Sex Differences in Teachers' Evaluative Feedback and Students' Expectancies for Success in Mathematics

Kirby A. Heller
Yale University

Jacquelynne Eccles Parsons
University of Michigan

HELLER, KIRBY A., and PARSONS, JACQUELYNNE ECCLES. *Sex Differences in Teachers' Evaluative Feedback and Students' Expectancies for Success in Mathematics.* CHILD DEVELOPMENT, 1981, **52,** 1015–1019. Two factors that were hypothesized to be related to sex differences in participation in mathematics courses—teachers' use of evaluative feedback and students' expectancies for success—were studied. Following the research of Dweck and her colleagues, 5 types of praise and criticism feedback and attribution statements by teachers were observed in 15 junior high school mathematics classrooms. In addition, 251 students in seventh and ninth grades completed questionnaires assessing expectancies for success on familiar and unfamiliar tasks in mathematics. Contrary to past research, no sex differences were found in the patterns of evaluative feedback used by teachers. Sex differences were also not found in students' expectancies for success on familiar tasks; however, girls had lower expectancies for success than boys on unfamiliar or future tasks. Possible explanations for the discrepancies between these results and those of Dweck, Davidson, Nelson, and Enna are discussed.

When high school mathematics courses are optional, girls choose to take fewer mathematics courses than do boys (Sherman & Fennema 1977). This sex difference in participation rates has both serious short- and long-term ramifications. Sherman and Fennema (1977) have argued that the failure of many females to take advanced mathematics courses may be an important cause of the sex differences traditionally found in mathematics achievement in adolescent males and females. When the number of years of studying mathematics is equated for males and females, few sex differences are typically found on high school achievement tests. Even more serious are the effects of underparticipation in mathematics by females on future educational and career options. Because mathematics is a prerequisite for many college majors, limited mathematics training effectively precludes women from entering numerous careers that include, but are not limited to, those in science and mathematics.

The purpose of the present study was to investigate sex differences in two factors that might be related to participation in advanced mathematics. The first factor, teachers' use of evaluative feedback in the classroom, has been shown to be related to students' achievement test scores and achievement attributions (Brophy & Good 1974; Dweck, Davidson, Nelson, & Enna 1978). To assess sex differences directly in the patterns of evaluative feedback used by teachers, Dweck et al. (1978) studied teacher-student interactions in three fourth- and fifth-grade classrooms. They found that less than one-third of the total criticism directed at boys was contingent on academic performance, while the majority of the criticism was for conduct or nonintellectual aspects of work. In contrast, more than two-thirds of the total criticism directed at girls was aimed at the quality or correctness of their work. The opposite pattern was seen in the use of praise. The percentage of total praise that was directed to the quality

This research is based on a dissertation submitted by the first author in partial fulfillment of the requirements for the doctoral degree at the University of Michigan under the supervision of the second author. It was supported by grant NIE-G-78-0022 from the National Institute of Education to the second author. We would like to thank Carol Kaczala and Judith Meece for their assistance in collecting and analyzing data. Requests for reprints should be sent to Kirby A. Heller, Committee on Child Development Research and Public Policy, National Academy of Sciences, 2101 Constitution Avenue, N.W., Washington, D.C. 20418.

of work was higher for boys than for girls. In addition, the teacher was more likely to attribute failure to a lack of effort for boys than for girls.

Dweck et al. (1978) hypothesized that since boys receive frequent criticism for academic as well as nonacademic behaviors, they can discount these negative evaluations as indications of their own abilities and attribute them to a characteristic or enduring attitude of the teacher. Since the criticism that girls received was more specific, it was more informative about their academic abilities. The same reasoning applies to the use of praise. Because girls received more generalized praise from the teacher, this feedback was less informative about their ability than the more specific type of praise that boys received.

The second factor studied in the present study was students' expectancies for success. Past research has shown that expectancies for success are linked to important achievement-related behaviors, such as performance, task persistence, and task choice (e.g., Battle 1966; Stein & Bailey 1973). The relation between expectancy for success and performance increases with age, and by adolescence expectancies are consistently related to the performance of both boys and girls (see Stein & Bailey 1973).

The existence of a sex difference in expectancies for success from mid-childhood onward, but most especially during the adolescent years, is a common finding (see Maccoby & Jacklin 1974). Although there are some inconsistencies in this research, the majority of studies using a diverse assortment of tasks has found lower expectancies in females than males, especially when subjects are asked to form expectancies about novel or unfamiliar tasks (e.g., Parsons, Ruble, Hodges, & Small 1976).

In the present study, students in seventh and ninth grades were asked to complete questionnaires assessing their expectancies for success in mathematics. This age range was chosen because previous research has identified the junior high school years as those in which sex differences in attitudes toward and achievement in mathematics begin to emerge (see Maccoby & Jacklin 1974). In addition, teacher-student interactions in junior high school mathematics classrooms were observed, and the teachers rated each student's likelihood of success in mathematics. Four sets of analyses were planned: (*a*) comparisons of the teachers' use of discriminant praise and criticism for boys versus girls; (*b*) comparisons of the teachers' use of praise and criticism for students having high versus low teacher expectancies; (*c*) comparisons of the teachers' causal attributions and expectancy statements for boys versus girls; and (*d*) comparisons of boys' and girls' expectations for their own math performance.

Method

Subjects

Eight seventh-grade mathematics classrooms and seven ninth-grade mathematics classrooms were observed. The 15 classrooms were drawn from junior high schools in middle- to upper-middle-class neighborhoods in a small midwestern city. Criteria for inclusion in the sample were: (*a*) voluntary agreement by teacher to participate, and (*b*) a mathematics curriculum that was at grade level or slightly advanced. The second criterion was especially relevant in the ninth grade, where at least four different mathematics courses were offered. Only one of the seventh-grade classes being observed was tracked; the rest had wide variability in the students' levels of ability.

The eight seventh-grade classrooms consisted of 107 boys and 98 girls. The teachers included six males and two females. In the seven ninth-grade classrooms, 100 boys and 96 girls were observed. Four of the ninth-grade teachers were male; three were female.

In 12 of these classrooms, expectancies for success in mathematics were assessed for the subset of the students who volunteered to complete questionnaires and received parental permission. Sixty-four girls and 57 boys in the seventh grade (59% of the total sample), and 72 girls and 59 boys in the ninth grade (67% of the total sample) completed the questionnaires.

Observational System

The observational system used was a modified version of two other systems: Brophy and Good's Teacher-Child Dyadic Interaction system (1970) and Dweck's observational procedures (Dweck et al. 1978). The focus of the classroom observations was on public and private interactions between the teachers and individual students. Every instance of praise and criticism to an individual was coded. The teacher's use of praise and criticism was subdivided into three categories: praise or criticism directed to the quality of the work (e.g., "You really are on the ball"; "You should know the answer by now"); praise or criticism directed to the form of the work (e.g., "This paper is very neat"; "You should be working in pencil, not pen"); and praise or criticism directed to conduct ("Look how quietly Ellen is working"; "Stop talking and pay attention").

Explicit causal attributional and expectancy statements made by the teachers were also coded. The attributional statements were explanations provided by the teacher of a student's response (e.g., "You did it too fast"; "You've forgotten how to do it? Well you've been sick"). These were coded into four categories: task difficulty; effort; ability; and incorrect use of mathematical operation, for example, "You used the wrong numerator." Expectancy statements were explicit statements regarding the teacher's expectation for a child's performance on a specific problem. These were categorized into four levels, based on how positive or negative the statement was; for example, level 1 (the most positive level) included statements such as "These are hard, but you shouldn't have any trouble with them." Level 4 (the least positive level) included statements such as "You'll get this if you take second-year algebra, if you get that far." Two coders rated each type of statement. The percentage of agreement for attributional statements was 84%; the percentage of agreement for expectancy statements was 91%.

Student Questionnaire

The questions relevant to this study were administered as part of a longer questionnaire. Students were asked six questions about their expectancies for success in mathematics. To test the hypothesis that sex differences are found only in expectancies for novel or less familiar tasks, the items were divided into those assessing expectancies for success on a familiar task, such as expectancies for success on the next mathematics test, and those measuring expectancies for success on a less familiar task, such as expectancies for success in advanced high school mathematics. Responses were based on a seven-point scale that ranged from "not at all well" to "very well." Cronbach's coefficient Alpha was used to assess reliability. The Alpha coefficients for the entire scale and for each of the two subscales ranged from .77 to .85. The correlation between the two subscales was $r\ (337) = .62$.

Teacher Questionnaire

Teachers were asked to rate how well each student was doing this year and how well the teacher expected each student to do in an advanced high school math course. Responses to the first item consisted of the teacher's ranking of each student's position in the class in terms of quintile, percentile rankings, for example, the top 20% of the class, etc. Responses to the second item were based on a seven-point scale ranging from 1 = "very poorly" to "very well."

Procedure

Four females and one male served as observers. Training lasted approximately 3 weeks. Discussion, the coding of written transcripts, the coding of videotapes of classrooms, and actual practice in the classrooms were used in training. The observers were in each classroom for 13–15 hours during a 2-month period in the spring. Observations were recorded for the last 10 of these hours. Classrooms were observed on different days of the week; data were not collected during class periods in which atypical events occurred (e.g., films, tests).

Observers' agreements were assessed for 3 or 4 hours per observer, using criterion-related agreement measures. Both the author and one other observer served as criterion coders. Reliability estimates were obtained prior to data collection and approximately halfway through the observations. The mean percentages of agreement for each observer ranged from 76% to 82% before data collection; they ranged from 75% to 86% during the study. Percentages of agreement for individual categories ranged from 70% to 100%.

The teacher and student questionnaires were administered after the classroom observations were completed.

Teachers, students, and observers were not told about the specific hypotheses of the study. They were told that the focus of the study was boys' and girls' behaviors in mathematics classrooms. All were given detailed information following the study.

Results

Discriminant Use of Praise and Criticism: Student Sex Effects

To test the Dweck et al. (1978) hypotheses, five variables were created: percentage of praise directed to the quality of the work; percentage of praise directed to the form of the work; percentage of criticism directed to the quality of the work; percentage of criticism directed to the form of the work; and percentage of criticism directed to conduct. Conduct was rarely praised and was, therefore, not included in these analyses. These variables were formed following the procedures recommended by Brophy and Good (1970). For example, the percentage of criticism directed to conduct was obtained by dividing the number of criticisms directed to conduct by the number of criticisms directed to the quality of work plus the number of criticisms directed to the form of the

work plus the number of criticisms directed to conduct.

Because there was wide variance among classrooms in the frequency of evaluative feedback, scores were transformed into standard score form. Each student in a class thus retained his or her position relative to others in the class for each variable, and units were comparable across classrooms. For initial analyses, the classroom was used as the unit of analysis and was treated as a random factor nested within grade. Thus, the mean scores of boys and girls within each classroom for each of the five variables were the dependent measures in the initial set of ANOVAs.

Neither significant main effects for sex or grade, nor interactions of the two were found for any of the five variables. Although the classroom was considered the appropriate unit of analysis, additional analyses of variance were also done on individual scores so that results could be compared directly to those of Dweck et al. (1978). Again, no significant effects were found.

While sex differences were not found in teachers' feedback for the sample as a whole, it is possible that patterns of feedback for boys and girls may differ in certain classrooms, or for a subset of the students. To investigate whether some teachers might be exhibiting the sex-differentiated pattern described by Dweck while other teachers might not, analyses of variance with sex of student (two levels) and teacher (15 levels) as the independent variables were run for each of the five dependent variables outlined earlier. By necessity, the individual was the unit of analysis for these ANOVAs. None of the analyses yielded significant main effects or interactions.

Praise and Criticism: Teacher-Expectancy and Student-Sex Effects

In a summary of the research on effects of teachers' expectancies on their behavior in the classroom, Brophy and Good (1974) concluded that sex and achievement level interact to influence the nature of the evaluative feedback received from the teacher. In general, high-achieving boys receive more praise than either low-achieving boys or high- and low-achieving girls. Girls of all achievement levels are treated more similarly. To test whether patterns of feedback differed for boys and girls for whom the teacher had high versus low expectancies, students were divided into two groups based on teachers' expectancies as measured in the teacher questionnaire. The classroom was the unit of analysis, and the dependent measure was the mean score for boys and girls within each expectancy group (high or low) and classroom. The 2 (student sex) $\times$ 2 (teacher expectancy) $\times$ 2 (grade) ANOVAs were run on these mean frequencies for praise and criticism. No expectancy-group main effects or sex $\times$ expectancy-group interactions were found for the use of praise. There was a significant main effect for sex on the amount of criticism. Boys received more criticism in dyadic interactions with the teacher than did girls ($F = 20.72$, $p < .001$). The sex $\times$ expectancy-group interaction for criticism was not significant. There were no significant effects involving the grade-level factor.

Attribution and Expectancy Statements: Sex- and Teacher-Expectancy Effects

Very few attributional statements were made by the teachers ($N = 76$), especially following successful outcomes. Consequently, only the attributions following unsuccessful outcomes were coded. The vast majority of attributions were to lack of effort ($N = 52$). The χ^2 analyses of the relation between student sex and type of attributional statement and between teacher-expectancy group and type of attributional statement were not significant ($p > .10$ in each case), indicating that teachers' attributional statements did not vary as a function of either sex or teacher expectancy.

Similar results emerged for expectancy statements. Very few were made ($N = 73$), and the majority fell into level 3 ($N = 40$), for example, "Come-on, think!" "We had this before" or "You should know that." The χ^2 analyses revealed no significant effects of either student sex or teacher expectancy on the probability of occurrence.

Student Expectations for Their Own Performance

The data from the expectancy of success scale, as well as the two subdivisions of the scale (expectancies of success for current performance and expectancies of success for future tasks), were each subjected to a 2 (sex of student) $\times$ 2 (grade level) ANOVA. Since the classroom was the unit of analysis, the dependent variable was the mean score of boys and girls within each classroom.

Neither significant main effects nor interactions were found for the total scale or for the scale measuring expectancies of success for current performance. There was a significant main effect for sex in expectancies of success for future tasks, $F(1,10) = 8.35$, $p < .02$. Girls had lower expectancies ($M = 4.54$) than boys (M

= 5.12). (The 2 × 2 ANOVAs using the individual as the unit of analysis yielded comparable results.)

Discussion

The results from this study clearly do not replicate those found by Dweck et al. (1978). There was no evidence supporting the hypothesis that teachers use evaluative feedback differentially or make differential attributions for boys and girls in junior high school mathematics classes.

This study differs from that of Dweck et al. (1978) in several ways. First, the greater number of classrooms observed in this study as compared to Dweck et al. provides a broader sampling of classrooms and schools. It is possible that the three teachers studied by Dweck et al. were not a representative group.

Second, the students in this study were observed in junior high school mathematics classrooms; the students in the study by Dweck et al. were in fourth and fifth grades and were observed in a variety of subject areas. Although the effects of grade level on the teacher's use of feedback have not been systematically studied, it seems likely that teachers' feedback is in part determined by the age of the students. For example, teachers in junior high school may give less feedback about the form of the work than teachers in elementary school, who may believe that learning to follow classroom rules and procedures is as important as specific academic skills.

It is important to note, however, that sex differences in expectancies for mathematics do not emerge with any consistent regularity until late junior high school (see Parsons, in press). Thus, it seems especially significant that we found no sex differences in the behaviors of junior high school teachers. As Popper (1979) has discussed, the failure to replicate the findings of Dweck et al. (1978) at this age level and for this subject area is as important to our understanding of the influences of school experiences on achievement expectancies as a significant effect would have been. Future research on teacher's influences is clearly needed before we will be able to evaluate the impact of teachers on children's expectations for their own achievement. Additional information on fourth- and fifth-grade classrooms is needed to determine whether our failure to replicate is a consequence of the different grade level tested or the broader generalizability of the Dweck et al. (1978) findings. Given the potential impact of the Dweck et al. results on teacher education, tests of the replicability and the range of generalizability of those results are called for.

The results from this study, however, do support previous research on sex differences in students' expectancies for success. Girls and boys had similar expectancies for current performance; girls had lower expectancies than boys for future or unfamiliar tasks. These findings suggest that the relation between expectancies for success and participation in advanced mathematics should be studied. Research is currently in progress to assess this relation.

References

Battle, E. Motivational determinants of academic competence. *Journal of Personality and Social Psychology*, 1966, **4**, 634–642.

Brophy, J. E., & Good, T. L. Good system (teacher-child dyadic interaction). In A. Simon & E. Boyer (Eds.), *Mirrors for behavior: an anthology of observation instruments continued, 1970 supplement*. Vol. A. Philadelphia: Research for Better Schools, 1970.

Brophy, J. E., & Good, T. L. *Teacher-student relationships: causes and consequences*. New York: Holt, Rinehart & Winston, 1974.

Dweck, C. S.; Davidson, W.; Nelson, S.; & Enna, B. Sex differences in learned helplessness, II: The contingencies of evaluative feedback in the classroom, and III: An experimental analysis. *Developmental Psychology*, 1978, **14**, 268–276.

Maccoby, E. E., & Jacklin, C. N. *Psychology of sex differences*. Stanford, Calif.: Stanford University Press, 1974.

Parsons, J. E. Expectancies, values and academic choice: origins and change. In J. Spence (Ed.), *Assessing achievement*. San Francisco: W. H. Freeman, in press.

Parsons, J. E.; Ruble, D. N.; Hodges, K. L.; & Small, A. W. Cognitive-developmental factors in emerging sex differences in achievement-related expectancies. *Journal of Social Issues*, 1976, **32**, 47–61.

Popper, K. R. *Objective knowledge: an evolutionary approach* (Rev. ed.). London: Oxford University Press, 1979.

Sherman, J., & Fennema, E. The study of mathematics by high school girls and boys; related variables. *American Educational Research Journal*, 1977, **14**, 159–168.

Stein, A. H., & Bailey, M. M. The socialization of achievement orientation in women. *Psychological Bulletin*, 1973, **80**, 345–364.