Journal of Educational Psychology
1998, Vol. 90, No. 3, 516–527

0022-0663/98/$3.00

Bias or Responsivity? Sex and Achievement-Level Effects on Teachers' Classroom Questioning Practices

Ellen Rydell Altermatt, Jasna Jovanovic, and Michelle Perry
University of Illinois at Urbana-Champaign

The authors examined rates of both teacher responsiveness and student participation in the classroom question-asking context. Participants were 165 students and their teachers in 6 science classrooms. Teachers in 3 of the 6 classrooms called on male students to answer questions more often than would have been expected on the basis of the number of boys in the classroom. In none of the classrooms, however, did teachers call on boys more often than would be expected on the basis of the heightened volunteering rates of their male students. No systematic sex or achievement-level differences were found in the types of questions that students responded to. These findings suggest the need to focus on the role that both teachers and their students play in creating and maintaining sex differences in the teacher–student interaction context.

Beginning early in children's schooling, boys have been found to have more positive interactions with their teachers than do girls, including more opportunities to answer questions (Becker, 1981; Cherry, 1975), more individual instruction (Stallings, 1985), and more encouragement and feedback (Morse & Handley, 1985; Serbin, O'Leary, Kent, & Tonick, 1973). At the same time, researchers (e.g., Jones & Wheatley, 1990; Serbin et al., 1973; Stake & Katz, 1982) have demonstrated that boys are more likely than girls to be the recipients of management-oriented contacts, including behavioral criticism and punishment. These findings—particularly the former—have received considerable attention among educators and the popular press and have been used to support the notion that girls and boys have very different experiences in U.S. classrooms. In particular, claims have been made that girls are shortchanged in the classroom because they are afforded fewer opportunities than boys to interact with teachers in academically meaningful ways (American Association of University Women [AAUW], 1992; Sadker & Sadker, 1990).

Underscoring these differential interactions have been suggestions that boys are involved in more high-level or higher order question exchanges with teachers than are girls (Sadker & Sadker, 1982, 1989). The implication is that teachers foster positive outcomes among boys more than girls not only by interacting more with boys but also by spending more quality academic time with boys. These discrepancies would be expected to be particularly detrimental for girls in stereotypically male sex-typed school subjects, such as mathematics and science, where it has been found that teacher–student interactions may be most likely to favor boys (Good & Slavings, 1988; Leinhardt, Seewald, & Engel, 1979; Morse & Handley, 1985). Sex differences in teacher–student interactions have, in fact, been hypothesized to contribute to sex-related differences in student performance and attitudes in these domains (Beller & Gafni, 1996; Catsambis, 1995; Jones, Mullis, Raizen, Weiss, & Weston, 1992; Kahle & Rennie, 1993; Licht, Stader, & Swenson, 1989; Weinburgh, 1995).

The role that students themselves may play in influencing the number of questions to which they are asked to respond has largely been ignored both in the empirical literature and by the popular media. This lack of attention has been particularly true of the literature that considers not only question quantity but question quality as well. Consistent with this perspective, claims of teacher bias have typically been based on findings that teachers call on boys at rates that are disproportionate to their numbers in the classroom. Efforts to combat this bias have, consequently, been focused on understanding and modifying teacher behavior. The

Ellen Rydell Altermatt, Department of Psychology, University of Illinois at Urbana-Champaign; Jasna Jovanovic, Department of Human and Community Development, University of Illinois at Urbana-Champaign; Michelle Perry, Department of Educational Psychology, University of Illinois at Urbana-Champaign.

This research was supported by grants from the National Science Foundation (RED 95-96017), the University of Illinois Research Board, and the Cooperative State Research, Education, and Extension Service, U.S. Department of Agriculture (45-323). The views expressed herein do not necessarily reflect those of the sponsoring agencies. Initial findings were presented at the 1997 annual meeting of the American Educational Research Association, Chicago, March 24–28.

We would like to thank the principals, teachers, and students from the schools who participated in the study. We thank the Fermi National Accelerator Laboratory Education Office staff for their assistance in identifying teachers for the study. Many thanks to Janis Jacobs, Reed Larson, and Eva Pomerantz for their valuable suggestions on previous drafts of this article and to G. Alfred Forsyth for his contributions to the early development of this work. We thank Douglas Simpson, Jay Verkuilen, Susan Butkus, and Olga Geling for their statistical consultation. Finally, we thank Candice Dreves, Sally Steinback King, Jessica Cook, and Cheri Tornow for their significant help on the project.

Correspondence concerning this article should be addressed to Jasna Jovanovic, Department of Human and Community Development, University of Illinois at Urbana-Champaign, 1105 West Nevada, Urbana, Illinois 61801. Electronic mail may be sent to jasna@uiuc.edu.

possibility that differences in the proportion of times that boys and girls are called on to answer questions may largely be accounted for by sex differences in student volunteering rates has, in contrast, been routinely neglected. The latter situation may still constitute a bias in some sense insofar as girls continue to receive less attention than boys. On the other hand, if teachers are responding to students at rates that are consistent with the heightened volunteering rates of boys, recommendations for change must also include dedication of resources to understanding the reasons for girls' relative nonparticipation and toward finding ways to encourage girls to become fuller participants in the classrooms question-asking context. A central goal of the present investigation, then, was to assess whether student volunteering rates could, in fact, account for sex differentiated student–teacher exchanges.

Classroom Questioning and Student Learning

Teacher-to-student questioning has long been a source of interest to researchers examining sex differentiation in schools. This situation is not surprising given the relative frequency and visibility of question-asking in the elementary and secondary school classroom. Skilled instructor-to-student questioning is generally believed to be helpful in eliciting thoughtful and reflective student responses that, in turn, may lead to the enhancement of student cognition (Carlsen, 1991; King, 1991; Martin & Pressley, 1991; Perry, Vanderstoep, & Yu, 1993; Wong, 1991).

Reviews of the questioning literature have, however, produced mixed results in determining whether questions help to focus student attention on the learning task and whether higher order questions, in particular, are related to increases in children's cognitive competence. For example, Winne (1979) and Samson, Strykowski, Weinstein, and Walberg (1987) found no consistent effects of teachers' use of higher order questions on student achievement. Redfield and Rousseau (1981), on the other hand, concluded that there is, in fact, strong support for the hypothesis that teachers' predominant use of higher level questions positively affects student achievement. Despite the inconsistencies in the findings, researchers have continued to study and, in many cases, to recommend the use of higher order teacher-to-student questions in enhancing student cognition (French & MacLure, 1981; Morse & Handley, 1985).

Sex Differences and Question Type

The notion that boys are afforded more opportunities than girls to respond to complex questions has been suggested anecdotally by Sadker and Sadker (1989, 1990) and by Stitt (1988). However, relatively few studies to date have been conducted to assess empirically the relationship between student sex and the questions that students are called on to answer. Early support for the hypothesis that boys are favorably differentiated from girls with regard to abstract questions (e.g., convergent and divergent as compared with memory questions) was provided by Sikes (1971) and Good, Sikes, and Brophy (1973). Good et al. (1973) obtained data from 16 seventh- and eighth-grade mathematics and social studies classrooms that showed not only that boys are provided with more response opportunities overall but also that they are given the chance to respond to more process questions (i.e., those involving an explanation of complex phenomenon or problem-solving strategies) than their female classmates.

Later studies, however, have provided somewhat mixed results. For example, Becker (1981) found that in geometry classes, boys were called on to respond to more higher order process questions than females. In contrast, Hillman and Davenport (1978) reported that boys were called on more than girls for product, but not for process, questions. Additionally, Jones and Wheatley (1990) reported no significant sex differences in the opportunities that high school science students had to answer abstract questions, whereas Scantlebury and Kahle (1993) found that boys were asked more higher order questions than girls in some, but not all, of the science classrooms investigated.

In short, although there is some indication that boys are asked more abstract questions than are girls, studies are clearly required to verify this finding. Regardless, even if researchers can conclude that at least in some classrooms boys are asked different types of questions than girls, the question still remains whether the behaviors of teachers, students, or both can account for these differences.

Teacher Bias or Teacher Responsivity?

In describing the equitable classroom, Scantlebury and Kahle (1993) suggested that the type and percentage of questions answered by boys and girls should approximate the percentage of boys and girls in the class. Although this type of analysis would clearly take into consideration the teacher's role in question-asking interactions (e.g., on whom the teacher calls), it fails to consider fully the influence that students themselves may have on the types and numbers of questions that they are asked (e.g., by considering whether boys and girls may be volunteering at rates that are disproportionate to their numbers in the classroom). In other words, it is likely that a bidirectional relationship exists between teachers and students wherein students are, in part, determiners of their own experiences (Bell, 1968; Lerner, 1982; Lerner & Busch-Rossnagel, 1981).

Eccles and Blumenfeld (1985) and Fennema and Peterson (1985) supported this contention, maintaining that preexisting differences in students' behavior are as responsible for differentiation in student–teacher interactions as are teacher biases. Similarly, Brophy (1985) concluded that gender differences in elementary school classroom experiences were due primarily to the gender role-related differences in the behavior of students themselves rather than to any general tendency for teachers of either sex to treat boys and girls differently. Whether this hypothesis holds true for research that examines both the number and types of questions that male and female students are asked to respond to in the classroom setting was a primary focus in this study.

Student Achievement as a Moderator of Sex Effects

Student ability or achievement level may also be an important factor influencing the types and numbers of questions that students volunteer for or are called on to answer or both. Early evidence for achievement-by-sex differences in student–teacher interactions is provided by Good et al. (1973). High-achieving students in the observed classrooms were found to be more active participants, to secure more teacher contacts, and to receive more positive teacher feedback than were low-achieving students. Similar results were found for teacher-to-student questioning patterns with high-achievers answering more open questions, process questions, and product questions than their peers. High-achievers were also more likely to call out responses to questions than other students. Although both high-achieving boys and high-achieving girls were found to have more and higher quality interactions with their teachers than low-achievers, the differences were most highly marked between high- versus low-achieving boys. High-achieving boys received the most favorable teacher treatment. In contrast, low-achieving boys had the poorest contact patterns with teachers. Similar findings are reported in studies by Wang and Weinstein (1980) and Parsons, Kaczala, and Meece (1982). In the present study we examined whether achievement would, in fact, moderate any sex effects in the types and numbers of questions that students were called on to answer.

The Present Study

In the present study, we tested four related hypotheses regarding the quantity and quality of questions to which boys versus girls are permitted to respond in the science classroom. First, we assessed whether boys are indeed asked to respond to more questions overall than are girls. Because previous studies support this disparity, we expected similar results. Second, we examined the hypothesis that boys are called on to respond to different types of questions than girls. Third, and primary to the present investigation, we tested the hypothesis that teacher responsiveness (i.e., on whom a teacher calls to answer a given question) is positively related to student responsiveness (i.e., boy–girl volunteering rates). More precisely, we expected that sex differences in the number and types of questions (i.e., memory, convergent, and divergent) to which students are asked to respond would not represent an overall teacher bias (e.g., in favor of boys), but, rather, a responsivity on the part of teachers to the heightened volunteering rates of their (male) students. This study is, in fact, the first to assess empirically whether boys are more likely than girls to be called on for different types of questions, accounting for possible differences in student volunteering rates both overall and for these different question types. Finally, we explored the hypothesis that student gender and achievement level may interact to influence the types and numbers of questions that students are called on to answer.

Method

Participants

Participants were 165 students in 6 fifth- through eighth-grade performance-based (i.e., hands-on) science classrooms. Fifty-three percent were girls, and the age of students ranged from 10.08 to 14.17 years, with a mean age of 12.21 years. Of the sample, 76% was European American, 5% was African American, 4% was Latino, 3% was Asian American, and 12% was of other ethnic origin.

The 6 science teachers (3 men, 3 women) associated with these science classrooms were participants in a short-term longitudinal study that examined the effects of performance-based science teaching on boys' and girls' attitudes and achievement in science (Jovanovic & King, in press). These 6 teachers were initially drawn from a pool of approximately 20 teachers nominated by science educators as exemplary hands-on instructors in Illinois. To control for teaching experience and to ensure that the sample of classrooms would be comparable with respect to socioeconomic status, teachers with 10 or more years of experience, teaching in middle- to upper-middle-class school districts were considered for inclusion.[1] From this sample, 10 teachers from several school districts in Illinois were interviewed to ensure that certain performance-based science teaching criteria were met by each teacher in his or her science classroom. A critical selection criterion was evidence of the teacher's interest in increasing girls' participation in science. As a result of this evaluation process, 6 teachers from five schools representing four school districts were selected to participate. All 6 teachers hold advanced degrees or have continued higher education experience in science. In addition, all but 1 of the teachers are members of the Illinois Science Teachers Association. Of the 6 teachers, 4 exclusively teach science. The science content covered during the school year varied both within and across classrooms. The fifth- and sixth-grade classrooms covered topics drawn from biology or life science, physical science, and chemistry. The seventh-grade classroom covered topics primarily in the biological or life sciences. The 2 eighth-grade classrooms focused on physical science and chemistry, although both classrooms included units involving life science.

Procedure

Whole-class science lessons (average time = 40 min) were observed in each of the classrooms one or two times per month over the course of 1 academic year. In total, 70 lessons were observed across the six classrooms. The teachers were instructed that the purpose of the study was to assess students' response to performance-based science instruction. Although the teachers understood that as part of this assessment we were interested in differences between boys and girls in the classroom, they were not specifically told that the frequency of teacher–student interactions by sex would be coded. Lessons were audio- and videotaped for subsequent, detailed coding. The video camera was situated such that the number of student volunteers per question and the identity of the student called on by the teacher would be maximally visible.

In brief, a two-step coding scheme was used: (a) to determine the cognitive type of each question asked by the teacher and (b) to record the number of boys and girls (by achievement level) who volunteered for each question. This step included identifying the sex and achievement level of the student called on by the teacher to

[1] In our sample, 61% of the mothers and 71% of the fathers of the students had college degrees or higher.

answer each question. The details of this procedure are given below.

Question types. Question transcripts were created from the videotapes made during the observed lessons. Subsequently, the cognitive operations established by Guilford (1979) were used to categorize each of the questions asked by teachers into one of three types. A simplified version of this classification system was used, differentiating questions into those involving memory and convergent and divergent cognitive processes. Procedural questions (e.g., "Who can read the next paragraph?") were also coded but were not included in subsequent analyses because of their relative infrequency.

Memory questions were those for which the student simply needed to recall information previously learned in class or to produce obvious factual information from everyday experiences. Questions such as "Who can remember any one of the three parts of the neuron?" and "Who can name a type of bubble gum?" are examples of memory questions from one of the classrooms observed for the current study. Questions were labeled as *convergent* if a correct response required students to incorporate already existing knowledge, material learned in class, or both, in such a way as to close in on a solution that had not been explicitly taught in class. Questions such as "What causes the meniscus?" and "If you fill a cup all the way up with water and ice, and let the ice melt, will it overflow?" are examples. Finally, *divergent* questions were those that, according to Guilford (1979), required "a broad searching or scanning" of existing knowledge to provide an answer that extends beyond this knowledge (p. 161). In most instances, these questions both allow for many possible responses and require students to come up with one or several of the correct alternatives. "What are some things that researchers from *Consumer Reports* could do to be fair and unbiased in their reporting?" and "What are some games that you play where you can control the distance by the amount of force you put on an object?" are examples.

Two raters used question transcripts to code each of the questions according to cognitive type. Disagreements were resolved through discussion. It should be noted that, in contrast to much of the research conducted in the area of classroom questioning, this study did not specifically classify questions as higher or lower order. Researchers have examined a number of problems that are evident in the higher level-lower level question distinction. For example, several researchers have suggested that it is inappropriate to assume that higher order questions are necessarily better than lower order questions (Blank & White, 1986; Carlsen, 1991; Cazden, 1986; Farrar, 1986; French & MacLure, 1981; Parker, 1983; Samson et al., 1987). Guilford's (1979) taxonomy does not make such distinctions but merely classifies questions in terms of the different types of thought processes necessary to answer memory, convergent, and divergent questions without any explicit or necessary hierarchical classifications.

Response opportunities. The second step of the coding procedure was accomplished using a portion of the Brophy-Good Teacher–Child Dyadic Interaction System (Brophy & Good, 1970). Within the Brophy-Good system, a response opportunity is coded only in those instances in which an individual child responds to a question posed by the teacher and in which the teacher acknowledges in some way that a response has been given (usually through feedback). Distinctions are made, moreover, between direct questions, call outs, and open questions. *Direct* questions are those in which the teacher calls on a student who has not volunteered to answer the question. *Call outs* are those in which a student spontaneously provides an answer to a question posed by the teacher without waiting to be called on to respond. Finally, the *open-question* category includes all instances in which the teacher asks a question, waits for the students to raise their hands, and then calls on one of the students whose hand is raised. Because the principal purpose of this study was to determine the relative importance of teacher and student influences on the types and numbers of questions that students are asked, it was of particular interest to consider all response opportunities that were partially teacher-afforded and partially child-created. Only open questions meet this criterion. Thus, although student-initiated questions, direct questions, and call outs were recorded, open questions were the primary focus for investigation (Brophy & Good, 1974). This was, in fact, the dominant question category (74%) in all six classrooms.

Student volunteering. Student volunteering rates were determined by recording, for each question, the identification number of each student who raised his or her hand. To ensure the accuracy and replicability of the primary rater's coding, both overall and over time, at least 25% of all questions were recoded by a second rater. Values for Cohen's kappa ranged from .80 to .91 across the six observed classrooms, indicating excellent interrater agreement.

Science achievement. Students' total scores on the Illinois Goals Assessment Program Standardized Test in Science (IGAP; Illinois State Board of Education [ISBE], 1993) were used as a measure of students' science ability.[2] The IGAP is designed to measure students' knowledge with respect to the state's goals for science learning. The examination is administered in all Illinois public schools in Grades 4, 7, and 11. Therefore, for all but the 7th-grade students, IGAP scores were generated before the start of the present study. IGAP scores range from 0 to 500, with a statewide mean of 250 and a standard deviation of 100. In the present sample, the mean score across classrooms was 310.21 ($SD = 93.07$). The mean scores for boys (324.97, $SD = 87.42$) and for girls (297.00, $SD = 96.45$) were not significantly different, $t(161) = 1.93, p = .06$.

In the current study, students were divided by a median split within sex in each classroom into high-achievement and low-achievement groups. The median split was conducted separately within each classroom because of significant differences in the mean IGAP scores for students in Classroom 4 and each of the other classrooms ($ps < .05$). This choice also seemed reasonable insofar as the classroom is the level at which teachers would be expected to differentiate between high- and low-achievers. Therefore, student achievement level refers to the performance of each student relative to those of others in his or her classroom. A breakdown of students by sex and achievement level within each classroom is provided in Table 1.

Results

A total of 677 nonprocedural questions were coded during the observed lessons across the six classrooms. Of these, 504 (74%) were open questions, 93 (14%) were direct questions, and 80 (12%) were call outs. Boys were asked 61% of all direct questions and were responsible for 54% of all call

[2] IGAP equivalency scores were computed for 25 students for whom IGAP results were not available. For these 25 students, either their final science grade in the year that observations were conducted, their final science grade from the previous academic year, or their score on an alternative science achievement test was used to equate these students with other classmates for whom the IGAP was available. The mean IGAP score among the classmates on which the student was equated (e.g., same final grade, same previous grade, or same achievement score) was then used as the equivalent IGAP score. This was considered reasonable given the moderate to high correlations found between IGAP scores and each of the three alternative achievement measures.

outs. These rates are consistent with those found for open questions, with boys answering 62% of open questions across the six classrooms. A breakdown of the number of memory, convergent, and divergent open questions that were asked by teachers in each of the classrooms is provided in Table 2.

Overall, boys accounted for 63% of all student volunteering. On average, there were 1.59 male volunteers and 1.01 female volunteers for each question. Of the 17 most responsive students (i.e., students constituting the top 10% of student volunteers across the six classrooms), 14 were boys. The 3 most responsive boys in each classroom accounted, on average, for 53% of total male volunteering and were called on to answer 51% of all open questions for which a boy was selected. Similarly, the 3 most responsive girls in each classroom accounted for an average of 58% of total female volunteering and answered 48% of all questions answered by girls.

Data Analyses

Chi-square analyses. Previous research on sex differences in student–teacher exchanges has most often relied on simple frequency counts to determine the proportion of times that boys versus girls are called on to answer questions in the classroom question-asking context. To maintain consistency with this prior research, a similar approach (i.e., chi-square) was taken in the current study. This approach, moreover, allowed us to contrast the results obtained when using the traditional approach of calculating expected rates of teacher responsiveness based simply on the number of boys and girls in the classroom and the current approach that considers the volunteering rates of these students as well.

Chi-square analyses were conducted to compare observed and expected rates of teacher responsiveness (i.e., which student the teacher selected to answer each question) and student volunteering for all open questions. Specifically, chi-square analyses were designed to test the degree to which (a) observed teacher responsiveness differed from that which would be expected based on the proportion of boys and girls (and high and low achievers) in the class, (b) observed student volunteering rates differed from expected rates based on the proportion of boys and girls (and high and low achievers) in the class, and (c) observed teacher responsiveness differed from expected teacher responsiveness based on observed student volunteering rates.

Table 1
Breakdown of Students by Sex and Achievement Level Within Each Classroom

			Achievement level and student sex			
Classroom	Teacher sex	Grade	High male	Low male	High female	Low female
1	Male	5	7	8	7	6
2	Male	7	7	7	8	8
3	Male	6	9	8	6	7
4	Female	6	4	4	8	7
5	Female	8	7	6	7	6
6	Female	8	6	6	8	8

Table 2
Frequencies for Memory, Convergent, and Divergent Questions in Each Classroom

	Type		
Classroom	Memory	Convergent	Divergent
1	29	20	20
2	46	71	59
3	39	26	17
4	24	37	23
5	17	13	5
6	22	14	22

The chi-square approach reported here does not correct for questions in which the volunteering pool was homogeneous (e.g., only high-achieving males volunteer). In this type of situation, the teacher cannot fail to be consistent with student volunteering rates. As such, the chi-square approach is potentially biased toward the null hypothesis that observed teacher responsiveness will not differ from expected teacher responsiveness based on student volunteering rates. To rule out the possibility that this bias was responsible for the nonsignificant results obtained in the present investigation, a measure that corrects for this potentiality was developed:

$$z_{ij} = \frac{\sum_k C_{ijk} - \sum_k \frac{v_{ijk}}{y_{ij-}}}{\sqrt{\sum_k \frac{v_{ijk}}{v_{ij+}}\left(1 - \frac{v_{ijk}}{v_{ij+}}\right)}}. \quad (1)$$

Here, i indexes sex, j indexes achievement, and k indexes the observation or question. c_{ijk} represents teacher responsiveness (i.e., the sex and achievement level of the student called on for each question), whereas v_{ijk} represents, for each question, the proportion of student volunteers representing each sex and achievement level combination. Results obtained using this procedure were identical to those obtained using the chi-square approach. That is, even after controlling for those instances in which the volunteering pool was homogeneous, observed teacher responsiveness did not significantly differ from that which would be expected based on student volunteering rates. For ease of interpretation, the chi-square values are reported.

Because of marked differences in the relative number of boys and girls in each classroom (see Table 1), in the number of questions that were available for analysis across classrooms (see Table 2) and, as is shown later, differences in the volunteering rates of boys and girls, it was necessary to conduct separate analyses in each of the six classrooms so that expected values would reflect these discrepancies. For ease of interpretation, results are reported separately for Sex × Question Type analyses and for Sex × Achievement × Question Type analyses.

Effect sizes. To provide an overall estimate of the strength of the reported relationships, effect sizes were computed for outcomes obtained within each classroom and these effect sizes were subsequently combined to yield an overall effect size across the six classrooms. The effect-size estimate used in the present analysis was r, the population correlation between Variables X (e.g., sex) and Y (e.g., times called on). For all effect sizes, the computation of r was based on the following formula, where P approximates r (Glass, McGaw, & Smith, 1981):

$$P = \sqrt{\frac{\chi^2}{N + \chi^2}} \qquad (2)$$

In the current study, positive effect sizes indicated greater teacher responsiveness to or greater participation rates from boys. Conversely, negative effect sizes indicated greater involvement by girls in these interactions.

To test the degree to which the effect sizes were consistent across classrooms, we also determined the homogeneity of effect sizes. The homogeneity statistic used was Q, which has an approximate chi-square distribution with $k - 1$ degrees of freedom, where k is the number of effect sizes (Hedges, 1982; Hedges & Olkin, 1985). A significant Q statistic indicates that the effect sizes do not arise from a single population. Using the same computer program that was used to calculate the effect sizes for the current study (DSTAT; see Johnson, 1989), outliers may be removed until the Q statistic is no longer significant. The unique characteristics of the removed classrooms may then be examined to determine possible moderating variables. Because of the limited number of classrooms included in the current investigation, close examination of these moderators is generally beyond the scope of the present investigation. Potential contributors are, however, addressed in the discussion portion of the article.

Sex and Question Type Analyses

Teacher responsiveness relative to the proportion of male and female students. Our first and second hypotheses were that boys would be called on to answer more questions overall than girls and, perhaps, that boys would be asked to respond to different types of questions (i.e., a greater number of convergent and divergent questions) than girls. To test these hypotheses, observed teacher responsiveness was compared with expected teacher responsiveness based solely on the distribution of boys and girls in the class. *Teacher bias,* as it has typically been defined, is evident in classrooms in which teachers call on students of one sex more frequently than would be expected by chance given the relative proportion of boys and girls in the classroom of interest. For example, teacher bias in favor of boys for convergent and divergent questions would traditionally be concluded if, given a classroom in which 50% of the students were boys and 50% of the students were girls, boys were called on to answer 50% of all memory questions but 75% of all convergent and divergent questions.

Because the number of boys and girls was unequal in each of the observed classrooms except Classroom 5, it was necessary to calculate expected chi-square cell values that accounted for these discrepancies. This was accomplished by multiplying the total number of questions asked by the teacher within each question type by the proportion of boys and girls in the class, respectively. In Classroom 2, for example, 47% ($n = 14$) of the students were boys. A total of 46 memory questions were asked in this classroom. By multiplying 46 (the number of memory questions asked) by .47 (the proportion of boys in the class), we determined that, based on the number of boys in the class, boys would be expected to be called on to answer 21.62 (i.e., 47%) of all memory questions. Boys were, however, asked more memory questions than would have been expected by chance in this class: 32 of the 46 questions posed by their teacher (see Table 3).

Significant differences between observed and expected teacher responsiveness based on the relative proportion of boys and girls in the classroom were, in fact, found in three of the six observed classrooms (Classrooms 2, 3, and 6). Consistent with the hypotheses for the current investigation, boys were called on more often than would be expected in each of these classrooms. Comparisons of observed and expected values revealed similar, but nonsignificant, patterns in the remaining three classrooms as well. The conclusion that, across all six classrooms, boys were called on more often than expected given their numbers is supported by the effect-size calculation ($r = .35$, $p < .0001$). The homogeneity statistic was, however, also significant ($Q = 12.14$, $p < .05$). Removal of the largest outlier, Classroom 4, reduced the value of Q to a nonsignificant level ($Q = 6.86$, $p = .14$).

On the basis of follow-up chi-squares, no clear patterns in teacher responsiveness across question types were revealed. In those classrooms in which the overall chi-square analyses were significant, boys were called on significantly more often than would be expected for memory questions in three of the classrooms, for convergent questions in two of the classrooms, and for divergent questions in two of the classrooms. Given these inconsistencies, the data do not lend support for the hypothesis that boys are asked different questions (i.e., more convergent and divergent questions) than girls.

Student volunteering rates. Our third hypothesis was that teachers' tendency to call on boys more than girls in the science classroom would not be the result of an overall teacher bias in favor of boys but, rather, a reflection of a responsivity on the part of teachers to the heightened volunteering rates of the boys. As a first step in testing this hypothesis, the second set of analyses compared observed student volunteering rates with expected volunteering rates, again on the basis of the relative number of boys and girls in the participating classrooms.

Expected student volunteering rates were calculated in a manner similar to that for the teacher responsiveness analyses. That is, expected cell values were determined by multiplying the total number of student volunteers for each question type by the proportion of boys and girls in the class,

Table 3
Summary of Chi-Square Analyses Comparing Observed Teacher Responsiveness With Expected Teacher Responsiveness on the Basis of the Proportion of Boys and Girls in Each Class

Classroom and question type	n	Boys		Girls		χ^2
		Observed	Expected	Observed	Expected	
Classroom 1	28					1.35
Memory		17	15.66	12	13.34	0.25
Convergent		10	10.80	10	9.20	0.13
Divergent		13	10.80	7	9.20	0.97
Classroom 2	30					46.82****
Memory		32	21.62	14	24.38	9.40**
Convergent		57	33.37	14	37.63	31.57****
Divergent		37	27.73	22	31.27	5.85*
Classroom 3	30					16.76***
Memory		34	22.23	5	16.77	14.49****
Convergent		18	14.82	8	11.18	1.59
Divergent		8	9.69	9	7.31	0.68
Classroom 4	23					1.74
Memory		11	8.40	13	15.60	1.24
Convergent		15	12.95	22	24.05	0.50
Divergent		8	8.05	15	14.95	0.00
Classroom 5	26					4.91
Memory		12	8.50	5	8.50	2.79
Convergent		9	6.50	4	6.50	1.92
Divergent		3	2.50	2	2.50	0.20
Classroom 6	28					20.85****
Memory		15	9.46	7	12.54	5.69*
Convergent		10	6.02	4	7.98	4.62*
Divergent		17	9.46	5	12.54	10.53***

Note. All overall chi-squares are calculated with 3 *df*s. All follow-up chi-squares are calculated with 1 *df*.
$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$. $^{****}p < .0001$.

respectively. Again, we use Classroom 2 as an example. Here, there were 162 student volunteers for memory questions. In the classroom, 47% of students were boys. Multiplying 162 by .47, a value of 76.14 is obtained, which (given that the total number of student volunteers remains constant) is the number of volunteers we would have expected to be boys. The observed value for male volunteers, 108, far exceeds this expected value.

Results from these analyses are summarized in Table 4. In four of the six classrooms observed (Classrooms 2, 3, 5, and 6), there was, in fact, a significant difference between observed and expected volunteering rates. In each case, boys volunteered significantly more often than would be expected. A similar, but nonsignificant, pattern was also evident in Classroom 1. These findings were confirmed across the six classrooms by the effect size calculations ($r = .32$, $p < .0001$). The hypothesis of homogeneity of effect sizes was again, however, rejected ($Q = 71.94$, $p < .0001$). Removal of Classrooms 1 and 4 reduced this value to nonsignificant levels ($Q = .90$, $p = .83$). Again, there was no clear pattern of sex differences in student volunteering rates across question types.

It is interesting that only in Classroom 4 was a pattern in the opposite direction found, with girls volunteering more often than boys, particularly for divergent questions. The fact that the teacher in this classroom was female does not appear to account for the heightened volunteering among girls: Significant differences in volunteering rates in favor of boys were found in each of the other two classrooms in which the instructor was female.

Teacher responsiveness relative to student volunteering. The final set of Sex × Question Type analyses again involved comparing observed teacher responsiveness with expected teacher responsiveness. This time, however, expected teacher responsiveness was not based on the number of boys and girls in the class but, rather, on the distribution of male and female student volunteers in the class. Our hypothesis, again, was that although teachers may call on boys more often than girls in some classrooms, this is probably because teachers are responding to the heightened volunteering rates of the boys in these classrooms. To examine this hypothesis, expected cell values for teacher responsiveness were calculated by multiplying the proportion of male and female volunteers, respectively, by the total number of questions within each type. In Classroom 2, for example, 108 (or 67%) of the 162 volunteers for memory questions were boys. There were a total of 46 memory questions asked during the observed lessons for this classroom. Multiplying 46 by .67, a value of 30.64 is obtained. This is the number of times we would expect a boy to be called on, assuming that male–female volunteering rates remained constant. Boys in this class were, in fact, called on to answer 32 of the 46 memory questions asked: just one more question than would be expected based on their

Table 4
Summary of Chi-Square Analyses Comparing Observed Student Volunteering Rates With Expected Volunteering Rates on the Basis of the Proportion of Boys and Girls in Each Class

Classroom and question type	n	Boys		Girls		χ^2
		Observed	Expected	Observed	Expected	
Classroom 1	28					5.43
Memory		49	41.58	28	35.42	2.88
Convergent		30	31.86	29	27.14	0.24
Divergent		27	22.14	14	18.86	2.32
Classroom 2	30					119.61****
Memory		108	76.14	54	85.86	25.15****
Convergent		183	116.09	64	130.91	72.76****
Divergent		113	82.25	62	92.75	21.69****
Classroom 3	30					55.69****
Memory		114	73.35	15	55.47	51.80****
Convergent		63	53.58	31	40.42	3.85*
Divergent		28	27.36	20	20.64	0.03
Classroom 4	23					7.11
Memory		19	18.90	35	35.10	0.08
Convergent		30	35.00	70	65.00	1.09
Divergent		15	24.85	56	46.15	6.01**
Classroom 5	26					11.27**
Memory		27	19.50	12	19.50	5.77*
Convergent		15	10.00	5	10.00	5.00*
Divergent		5	4.00	3	4.00	0.50
Classroom 6	28					35.43****
Memory		20	12.90	10	17.10	6.86**
Convergent		20	17.20	20	22.80	0.80
Divergent		43	23.65	12	31.35	27.76****

Note. All overall chi-squares are calculated with 3 *dfs*. All follow-up chi-squares are calculated with 1 *df*.
$*p < .05$. $**p < .01$. $****p < .0001$.

heightened volunteering rates compared with girls in this class.

Confirming our predictions, in none of the participating classrooms did observed teacher responsiveness differ significantly from expected teacher responsiveness when we used student volunteering rates to calculate expected values (see Table 5). The degree of overlap was quite striking in five of the six classrooms, with p values ranging from .64 to .99. When we divided the number of times each student was called on by the total number of times that he or she volunteered, we found that girls were, in fact, slightly more likely than boys to be called on when they volunteered. Specifically, a girl was, on average, selected to answer a question 46% of the time when her hand was raised; a volunteering boy was selected in 44% of these instances. Moreover, a similar pattern of results was found for the 3 most responsive boys and girls in each classroom. Within this subset of students, boys and girls were called on to answer questions 42% and 41% of the time, respectively, when they volunteered. The remaining boys and girls were, in contrast, called on to answer questions 47% and 48% of the time, respectively, when they volunteered to answer a question.

Interestingly, only in one classroom (Classroom 4) was there any evidence of a pattern that was inconsistent with our predictions. In this classroom, which was the only classroom in which girls tended to volunteer more often than boys, there was a nonsignificant trend ($p = .12$) for the teacher to call on boys more often than would be expected based on their somewhat lower volunteering rates relative to girls'. The effect-size calculation across classrooms was, as anticipated, nonsignificant ($r = .05$, $p = .2661$). The homogeneity statistic was, moreover, nonsignificant ($Q = 7.056$, $p = .22$), indicating that the pattern of results found in Classroom 4 was not sufficiently deviant for the hypothesis of homogeneity of effect sizes to be rejected.

Sex, Question Type, and Achievement Analyses

Analyses identical to those performed above were conducted with relative student achievement (high vs. low) within each classroom considered. Because of small cell sizes that were due both to the numbers of questions asked by teachers and the volunteering rates of students in Classrooms 5 and 6, results are meaningfully interpretable only for the first 4 classrooms.

As might be expected, results were similar to those reported for the Sex × Question Type analyses. In both Classrooms 2 and 3, significant overall differences were found between observed teacher responsiveness and the pattern to be expected on the basis of the proportion of high- and low-achieving boys and girls in these classrooms, $\chi^2(9, N = 30) = 74.70$, $p < .001$, and χ^2 $(9, N = 30) =$

Table 5
Summary of Chi-Square Analyses Comparing Observed Teacher Responsiveness With Expected Teacher Responsiveness on the Basis of Boys' and Girls' Volunteering Rates

Classroom and question type	n	Boys		Girls		χ^2
		Observed	Expected	Observed	Expected	
Classroom 1	28					0.35
Memory		17	18.50	12	10.50	0.34
Convergent		10	10.16	10	9.84	0.01
Divergent		13	13.16	7	6.82	0.01
Classroom 2	30					1.69
Memory		32	30.64	14	15.36	0.18
Convergent		57	52.61	14	18.39	1.41
Divergent		37	38.17	22	20.89	0.10
Classroom 3	30					0.97
Memory		34	34.40	5	4.60	0.05
Convergent		18	17.71	8	8.32	0.02
Divergent		8	9.95	9	7.06	0.92
Classroom 4	23					5.76
Memory		11	8.42	13	15.58	1.21
Convergent		15	11.10	22	25.90	1.96
Divergent		8	4.85	15	18.15	2.59
Classroom 5	26					0.28
Memory		12	11.75	5	5.25	0.03
Convergent		9	9.79	4	3.21	0.25
Divergent		3	3.13	2	1.88	0.01
Classroom 6	28					0.04
Memory		15	14.61	7	7.39	0.03
Convergent		10	9.93	4	4.07	0.00
Divergent		17	17.16	5	4.84	0.01

Note. All overall chi-squares are calculated with 3 *df*s. All chi-square values are nonsignificant.

21.61, $p < .01$, respectively. In Classrooms 1, 2, and 3, observed student volunteering rates differed from expected rates based on this same criterion, χ^2 (9, $N = 28$) = 27.77, $p < .001$; χ^2 (9, $N = 30$) = 224.82, $p < .001$; and χ^2 (9, $N = 30$) = 69.37, $p < .001$, respectively.

Few consistent discriminatory patterns in teacher responsiveness or in student volunteering rates on the basis of relative student achievement were noted in the follow-up analyses. Likewise, no consistent patterns were found across question types. In some classrooms and at some question types, both high- and low-achieving boys were found to receive more teacher attention and to volunteer more often than would be expected based on their numbers in the classroom. In the same way, both low- and high-achieving girls were found, at some point, to receive less attention or to volunteer less often than would be expected. Interestingly, significant alternative patterns (i.e., indicating greater attention to or volunteering from girls or indicating lesser attention to or volunteering from boys) were never found, regardless of student achievement. This finding provides some evidence for the conclusion that, at least in this sample, observed differences in teacher responsiveness (on the basis of the proportion of boys and girls in class) and in student volunteering are more strongly governed by student sex than by an interaction of student sex and relative achievement level.

Most important to the current study is the finding that, as hypothesized, in none of these four classrooms did observed teacher responsiveness differ significantly from what would be expected based on the distribution of high- and low-achieving male and female volunteers within each classroom. Chi-square values ranged from 4.04 to 7.73 across the four classrooms. The degree of overlap was, again, striking, with p values ranging from .56 to .91. Clearly, teachers in these classrooms were calling on students at rates that were consistent with patterns of student volunteering. These findings were again confirmed by the effect-size calculations ($r = .08$, $p = .12$). Closer examination of the results, using relative achievement level as a potential moderating variable, indicated that teacher responsiveness did not differ significantly from expected levels based on student volunteering rates for either high ($r = .09$) or low ($r = .06$) achievers ($Q = .01$, $p = .79$).

Discussion

Often ignored in the literature on sex differences in classroom interactions is the role that students themselves may play in influencing the numbers and types of questions to which they are asked to respond (e.g., through volunteering to answer questions posed by their teacher). The purpose of this study was to address this gap in the current literature by examining rates of both teacher responsiveness and student participation (i.e., volunteering) in the classroom question-asking context. Although, as in previous research, we expected that teachers would direct more of their questions to their male than to their female students, we also anticipated that teachers would do so at rates that were consistent with the heightened volunteering rates of boys in these classrooms.

Replicating the findings of several previous studies (e.g., Becker, 1981; Good et al., 1973; Morse & Handley, 1985), we found that teachers in three of our six classrooms did call on boys significantly more often than would be expected on the basis of the relative proportion of boys to girls in the class. Patterns in this same direction were evident in each of the three remaining classrooms. Although several researchers (e.g., Becker, 1981; Good et al., 1973; Scantlebury & Kahle, 1993) have also found support for the hypothesis that boys are asked more higher level questions than are girls, we found evidence in support of the alternative claim that boys and girls are asked similar questions (see Hillman & Davenport, 1978; Jones & Wheatley, 1990). In the same way, we found that question type was not differentiated by student achievement level. As anticipated, when we calculated expected rates of teacher responsiveness based on the number of boys and girls in the volunteering pool from which the teacher selected a respondent, boys were no longer favored over girls. A similar pattern of results was evident when student sex and achievement level were considered together.

Further evidence that teachers were not demonstrating an overall bias in favor of boys in these classrooms can be derived from the result that boys and girls were equally likely to be called on when they volunteered. In the same way, neither male nor female target students (i.e., the 3 most responsive students of each sex in each classroom) were more likely than nontargets to be called on when their respective volunteering rates were taken into consideration. In fact, a pattern in the opposite direction emerged with nontargets somewhat more likely to be called on than targets when they volunteered. Therefore, not only were teachers calling on students at rates that were fairly consistent with their volunteering rates, but it appeared that teachers were, in fact, somewhat more responsive to students for whom levels of participation were reduced.

In only one classroom did the pattern of results differ (although not significantly so) from our prediction that rates of teacher responsiveness would be consistent with rates of student volunteering. In Classroom 4, girls volunteered to answer more questions than would be expected relative to boys. A likely feature of this classroom that may have permitted girls' heightened volunteering rates was their numerical dominance: 63% of the students in this class were girls. Furthermore, this is the only classroom in which there was a trend for teachers to respond to students at rates that were inconsistent with their volunteering rates. That is, girls were called on somewhat less often than might have been expected on the basis of their volunteering rates.

One possible interpretation of this finding is that the Classroom 4 teacher may have underestimated and, thus, not have fully compensated for her girls' increased responsiveness. Despite the fact that women are often stereotyped as loquacious, in a variety of contexts including the classroom context, it is men and boys who talk more (Spender, 1982). It is also plausible that the teacher in Classroom 4 may have been using questions as a form of discipline or a means to draw boys' attention to the front of the class (Eccles & Blumenfeld, 1985; Pintrich & Blumenfeld, 1985). In this classroom, the boys clustered toward the back of the room.

Our findings are critical in two respects. First, we provide a possible alternative explanation for why sex differences in teacher–student interactions continue to be found, even among these exemplary science teachers. Quite simply, teachers may be calling on boys more than girls not because they are biased (at least not in the traditional sense) in favor of boys but, rather, because they are selecting from a volunteering pool that is dominated by boys. Second, we imply from our findings that it may be necessary to reconsider the focus of recommendations to educators. If student volunteering remains a critical feature of teacher–student interactions, simply instructing a teacher to call on girls more often may be unreasonable when such a teacher is faced with a volunteering pool that is 70% male (as in Classroom 2). Efforts should be made to understand the reasons for girls' relative nonparticipation so that steps can be taken to encourage girls to become fuller participants in the classroom question-asking context. These efforts may be most critical in the elementary years. That is, if elementary school teachers hold different expectations for participation from their male and female students, particularly in subject domains such as science (Kelly, 1988; Shepardson & Pizzini, 1992), these expectations may play a critical role in socializing later sex differences in student participation.

Some researchers have suggested that one way in which teachers may alleviate the negative consequences of diminished volunteering rates among girls is to use teaching strategies that defend against male dominance. Eccles and Blumenfeld (1985), for example, have noted that some girls may feel threatened in classrooms characterized by a high reliance on participation through student volunteering. In contrast, it is believed that classrooms characterized by high levels of teacher-controlled questioning (i.e., direct questions) and more private student–teacher interactions may facilitate heightened participation among these students.

Whatever the recommendations, care must be taken to examine fully their implications. Recommendations to dedicate resources to understand the reasons for girls' relative nonparticipation and to encourage greater volunteering among girls may elicit fears that girls are being blamed for their lower participation rates. At the same time, recommendations to restructure the classroom to appear more "friendly" to girls may raise concerns that girls are being overprotected. The latter recommendation will also require sex-equitable distribution of direct questions and careful management of call outs, both of which favored boys (although at relatively moderate levels) in the current sample.

In addition to addressing sex-related differences in volunteering rates specifically, recommendations may be aimed more generally at creating equitable experiences in the classroom question-asking context for all students. Consistent with findings reported by Tobin and Gallagher (1987), we found that a very small percentage of students (i.e., 3 targets of each sex in each class) accounted for a large proportion of the total volunteering pool (53% and 58% for boys and girls, respectively). As might be expected given the group-level consistency between teacher responsiveness and student volunteering rates, these same target students were called on by their teachers to answer a disproportionate number of the open questions posed by their teachers (51% and 48% for boys and girls, respectively). Boys were much more likely than girls to be a part of this target group of students. It is not, however, the case that teacher–student interactions were dominated by a small percentage of students only among boys. This finding, in particular, suggests that increasing participation among nontarget students regardless of gender will be a critical feature of creating an equitable teacher–student interaction environment.

Several limitations of the current study should be noted. First, because the purpose of this study was to examine group-level differences in teacher responsivity to student volunteering rates, the quantity and quality of teacher responses to student answers were not evaluated. Sex-related differences in the amount and types of feedback that students receive have been reported in the empirical literature (e.g., Becker, 1981; Jones & Wheatley, 1990; Morse & Handley, 1985). Certainly, then, the impact that teacher feedback may have both on student volunteering rates and on attitudes toward classroom participation should be examined in future research.

A second limitation of the present study is that the participating teachers were, clearly, a small and select sample. All were deemed exemplary science teachers and, furthermore, were regarded as being sensitive to issues of gender equity in their teaching of science. Continued research will be necessary to generalize the finding that teachers are responding to student volunteering rates rather than to student sex or achievement level to other classrooms. Because the current investigation was limited to fifth- through eighth-grade science classrooms, additional research will be particularly helpful in elucidating whether the pattern of results remains similar across academic subject area (e.g., stereotypically feminine subject areas like English and reading) and grade level. Moreover, this work will be critical in providing evidence regarding the degree to which teachers do or do not respond to students at rates that are consistent with volunteering rates (in rare instances) when girls are the group to dominate in the question-asking context.

On the other hand, the select sample may be considered a strength of the current investigation insofar as it permitted us to demonstrate that *teacher bias* as it has been typically defined (i.e., calling on one group of students at rates that are inconsistent with their number in the classroom) is apparent even among these exemplary teachers. In so doing, our point that research must concern itself with the role that both teachers and students play in these interactions was highlighted. Focusing on the former while ignoring the latter may conceal the important notion that exemplary teachers, and perhaps most teachers, are remarkably accurate in responding to students at rates that are consistent with student volunteering rates.

In summary, the present investigation provides an alternative explanation for observed sex differences in teacher–student interactions in the classroom question-asking context. Although teachers do not appear to be the major source of the consolidation of sex-differentiated questioning patterns in at least the late-elementary and middle school science classrooms, the inequalities that remain cannot be ignored. Focusing on the role that both teachers and their students play in creating and maintaining observed sex differences in the teacher–student interaction context will be critical in alleviating the effects that nonparticipation may bring.

References

American Association of University Women. (1992). *How schools shortchange girls: The AAUW report*. New York: Marlowe.

Becker, J. R. (1981). Differential treatment of females and males in mathematics classes. *Journal for Research in Mathematics Education, 12*, 40–53.

Bell, R. Q. (1968). A reinterpretation of the direction of effects in studies of socialization. *Psychological Review, 75*, 81–95.

Beller, M., & Gafni, N. (1996). The 1991 international assessment of educational progress in mathematics and sciences: The gender differences perspective. *Journal of Educational Psychology, 88*, 365–377.

Blank, M., & White, S. J. (1986). Questions: A powerful but misused form of classroom exchange. *Topics in Language Disorders, 6*, 1–12.

Brophy, J. (1985). Interactions of male and female students with male and female teachers. In L. Wilkinson & C. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 115–142). New York: Academic Press.

Brophy, J., & Good, T. (1970). *Teacher–child dyadic interaction: A manual for coding classroom behavior.* Austin, TX. The Research and Development Center for Teacher Education.

Brophy, J., & Good, T. (1974). *Teacher–student relationships: Causes and consequences*. New York: Holt, Rinehart & Winston.

Carlsen, W. S. (1991). Questioning in the classrooms: A sociolinguistic perspective. *Review of Educational Research, 61*, 157–178.

Catsambis, S. (1995). Gender, race, ethnicity, and science education in the middle grades. *Journal of Research in Science Teaching, 32*, 243–257.

Cazden, C. B. (1986). Classroom discourse. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 432–463). New York: Macmillan.

Cherry, L. (1975). The preschool teacher–child dyad: Sex differences in verbal interaction. *Child Development, 46*, 532–535.

Eccles, J. S., & Blumenfeld, P. (1985). Classroom experiences and student gender: Are there differences and do they matter? In L. C. Wilkinson & C. B. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 79–114). New York: Academic Press.

Farrar, M. T. (1986). Teacher questions: The complexity of the cognitively simple. *Instructional Science, 15*, 89–107.

Fennema, E., & Peterson, P. (1985). Autonomous learning behavior: A possible explanation of gender-related differences in mathematics. In L. C. Wilkinson & C. B. Marret (Eds.), *Gender influences in classroom interaction* (pp. 17–35). New York: Academic Press.

French, P., & MacLure, M. (1981). Teachers' questions, pupils' answers: An investigation of questions and answers in the infant classroom. *First Language, 2*, 31–45.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.

Good, T. L., Sikes, J. N., & Brophy, J. E. (1973). Effects of teacher sex and student sex in classroom interaction. *Journal of Educational Psychology, 65*, 74–87.

Good, T. L., & Slavings, R. L. (1988). Male and female student question-asking behavior in elementary and secondary mathematics and language arts classes. *Journal of Research in Childhood Education, 3,* 5–23.

Guilford, J. P. (1979). *Cognitive psychology with a frame of reference.* San Diego, CA: Edits.

Hedges, L. V. (1982). Fitting categorical models to effect sizes from a series of experiments. *Journal of Educational Statistics, 6,* 107–128.

Hedges, L. V., & Olkin, L. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Hillman, S. B., & Davenport, G. G. (1978). Teacher–student interactions in desegregated schools. *Journal of Educational Psychology, 70,* 545–553.

Illinois State Board of Education. (1993). *Illinois Goal Assessment Program: Sample tests of Illinois goals in science.* Springfield, IL: Author.

Johnson, B. T. (1989). *DSTAT: Software for the meta-analytic review of research literatures.* Hillsdale, NJ: Erlbaum.

Jones, G. M., & Wheatley, J. (1990). Gender differences in teacher–student interactions in science classrooms. *Journal of Research in Science Teaching, 27,* 861–874.

Jones, L. R., Mullis, I. V., Raizen, S. A., Weiss, I. R., & Weston, E. A. (1992). *The 1990 science report card: NAEP's assessment of fourth, eighth, and twelfth graders.* Washington, DC: National Center for Education Statistics.

Jovanovic, J., & King, S. S. (in press). Boys and girls in the performance-based science classroom: Who's doing the performing? *American Educational Research Journal.*

Kahle, J. B., & Rennie, L. J. (1993). Ameliorating gender differences in attitudes about science: A cross-national study. *Journal of Science Education and Technology, 2,* 321–334.

Kelly, A. (1988). Sex stereotypes and school science: A three year follow-up. *Educational Studies, 14,* 151–163.

King, A. (1991). Effects of training in strategic questioning on children's problem-solving performance. *Journal of Educational Psychology, 83,* 307–317.

Leinhardt, G., Seewald, A. M., & Engel, M. (1979). Learning what's taught: Sex differences in instruction. *Journal of Educational Psychology, 71,* 432–439.

Lerner, R. M. (1982). Children and adolescents as producers of their own development. *Developmental Review, 2,* 342–370.

Lerner, R. M., & Busch-Rossnagel, N. A. (1981). Individuals as producers of their development: Conceptual and empirical bases. In R. M. Lerner & N. A. Busch-Rossnagel (Eds.), *Individuals as producers of their development: A life-span perspective* (pp. 1–36). San Diego, CA: Academic Press.

Licht, B. G., Stader, S. R., & Swenson, C. C. (1989). Children's achievement-related beliefs: Effects of academic area, sex, and achievement level. *Journal of Educational Research, 82,* 253–260.

Martin, V. L., & Pressley, M. (1991). Elaborative-interrogation effects depend on the nature of the question. *Journal of Educational Psychology, 83,* 113–119.

Morse, L. W., & Handley, H. M. (1985). Listening to adolescents: Gender differences in science classroom interaction. In L. C. Wilkinson & C. B. Marret (Eds.), *Gender influences in classroom interaction* (pp. 37–56). New York: Academic Press.

Parker, S. J. (1983). *Thinking about thinking.* East Lansing, MI: National Center for Research on Teacher Learning. (Eric Document Reproduction Service No. ED 270725)

Parsons, J. E., Kaczala, C. M., & Meece, J. L. (1982). Socialization of achievement attitudes and beliefs: Classroom influences. *Child Development, 53,* 322–339.

Perry, M., Vanderstoep, S. W., & Yu, S. L. (1993). Asking questions in first-grade mathematics classes: Potential influences on mathematical thought. *Journal of Educational Psychology, 85,* 31–40.

Pintrich, P. R., & Blumenfeld, P. C. (1985). Classroom experience and children's self-perceptions of ability, effort, and conduct. *Journal of Educational Psychology, 77,* 646–657.

Redfield, D. L., & Rousseau, E. W. (1981). A meta-analysis of experimental research on teacher questioning behavior. *Review of Educational Research, 51,* 237–245.

Sadker, M., & Sadker, D. (1982). *Sex equity handbook for schools.* New York: Longman.

Sadker, M., & Sadker, D. (1989). *The cost of sex bias in schools and society: The report card #1* (Report No. SO-021144). East Lansing, MI: National Center for Research on Teacher Learning. (ERIC Document Reproduction Service No. ED 329473)

Sadker, M., & Sadker, D. (1990). Confronting sexism in the college classroom. In S. L. Gabriel & I. Smithson (Eds.), *Gender in the classroom: power and pedagogy* (pp. 176–187). Urbana: University of Illinois Press.

Samson, G. E., Strykowski, B., Weinstein, T., & Walberg, H. J. (1987). The effects of teacher questioning levels on student achievement: A quantitative synthesis. *Journal of Educational Research, 80,* 290–295.

Scantlebury, K., & Kahle, J. B. (1993). The implementation of equitable teaching strategies by high school biology student teachers. *Journal of Research in Science Teaching, 30,* 537–545.

Serbin, L. A., O'Leary, K. D., Kent, R. N., & Tonick, I. J. (1973). A comparison of teacher response to the preacademic and problem behavior of boys and girls. *Child Development, 44,* 796–804.

Shepardson, D. P., & Pizzini, E. L. (1992). Gender bias in female elementary teacher's perceptions of the scientific ability of students. *Science Education, 76,* 147–153.

Sikes, J. N. (1971). *Differential behavior of male and female teachers with male and female students. Dissertation Abstracts International, 33,* 217A. (University Microfilms No. 72-19670)

Spender, D. (1982). *Invisible women: The schooling scandal.* London: Writers and Readers Publishing Cooperative.

Stake, J. E., & Katz, J. F. (1982). Teacher–pupil relationships in the elementary school classroom: Teacher–gender and pupil–gender differences. *American Educational Research Journal, 19,* 465–471.

Stallings, J. (1985). School, classroom, and home influences on women's decisions to enroll in advanced mathematics courses. In S. F. Chapman, L. R. Brush, & D. M. Wilson (Eds.), *Women and mathematics: Balancing the equation* (pp. 199–233). Hillsdale, NJ: Erlbaum.

Stitt, B. A. (1988). *Building gender fairness in schools.* Carbondale: Southern Illinois University Press.

Tobin, K., & Gallagher, J. J. (1987). The role of target students in the science classroom. *Journal of Research in Science Teaching, 24,* 61–75.

Wang, M. C., & Weinstein, W. J. (1980). Teacher expectations and student learning. In Fyans, L. J. (Ed.), *Achievement motivation: Recent trends in theory and research* (pp. 417–444). New York: Plenum Press.

Weinburgh, M. (1995). Gender differences in student attitudes toward science: A meta-analysis of the literature from 1970 to 1991. *Journal of Research in Science Teaching, 32,* 387–398.

Winne, P. H. (1979). Experiments relating teachers' use of higher cognitive questions to student achievement. *Review of Educational Research, 49,* 13–50.

Wong, E. D. (1991). Beyond the question/nonquestion alternative in classroom discussion. *Journal of Educational Psychology, 83,* 159–162.

Received September 12, 1997
Revision received February 3, 1998
Accepted February 15, 1998 ■