathematics performance and student willingness to perform mathematics. At all grade levels, male students were more confident than female students. When presented with an unfamiliar task, eighth grade males were more willing to
Confidence in Relation to Mathematics

Sex differences in confidence in relation to mathematics performance have been reported for secondary students and adults (Chipman & Wilson, 1985; Fennema, 1983; Fennema & Sherman, 1977; Fennema & Sherman, 1978; Fox, Brody, & Tobin, 1980; Stage, Kreinberg, Eccles, & Becker, 1985). The present study extends the observation of sex differences of confidence to younger students as well. Use of a specific mathematics task as a referent may have enabled younger students to respond to this confidence measure.

In contrast to previous studies that measured confidence in relation to broad categories of academic achievement or general mathematics, this study also measured confidence in relation to a particular mathematics task in which students were engaged during part of and prior to this interview. Thus, the confidence scale includes a task-specific and a generalized referent for mathematics. Shavelson & Bolus (1982) indicate the importance of relating self-concept to specific academic attainment rather than to general beliefs. This direction guided the development of the confidence measure included in this study.

Important potential consequences of sex differences in confidence are reflected in differential performance expectations for future courses by these students (Stage, Kreinberg, Eccles, & Becker, 1985), as well as differential
course selection by high school and college students (Chipman & Wilson, 1985). It should not be surprising that performance expectations and course selection are related. What emerges, then, is a picture of student behavior closely linked to confidence. Identifying sex differences in confidence in elementary students will not be reflected in differential course selection at that level, but may create an expectation of failure, which leads to avoidance of mathematics-related situations later on. Since it appears that female students do tend to avoid mathematics and mathematics courses once they can effect that choice, the roots of that decision may be deeper and start earlier than previously assumed.

Willingness to Attempt Mathematics

The behavior of the eighth grade students in this study appears to be consistent with the pattern summarized by Dweck (1986), which states that bright girls tend to prefer tasks on which they can do well and on which they are certain they will succeed, while bright boys prefer tasks that present a challenge to mastery. When presented with a novel challenge, males were engaged while females refused to participate.

One possibility is that unwillingness to attempt mathematics may be similar to the "resistance" Brown & Gilligan (1990) observed in female students moving from primary to secondary school. These researchers suggest that students experience a withdrawal of relatedness as teachers' attention shifts from
student to content area. Resistance may be a self-protective response to this shift. In addition, Brown & Gilligan identify female adolescent investment in maintaining an ideal life, which includes perfection in many aspects of self-presentation at the cost of losing connection with reality.

This combination of factors may have had an influence on the female eighth grade students in the present study. They may be experiencing a disconnection from the relatedness of their elementary school teacher that is combined with a sense of the potential loss of their ideal presentation. Refusal to participate is preferable to attempting a problem and failing. By refusing, their ideal is preserved while attempting allows the possibility that their fragile ideal might be shattered and replaced by the real. Brown & Gilligan recommend supporting the efforts of those students who resist the pull of the ideal to experience their reality. For educators, this may be an important issue to consider.

Causal Attributions

The expected pattern for success attributions (Gore & Roumagoux, 1983; Wolleat, Pedro, Becker, & Fennema, 1980) suggests that males would attribute success primarily to ability, while females would attribute success primarily to effort. In the present study, both male and female students in the third and fifth grades attributed success primarily to effort, while eighth grades students of both sexes attributed success primarily to
ability. This attributional pattern is consistent with the occurrence of a developmental shift in the conception of ability identified by Nicholls (1984), but does not explain the absence of expected sex differences in eighth grade students' attributions.

One possible explanation of these results stems from students' recent successful computation experience. Attributions of effort may be related to having worked hard to solve the mathematics tasks. Attributions of ability may be related to students' sense of having known how to solve the problems.

The expected pattern for failure attributions (Gore & Roumagoux, 1983; Wolleat, Pedro, Becker, & Fennema, 1980) suggests that females would attribute failure primarily to lack of ability and task difficulty more than would males. In the present study, all students, regardless of sex, attributed failure primarily to lack of effort. As with attributions for success, attributing failure primarily to lack of effort may reflect a child's developing conception of ability as learning through effort. In this way, failure means one did not apply sufficient effort. It may be that most students in this study hold this conception of ability.

Predicting Standardized Mathematics Achievement Test Scores

The present study examined the interplay of cognitive and motivational factors to predict performance on a standardized mathematics achievement test. Neither sex of student nor accuracy
on the computation task predicted achievement scores. However, a combination of confidence and failure attributions to luck was the best predictor of achievement.

The literature suggests that high confidence and attributing failure to categories other than luck would predict achievement scores. Fennema and Sherman (1977) found that confidence had a significant positive correlation with mathematics achievement for elementary students. Fennema (1983) found that confidence had consequences for how hard high school students studied, how much they learned, and their willingness to participate in mathematics courses.

In contrast to research (Chipman & Wilson, 1985; Fennema & Sherman, 1978) that associates higher confidence with higher achievement scores, this study found sex differences in standardized test scores for eighth grade students only, but sex differences in confidence over all three grades. This shift in the influence of confidence suggests that, for younger students, confidence may be a silent but potentially potent factor affecting later performance.

In summary, the results of this study support a view that sex differences in motivational factors rather than cognitive deficits in computation problems in young students are important for their performance and the performance of older students on mathematics tasks.
REFERENCES


Freeman.


APPENDIX B

QUESTIONS INCLUDED IN THE CONFIDENCE SCALE

1. (Asked after first computation problem is presented, but before it is attempted.) Before you answer, do you think you'll get it right?
2. (Asked after second computation problem is presented, but before it is attempted.) Before you answer, do you think you'll get it right?

3. How do you expect to do in math?
4. How do you expect to do in math computations?

The following question is based on a drawing of a same-sex student seated at a desk in a bedroom.
5. How does s/he think s/he's going to do on her/his math test?

The following questions refer to a drawing of a same-sex student in a (math) classroom situation.
6. (Let's pretend the teacher just showed the class a new way to do the example.) Does s/he think s/he will be able to do the work?
7. (Let's pretend that the teacher just corrected her/his paper. S/he got them all right.) How will s/he do the next time?
8. (Let's pretend s/he got most of them wrong.) How will s/he do
the next time?

9. (When the teacher teaches you something new in math) Do you think you'll be able to do it?
10. How do you feel when the teacher gives you familiar work to do in math?
11. (familiar work) Do you think you'll be able to do it?

For each of the following questions, students were asked to agree along a five-point Likert-type scale.
12. I'm good in math.
13. I can get good grades in math.
14. For some reason, even though I study, math seems very hard for me.
15. I'm no good in math.
16. I'm good at doing computation examples.
17. For some reason, even though I study, computation examples seem very hard to me.
18. Math is my best subject.
19. Math is my worst subject.
20. When I start some new math work, I always think I will get it right.
APPENDIX C

QUESTIONS INCLUDED IN THE ATTRIBUTION MEASURES

SUCCESS ATTRIBUTION MEASURE

1. Let's say you got the right answer in math. Is it because:
   (a) you have a special talent for math?
   (b) math is easy?
   (c) you studied math for a long time?
   (d) the teacher explained math very well?
   (e) you made some guesses and you were lucky?

2. Let's say you got the right answer in computation (like those two examples we did before). Is it because:
   (a) you have a special talent for computation examples?
   (b) computation examples are easy?
   (c) you studied computation examples for a long time?
   (d) the teacher explained computation examples very well?
   (e) you made some guesses and you were lucky?

3. (How do you expect to do in math?) Why do you say that?

4. (How do you expect to do in math computations?) Why do you say that?

5. (When you learn something new in math, do you think you'll be able to do it?) Why do you say that?

6. (When you learn something new in math computations, do you think you'll be able to do it?) Why do you say that?
The following questions are based on a drawing of a same-sex student seated at a desk in a bedroom.

7. (How does s/he think s/he's going to do on her/his math test?) Why do you say that?
8. Let's say s/he thinks s/he's going to get 100% on her/his math test, why does s/he feel this way?

The following questions refer to a drawing of a same-sex student in a (math) classroom situation.

Let's pretend that the teacher just corrected her/his paper. S/he got them all right.)

9. Why did s/he get them correct?
10. Is it because:
    (a) s/he has a special talent for math?
    (b) the math is easy?
    (c) s/he studied for a long time?
    (d) the teacher explained math very well?
    (e) s/he made some guesses and was lucky?
11. (How will s/he do the next time?) Why do you say that?
12. (Let's say this student learned all of his/her math work very well.) Why did s/he do this?
13. (If you were that student, and you had to learn something new in math, do you think you'll be able to do it?) Why do you say that?
14. (When the teacher gives you familiar work to do in math, do you think you'll be able to do it?) Why do you say that?
FAILURE ATTRIBUTION MEASURE

1. Make believe you got the wrong answer in math. Is it because:
   (a) you do not have a talent for math?
   (b) math is hard?
   (c) you did not study math very long?
   (d) the teacher did not explain math very well?
   (e) you made some guesses and you were not lucky?

2. Make believe you got the wrong answer in computation (like those two examples we did before). Is it because:
   (a) you do not have a talent for computation?
   (b) computation examples are hard?
   (c) you did not study computation examples very long?
   (d) the teacher did not explain computation examples very well?
   (e) you made some guesses and you were not lucky?

The following question is based on a drawing of a same-sex student seated at a desk in a bedroom.

3. Let's say s/he thinks s/he's going to fail her/his math test, why does s/he feel this way?

The following questions refer to a drawing of a same-sex student in a (math) classroom situation.

Let's pretend that the teacher just corrected her/his paper.

S/he got most of them wrong.

4. Why did s/he get them wrong?
5. Is it because:
   (a) s/he does not have a talent for math?
   (b) the math is hard?
   (c) s/he didn't study?
   (d) the teacher did not explain the math well?
   (e) s/he made some guesses and was not lucky?

6. (How will s/he do the next time?) Why do you say that?