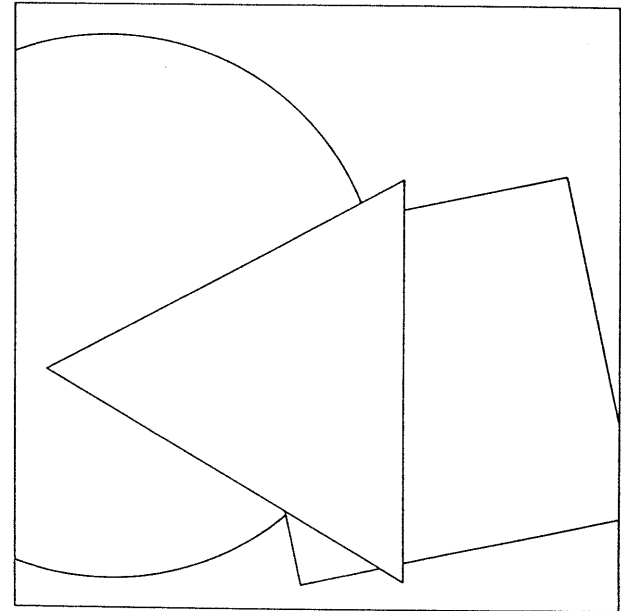


# STEREOTYPE ACCURACY

TOWARD APPRECIATING GROUP  
DIFFERENCES



Edited by Yueh-Ting Lee, Lee J. Jussim,  
and Clark R. McCauley

American Psychological Association  
Washington, DC

1995

## Are Teacher Expectations Biased by Students' Gender, Social Class, or Ethnicity?

Lee J. Jussim and Jacquelynne Eccles

For some time now, social science perspectives on stereotypes have generally been divided between those emphasizing error, bias, and inaccuracy (e.g., American Psychological Association [APA], 1991; Fiske & Taylor, 1991; Jones, 1990; Marger, 1991; Miller & Turnbull, 1986) and those arguing that the error/bias/inaccuracy issue is a largely unanswered empirical question (Campbell, 1967; Judd & Park, 1993; Jussim, 1990, 1991; Jussim, McCauley, & Lee, chapter 1, this volume; McCauley, Stitt, & Segal, 1980). There are many ways in which stereotypes may go awry. First, they may lead people to errors and biases in their beliefs about social groups. Those beliefs may be factually incorrect, they may exaggerate real differences, or they may lead people to perceive outgroups as all alike. Although these issues are extremely important, they are not the focus of this chapter.

---

We gratefully acknowledge the valuable suggestions regarding this research provided by attendees of the 1994 Conference on Stereotype Accuracy, especially those of Richard Ashmore, Clark R. McCauley, and Victor Ottati. Preparation of this chapter was supported in part by National Institute on Child Health and Human Development Grant 1 R29 HD28401-01A1 to Lee J. Jussim. Correspondence regarding this chapter may be sent to Lee Jussim at Department of Psychology, Rutgers University, New Brunswick, New Jersey 08903 (phone: 908-932-2070; fax: 908-932-0036).

Presumably, however, erroneous stereotypes are a social problem primarily if they lead to biases and discrimination (if some people hold inaccurate social beliefs, but do not act any differently than others who hold accurate social beliefs, inaccuracy is not a problem). Inaccuracy becomes a problem when perceivers treat or evaluate one group differently than another as a result of that inaccuracy. Furthermore, many social psychologists believe both that stereotypes are frequently inaccurate and that they lead to all sorts of biases—consequently, they are frequently accused of being the cognitive culprits in prejudice and discrimination (e.g., Fiske & Taylor, 1991; Hamilton, Sherman, & Ruvolo, 1990). Others, however, have argued that the empirical evidence supporting the conclusion that stereotypes are generally inaccurate and lead to biases and discrimination is actually sparse, weak, and equivocal (see reviews by Jussim, 1990, 1991; Jussim et al., chapter 1, this volume; McCauley et al., 1980). Therefore, this chapter focuses on the role of stereotypes in leading to errors and biases in a context of critical importance for issues of justice, fairness, and equality of opportunity: education. Specifically, this chapter addresses whether teachers' stereotypes lead them to evaluate students from different sex, social class, and ethnic groups differently, when they do not deserve to be evaluated differently. This question goes to the heart of some of the alleged problems with stereotypes.

### PROCESS AND CONTENT IN RESEARCH ON STEREOTYPES

The overwhelming majority of social psychological research on stereotypes has been experimental laboratory studies. This research has several important merits. Tightly controlled studies aptly highlight some of the social and psychological processes relating stereotypes to person perception (e.g., Bodenhausen, 1988; Darley & Gross, 1983; Fiske & Neuberg, 1990; Krueger & Rothbart, 1988; Linville, 1982; Locksley, Borgida, Brekke, & Hepburn, 1980).

The experimental laboratory studies, however, also suffer several im-

portant limitations. They often use artificial or impoverished social stimuli (see, e.g., studies cited in APA, 1991, and reviews by Funder, chapter 6, this volume; Jussim, 1990, 1991, 1993). Perceivers often do not engage in face-to-face interactions with targets at all; they make judgments based on written descriptions of targets, slides, videotapes, and so on. Even when they do actually engage in a face-to-face interaction, it is usually with a stranger, for a period of an hour or less. And, of course, the laboratory studies primarily use college students as research subjects.

All of these factors may limit the generalizability of the findings from the experimental laboratory studies. However, the laboratory studies also suffer an extremely important conceptual or theoretical limitation. Studies that focus exclusively on identifying social-cognitive *processes* involved in stereotyping are completely incapable of drawing inferences about the accuracy of the *content* of stereotypes. For example, showing that categorization leads people to evaluate one group differently than they evaluate another group provides no information about whether the evaluation of either group is correct. Therefore, the implications of much of social psychology's knowledge base for understanding the accuracy of social stereotypes under naturalistic conditions are not clear.

Identifying accuracy or inaccuracy in the content and use of social stereotypes can be accomplished only when (a) the targets are real people with real attributes (as opposed to artificially created social stimuli), (b) there is some means of measuring those attributes (a criterion), and (c) perceivers' judgments are compared with the criterion.

The research described in this chapter was performed to help begin redressing this limitation to research on stereotypes by studying naturally occurring person perception and by comparing those perceptions to clear criteria. The first study addressed accuracy by comparing teacher perceptions of performance, talent, and effort differences among students from differing sex, socioeconomic, or ethnic groups to actual differences among those students. The second study examined the processes leading to accuracy and inaccuracy in teachers' perceptions of students from the differing groups.

## TEACHER EXPECTATIONS

There are few contexts more important for investigating stereotypes than teachers' expectations for their students. Ever since Rosenthal and Jacobson's (1968) seminal and controversial (e.g., Elashoff & Snow, 1971) Pygmalion study, writers in both scholarly journals and the popular press have implicated teacher expectations as a major perpetrator of injustices and inequalities based on ethnicity, social class, and sex (see Wineburg, 1987, for a review). In this chapter, we present evidence suggesting that such claims present a greatly oversimplified picture of the role of teacher expectations in perpetuating social inequalities. This evidence will convey two main points. First, teachers generally perceive only small differences among social groups (e.g., ethnic groups, social class groups, and sex groups), that is, stereotypes do not seem to be a powerful influence on their expectations. Second, many of the differences that they do perceive are reasonably accurate; many (though not all) of the differences they perceive among different groups correspond to preexisting objective differences among those groups.

Perhaps the most comprehensive analysis to date of the role of stereotypes in the development of teacher expectations remains Dusek and Joseph's (1983) meta-analysis, which showed that teachers perceived moderate differences between students based on social class ( $r = .23$ ) and little difference based on student sex or ethnicity ( $r_s = .04$  to  $.10$ ). However, the relevance of these findings to naturally occurring teacher-student interactions is not clear. Most studies included in Dusek and Joseph's meta-analysis were experiments that suffered from two important limitations: (a) Targets were fictitious manipulations (they were not real students), and (b) although perceivers had access to some information about students, they generally had no opportunity to interact with students or observe their achievement over an extended period (as do real teachers interacting in real classrooms with real students). Therefore, whether in-service teachers perceive differences among students comparable to those obtained in Dusek and Joseph's (1983) meta-analysis is unknown.

## STUDY 1

### Three Main Research Questions

Study 1 addressed three main questions: (a) Do teachers perceive sex, social class, or ethnic differences in performance, talent, and effort? (b) How accurate are the differences (or lack of differences) teachers perceive among students from different sex, social class, and ethnic groups? (c) Do sex, social class, and ethnic stereotypes lead to biases and errors in teachers' perceptions of students?

The strategy for addressing these questions was straightforward. First, we identified whether teachers perceived performance, talent, and effort differences among students from the different demographic groups. Next, we compared the differing groups on measures of performance, talent, and effort. We concluded that teachers were accurate when the size of the difference they perceived approximately corresponded to the size of the actual difference among students. Teachers' perceptions were inaccurate when the differences they perceived among students from the different groups substantially deviated from the actual differences. They could be inaccurate in either of two directions: (a) They might overestimate differences among groups (in the extreme, they might see a difference where none existed), or (b) they might underestimate differences between groups (in the extreme, they might perceive no difference when one existed).

### The Data

This study was based on the Michigan Study of Adolescent Life Transitions Project (Eccles, 1988), which assessed a variety of social, psychological, demographic, and achievement-related variables in a sample that included about 100 teachers and 2,600 students in sixth-grade math classes. Three teacher expectation variables were assessed in early October of sixth grade: teacher perceptions of students' performance, talent, and effort at math. Student motivation, which was also assessed in early October (just before the assessment of teacher perceptions), included self-concept of math ability and self-perceptions of effort and time spent on math home-

work. All measures were reliable and valid (for more detail, see Eccles, 1988; Eccles [Parsons], Adler, & Meece, 1984; Jussim, 1987, 1989; Jussim & Eccles, 1992; Parsons, 1980).

Final marks in fifth-grade math classes were the primary measure of performance. Scores on standardized achievement tests taken in late fifth or early sixth grade were the primary measure of talent. Although both measures are imperfect, we believe that they provide reasonable criteria with which to compare teacher perceptions. Grades are imperfect because they may reflect not only performance, but also neatness, assignment completion, cooperativeness, and teacher bias. Standardized tests are imperfect because in addition to underlying competencies, they may also reflect motivation, illness, and so on. Despite these imperfections, grades primarily represent the quality of students' performance over the course of the school year. If this were not true, the correlation between grades and standardized test scores (which are not influenced by neatness, cooperativeness, or teacher bias) would not be so high (e.g., Jussim, 1987; Jussim & Eccles, 1992). Also, standardized tests are intended to assess students' enduring competencies, knowledge, and skills, and in general, they are usually quite successful at doing so (e.g., Anastasi, 1982).

## Results

### *Student Sex*

These analyses were based on 942 girls (coded as 1) and 847 boys (coded as 2). This was the subsample that had valid data on all variables necessary for analyses involving student sex. Did teachers perceive differences between boys and girls? They did, albeit small ones. Teachers perceived girls as performing slightly more highly ( $r = -.08, p < .001$ ) and as trying harder ( $r = -.16, p < .001$ ). They perceived no difference in boys' and girls' talent ( $r = .02$ ).

Were these perceptions accurate? For performance and talent, the answer is yes. Girls had performed slightly higher than did boys in fifth-grade math classes ( $r = -.07, p < .01$ ), a real but small difference that corresponded closely to the real but small perceived difference in performance.

Similarly, there was no sex difference in standardized test scores ( $r = .00$ ), which corresponded with teachers' perceptions of no talent difference.

There was no evidence in these data that teacher perceptions of sex differences in effort were accurate. Boys and girls reported exerting the same amounts of effort ( $r = .00$ ) and spending the same amount of time on homework ( $r = -.03$ ). Self-concept of ability was considered a motivational variable because of its crucial role in leading to effort and persistence, according to several motivational theories (e.g., Bandura, 1977; Eccles & Wigfield, 1985; Weiner, 1979). However, boys actually had slightly higher self-concepts of math ability than did girls ( $r = .09, p < .001$ ).

Were teachers biased by students' sex? For performance and talent, the answer is no; for effort, the answer is yes. Were teachers biased against girls? No; if anything, they seemed biased in favor of girls. They evaluated girls as trying harder than boys, even though boys and girls claimed to be working equally hard and even though boys had higher self-concepts of ability.

### *Social Class*

These analyses assessed whether teachers perceived differences among students from differing socioeconomic backgrounds. To address this question, we obtained the multiple correlation of parental education and family income with each of the three teacher-perception variables. Parental education (for 98% of the students, this was the mother's education) and family income information was available for 1,066 students.

Did teachers perceive social class differences in performance and talent? They did. Teachers perceived students from higher social class backgrounds as performing more highly ( $R = .21$ ) and as more talented ( $R = .26$ , both  $ps < .01$ ). Were there real social class differences? There were. Parental income and education correlated with fifth-grade final grades ( $R = .27$ ) and previous standardized achievement test scores ( $R = .31$ , both  $ps < .001$ ).

Did teachers perceive social class differences in effort? They did. Teachers saw students from higher social class backgrounds as trying harder ( $R = .18, p < .01$ ). Were there real social class differences in effort? Although there were no social class differences in self-reported effort or time spent on homework (both  $Rs < .05, ns$ ), students from higher social class

backgrounds had higher self-concepts of math ability ( $R = .15, p < .01$ ). Thus, teacher perceptions of effort corresponded reasonably well with student social class differences in self-concept of ability.

Overall, therefore, these results provided little evidence that teachers were biased by students' social class. There was no evidence at all that teachers were biased against students from lower socioeconomic backgrounds.

### *Ethnicity*

These analyses focused on teacher perceptions of African American students and White students. There were too few students from other ethnic backgrounds to warrant performing analyses.

Did teachers perceive differences between African American students and White students? Answering this question turns out to be more difficult than it may seem, primarily because the continuing patterns of residential segregation in the United States were largely reflected in these data. Specifically, 10 school districts had predominantly White students. None had fewer than 88% White students, and as a group, 95% of the students in these 10 districts were White.

One district was integrated (63% White, 34% African American, and 3% other). One district included predominantly (93%) African American students. Because of these differences between districts, we performed two separate sets of analyses. The first set of analyses examined teacher perceptions and student differences among the White students in the White segregated districts, the African American students in the White segregated districts, and the African American students in the African American segregated districts. There were only four White students in the African American segregated district. Because this number is too small to permit meaningful analyses, data on this group were not included in any of the results we report below. The second set of analyses examined teacher perceptions of and student differences between the African American students and White students in the integrated district.

We report means, rather than correlations for analyses focusing on the segregated districts, to keep clear differences between teacher perceptions

of African American students in the segregated White versus segregated African American districts. Also, because there were virtually no White students in the segregated African American district, it would be impossible to compute correlations involving ethnicity in that district. Because of the large sample size, very small and trivial differences among groups sometimes yielded statistically significant  $F$  values. Therefore, when results reach statistical significance, we also report the effect size ( $etas$ ), and we used Scheffé's method (which is conservative) for testing post hoc comparisons among groups.

This first set of analyses actually combine two different sets of comparisons. Ethnic differences within the White segregated districts involve teacher ratings of different students in their classes. The questions here were, "In the White segregated district, do teachers perceive ethnic differences? And how well do their perceptions correspond to actual ethnic differences?" However, comparisons involving the segregated African American students involve *between-districts* comparisons. The questions here were, "Do teachers in the segregated African-American district view their students differently than teachers view African-American and White students in the White segregated district? And are there differences between students in the segregated African-American district and either the White or African-American students in the segregated White district?" Although these questions are quite different from one another, both are important. Therefore, the first set of analyses addressed these questions.

*The segregated districts.* In the segregated districts, none of the differences in teachers' perceptions of African American versus White students were statistically significant (all  $F$ s  $< 2.5$ , all  $p$ s  $> .05$ ). Teachers perceived little difference in the performance, talent, and effort between African American students and White students.

Were teacher perceptions of no performance or talent differences justified? The differences between the African American and White students in the segregated districts were significant for both standardized test scores,  $F(2, 2034) = 35.75, p < .001, \eta = .18$ , and final grades,  $F(2, 1947) = 11.88, p < .001, \eta = .11$ . Table 1 contains the mean previous grades and standardized test scores for students in the segregated districts and shows that

Table 1

Were There Ethnic Differences in Achievement in the Segregated Districts of Study 1?

Measure	White students in White districts	African American students in White districts	African American students in the African American district
Standardized test scores	61 <sub>a</sub>	52 <sub>a</sub>	38 <sub>b</sub>
<i>n</i>	1,907	39	91
Fifth grade final marks	B-/B <sub>a</sub>	C+/B- <sub>a,b</sub>	C+
<i>n</i>	1,865	26	59

NOTE: Standardized test scores listed in percentile ranks. Across rows, means that share at least one subscript are not significantly different at  $p < .05$ .

teacher perceptions were partially justified. In the segregated White districts, although there was a slight tendency for the White students to perform more highly than did African American students, neither the standardized test score differences nor the grade differences were statistically significant (all  $t_s < 1.4$ , all  $p_s > .1$ ). Therefore, teachers were justified in perceiving few differences between African American students and White students in the White segregated districts.

However, Table 1 also shows that the African American students in the segregated African American district performed significantly more poorly in terms of both standardized test scores and previous grades than did White students (both  $t_s > 3$ , both  $p_s < .01$ ). Therefore, teachers were not justified in evaluating the performance of the African American students in the segregated African American district as favorably as teachers judged the performance of the White students in the segregated White districts.

Teachers were also reasonably justified in perceiving no differences in the effort exerted by the different groups of students. The differences among students on the three motivation variables reached statistical significance for time spent on homework,  $F(2, 2383) = 4.68, p < .01$ ; self-concept of math ability,  $F(2, 2383) = 4.58, p < .02$ ; and marginal significance for self-perceptions of effort,  $F(2, 2388) = 2.81, p < .07$ . However, only one of the

post hoc comparisons was significant: African American students in the segregated African American district had higher self-concepts of math ability than had the White students in the White segregated district ( $t = 2.11, p < .05$ ). Furthermore, all of the *etas* were below .07, indicating that although statistically significant, the differences were minor.

*The integrated district.* Did teachers perceive the 22 African American students differently than they perceived the 40 White students in the integrated district? They did. Teachers perceived White students as performing more highly ( $r = -.27, p < .05$ ), as more talented ( $r = -.26, p < .05$ ), and as exerting more effort, although this last difference did not reach statistical significance ( $r = -.20, p = .12$ ).

Were these perceptions justified? The African American students did have lower fifth-grade marks than did White students in this district (C/C+ vs. B-,  $r = -.21$ ). However, because not all of the students in sixth grade attended this district in fifth grade, this difference was based on only 32 White students and 14 African American students, and it was not statistically significant ( $p = .16$ ). It was, however, of about the same magnitude as the differences that teachers perceived. Unfortunately, no standardized test had been given in this district in fifth grade.

Did teacher perceptions of effort differences correspond to ethnic differences in the motivation variables? African American students and White students in this district all claimed to be exerting about the same amount of effort and spending about the same amount of time on homework; all had similar self-concepts of ability (all  $r_s < .07$ , all  $p_s > .6$ ). Thus, there was a slight, but nonsignificant, tendency for teachers to perceive White students as trying harder, and no evidence in the student effort and motivation variables to suggest any ethnic difference in effort.

## Discussion

### *Teacher Expectations and Stereotypes: Preliminary Conclusions*

Some answers to the three questions guiding this first study are now available. Did teachers perceive differences between boys and girls, middle-class

and lower-class students, and African American and White students? They did. They perceived girls as performing slightly higher and as trying harder than boys, but they also evaluated their natural talent at math as similar. Teachers also viewed the performance, talent, and effort of middle-class students more favorably than those of lower-class students.

The results regarding ethnicity were mixed. Teachers perceived no differences among African American students and White students in the segregated districts. In the integrated districts, however, they evaluated the White students more favorably.

How accurate were the differences and similarities that teachers perceived? For the most part, they were accurate. For all three demographic groups, teachers' perceptions of students' performance or talent generally corresponded quite closely to the actual differences or similarities in those groups' previous grades and standardized test scores. There was only one exception to this pattern: Teachers rated African American students in the predominantly African American district as favorably as other students, when, in fact, both their grades and standardized test scores were not as high as those of other students.

The pattern for teacher perceptions of effort was more mixed, providing evidence of both accuracy and inaccuracy. Teachers believed that girls tried harder than boys, but there was no difference between the sexes on the effort measures, and boys felt they had more math ability than girls felt they had. Thus, there was no evidence of accuracy here. In contrast, however, teacher beliefs that middle-class students tried harder than did lower-class students corresponded closely to student social class differences in self-concept of math ability. This belief, therefore, seems to have been reasonably accurate. Similarly, in both the integrated and segregated school districts, the teachers perceived few differences in the effort of African American versus White students. These perceptions were also reasonably accurate; few differences emerged on either the effort measures or on self-concept of ability.

Were teachers biased by student sex, class, or ethnicity? Sometimes it seemed that they were. However, there was no evidence that teachers were biased against girls, students from lower-class backgrounds, or African American students. Instead, teacher perceptions of effort were biased in

favor of girls, and their perceptions of performance and talent were biased in favor of the African American students in the predominantly African American district. They also perceived slightly smaller social class differences than really existed. Even these biases, however, were relatively modest. Overall, therefore, these data provide little evidence of pervasive or powerful biases based on sex, social class, or ethnicity.

#### *Why Was There So Little Evidence of Bias?*

*Social desirability.* Perhaps teachers attempted to appear socially desirable in providing favorable evaluations of girls and African American students. This seems unlikely for several methodological and conceptual reasons. First, teachers were simply asked to evaluate the students in their classes; the role of student demographics was never mentioned. Thus, issues involving stereotypes should not have been particularly salient. Second, if the social desirability explanation were true for ethnicity, there should have been more reverse bias in classes where teachers had both African American students and White students than in classes where there were only African American students. This is because interethnic comparisons should be far more salient in the mixed classes. In fact, however, the pattern we obtained was the opposite: There was no evidence of reverse bias in the mixed classes, whereas there was some bias in the segregated African American classes.

The main reverse bias for sex involved effort: Teachers perceived girls as trying harder than boys. Other analyses based on these same data, however, showed that teachers take their effort perceptions quite seriously; teachers provide higher grades to students believed to be trying harder (Jussim, 1989; Jussim & Eccles, 1992). This does not seem to be the behavior of teachers simply trying to act in a socially desirable manner. Furthermore, even the current data could be construed as providing at least a hint of bias against girls. That is, even though girls were perceived as performing more highly than boys, they were seen as "merely" equally talented, but as trying harder.

Another version of the social desirability explanation suggests that teachers may be more reluctant to rely on stereotypes than they were 20 or 30 years ago. That is, modern teacher training may often include dis-



cussion of self-fulfilling prophecy in general, or at least the original Rosenthal and Jacobson (1968) study, thereby reducing teachers' susceptibility to basing their expectations on erroneous, superficial information. In fact, however, there is no evidence that teachers' beliefs in demographic differences have changed much over the past 30 years (see Jussim, Madon, & Chatman, 1994, for a review); and self-fulfilling prophecies effect sizes have remained remarkably stable since the original Rosenthal and Jacobson study (see Jussim & Eccles, 1995, for a review).

*Change of standards.* Perhaps, however, the slight positive bias in favor of African American students in the segregated district resulted from teachers using differing standards for their students than were used in the other districts. When evaluating students on a subjective rating scale and when faced with overall lower levels of performance, perhaps a teacher's subjective rating of a student as meriting a 3 in performance in the segregated African American district reflects a somewhat lower level of actual performance than a 3 typically reflects in the White segregated districts (see, e.g., Biernat & Manis, 1994). Thus, even one of the results that looked like bias, teachers rating the students in the segregated African American district as highly as teachers rated students in the segregated White districts, may be more apparent than real.

*They held no stereotypes.* The simplest explanation for the lack of bias is that teachers held no negative stereotypes about girls, students from lower social class backgrounds, or African American students. If they held no stereotype, then there would be no stereotype to bias their perceptions. Unfortunately, this possibility cannot be tested directly, because teachers' social stereotypes were not assessed.

However, this explanation seems highly implausible for at least two reasons. First, nearly all of the teachers in the current study would need to have no such stereotypes. If even a substantial minority held (and used) stereotypes, there still should have been some evidence of bias. Furthermore, abundant research in the social sciences attests to the widespread existence and importance of these stereotypes (e.g., APA, 1991; Darley & Gross, 1983; Dusek & Joseph, 1983; Fiske & Taylor, 1991; Jones, 1990;

Marger, 1991). Thus, the likelihood that this sample of teachers was unique in that virtually none held stereotypes seems vanishingly small.

Second, some researchers have argued that one does not need even to subscribe to a stereotype for that stereotype to influence social perception (Devine, 1989; Sedikides & Skowronski, 1991). Mere knowledge of a cultural stereotype (regardless of whether one accepts it oneself), they argue, is sometimes sufficient to produce biases. Thus, one would need to argue not only that nearly our entire sample of teachers did not subscribe to stereotypes, but that they were all oblivious to them. This, too, seems highly implausible.

*Teachers did not use their stereotypes.* Another explanation could be that teachers did hold stereotypes regarding these groups, but did not use them in evaluating students. Research in education and social psychology suggests considerable plausibility for this explanation. Research on classrooms has consistently demonstrated considerable accuracy in teacher perceptions of students (Jussim, 1989; Jussim & Eccles, 1992; West & Anderson, 1976; Williams, 1976; see Brophy, 1983; Brophy & Good, 1974; Jussim, 1991, for reviews). Teachers generally judge students far more on the basis of their achievement and motivation than on teachers' own social stereotypes. And teachers are not the least bit unique. Whether individual targets are men and women, upper class and lower class, African American and White perceivers generally judge them far more on the basis of their personal characteristics than on their membership in these social groups. This occurs both in laboratory studies and in naturalistic studies (see Jussim, 1990, 1993, for reviews), and this possibility was directly tested in Study 2.

*Teachers used accurate stereotypes.* Another explanation could be that teachers do hold stereotypes regarding these groups, and they did, at least partially, rely on those stereotypes when judging students. Then why was there so little evidence of bias? If teachers relied on accurate stereotypes, they would have little or no tendency to exaggerate differences among the groups of students. This possibility could be tested with these data, and such a test is provided in the next study.

## STUDY 2

## Were Teachers Relying on Accurate Stereotypes?

Study 1 showed that teachers' perceptions of differences between students in the various groups were mostly accurate. In Study 2, additional analyses using the same data and teacher-student samples as in Study 1 were performed to compare the "teachers did not use their stereotypes" and "teachers used accurate stereotypes" alternative explanations for the lack of bias in Study 1. The question addressed in Study 2 was, Did relying on an accurate stereotype facilitate accuracy in teacher perceptions? Addressing this question requires answering two subquestions: (a) Did teachers rely on stereotypes when judging students, and (b) if so, did relying on stereotypes enhance or undermine their accuracy? Thus, whereas Study 1 focused exclusively on issues of content (e.g., were teacher perceptions of students from different groups accurate?), Study 2 focused on issues of process.

How can one discover if teachers relied on stereotypes when stereotypes were not assessed? One can do so indirectly, using the methods first developed in experimental social psychological laboratory studies of stereotypes and person perception. The prototypical and classic studies in this area involved no assessment of stereotypes. Instead, social psychological studies of the role of stereotypes in person perception typically manipulate targets' social group membership, hold constant or manipulate information about targets' personal characteristics, and assess whether perceivers judge targets from one group differently than targets from another group (e.g., Bodenhausen, 1988; Darley & Gross, 1983; Fiske & Neuberg, 1990; Krueger & Rothbart, 1988; Linville, 1982; Locksley et al., 1980; see reviews by Darley & Fazio, 1980; Fiske & Taylor, 1991; Hamilton et al., 1990; Jussim, 1990). If perceivers judge targets from different groups differently (holding constant targets' behavior or attributes), perceivers are assumed to be relying on their stereotypes when judging targets.

This is the strategy we used for identifying whether teachers relied on stereotypes in evaluating their students. Analyses assessed whether teachers perceived differences based on student sex, class, or ethnicity, when

holding constant statistically students' achievement and motivation. Specifically, we performed a series of regressions in which students' performance and motivation, and their social group memberships, predicted teacher perceptions. Operationally, therefore, the "relied on stereotypes" hypothesis was that teacher perceptions would be based on student group membership, even after controlling for student performance and motivation.

## Results

*Sex Stereotypes*

Three regressions were performed, in which student sex, previous grades, standardized test scores, self-concept of math ability, time spent on homework, and self-perceptions of effort predicted each of the three teacher-perception variables. The main questions here were (a) Would student sex predict teacher perceptions, independent of the other variables, (b) if so, did the student sex effect enhance or undermine accuracy, and (c) to what extent did teachers rely on their sex stereotypes versus individuating information (previous achievement and motivation)?

Table 2 summarizes the results from these analyses. These results showed that teachers seemed to be relying on an accurate stereotype when judging students' performance. The *beta* relating student sex to teacher perceptions of performance was  $-.09$ , which closely corresponded to the small sex differences in grades of  $-.07$  (found in Study 1). Although teachers also judged students on the basis of their performance, doing so was not the main source of the correlation between teacher perceptions and student sex. The effect of student sex on teacher perceptions ( $-.09$ ) accounted for most of the correlation between sex and teacher perceptions ( $-.07$ ). This means that teachers apparently stereotyped girls as performing at a higher level than boys, independent of the actual sex differences in performance. However, the extent to which they did so corresponded reasonably well with the actual small sex difference in performance.

Results for teacher perceptions of talent provided no evidence of

**Table 2**

**Did Teachers Rely on Students' Sex, Independent of Students' Achievement and Motivation in Study 2?**

Predictor	Teacher's perception		
	Performance	Talent	Effort
Student sex	-.09*	.02	-.16*
Fifth-grade final marks	.23*	.21*	.22*
Standardized test scores	.36*	.42*	.25*
Self-concept of math ability	.22*	.18*	.15*
Effort self-perceptions	.05*	.00	.11*
Time spent on homework	-.06*	-.05*	-.02
<i>R</i> <sup>2</sup>	.47*	.47*	.32*

NOTE: All entries are standardized regression coefficients. *N* = 1,789 (942 girls; 847 boys). \**p* < .01.89 (942 girls and 847 boys)

teachers relying on a stereotype. The *beta* relating student sex to teacher perceptions of talent was .02 (*ns*), corresponding closely to a .00 correlation of student sex with standardized test scores.

Results for teacher perceptions of effort suggested reliance on an inaccurate stereotype. The *beta* relating student sex to teacher perceptions of effort was -.16 (*p* < .001), even though the correlations of student sex with self-concept of ability, time spent on homework, and self-perceptions of effort were .09, -.03, and .00, respectively. Teachers apparently erroneously stereotype girls as trying harder, oblivious to boys' higher motivation (as indicated by self-concept of ability) and the similarities between boys' and girls' effort.

Which was a more powerful influence on teacher perceptions, sex stereotypes or individuating information? Table 2 clearly shows that all three teacher perceptions were based far more on students' grades, standardized test scores, and self-concept of ability than they were based on student sex.

*Social Class Stereotypes*

Analyses were identical to those examining teachers' sex stereotypes, except that instead of student sex, parental education and income were included in the equations predicting teacher perceptions of performance, talent, and effort. Results are presented in Table 3.

These analyses provided no evidence that teachers relied on social class stereotypes. The *R*<sup>2</sup> increment associated with adding parental income and education to the regression equations never exceeded .05 and was never statistically significant (all *F*s < 2.3, all *p*s > .1). Of the six possible individual relations between income and education and the three teacher-perception variables, only one was statistically significant (education predicted teacher perceptions of talent, *p* < .05), but the *beta* was very small (.05). Apparently, the accuracy of teacher perceptions of social class differences in performance, talent, and effort occurred because teachers evaluated students on the basis of their achievement and motivation—factors that correlated with social class.

**Table 3**

**Did Teachers Rely on Students' Social Class, Independent of Students' Achievement and Motivation in Study 2?**

Predictor	Teacher's perception		
	Performance	Talent	Effort
Parental income	-.02	.00	.00
Parental education	.02	.05*	.02
Fifth-grade final marks	.26**	.21**	.27**
Standardized test scores	.35**	.42**	.21**
Self-concept of math ability	.20**	.16**	.11**
Effort self-perceptions	.09**	.03	.13**
Time spent on homework	-.06*	-.05*	-.03
<i>R</i> <sup>2</sup>	.48**	.48**	.30**

NOTE: All entries are standardized regression coefficients. *N* = 1,066. \**p* < .05. \*\**p* < .01.

Which was a more powerful influence on teacher perceptions, social class stereotypes or individuating information? Table 3 clearly shows that teachers almost always based their perceptions more (and often much more) on the individuating information (previous achievement and motivation) than on student social class as indicated by parental income and education.

### *Ethnic Stereotypes*

For these analyses, students' ethnicity, grades, self-concept of ability, effort, and time spent on homework predicted teacher perceptions. Standardized test scores were not used as predictors because in the integrated districts, students had not taken a standardized test in fifth or early sixth grade. Therefore, using standardized test scores as a predictor would have had the undesirable effect of excluding these students from the analyses. These analyses included 1,873 White students and 96 African American students. Results are presented in Table 4.

**Table 4**

**Did Teachers Rely on Students' Ethnicity, Independent of Students' Achievement and Motivation in Study 2?**

Predictor	Teacher's perception		
	Performance	Talent	Effort
Student ethnicity	.00	.03	.06*
Fifth-grade final marks	.41*	.42*	.36*
Self-concept of math ability	.32*	.29*	.21*
Effort self-perceptions	.02	-.02	.10*
Time spent on homework	-.06*	-.06*	-.03
$R^2$	.38	.36	.26

NOTE: All entries are standardized regression coefficients.  $N = 1,969$  (1,873 White students and 96 African American students).

\* $p < .01$ .

These analyses provided little evidence that teachers relied on ethnic stereotypes. Student ethnicity had no significant effect on teacher perceptions of performance ( $\beta = .00$ , *ns*) and talent ( $\beta = .03$ ). Ethnicity did, however, have a small effect on teacher perceptions of effort ( $\beta = .06$ ,  $p < .01$ ), indicating a slight tendency to see African American students as trying harder.

Which was a more powerful influence on teacher perceptions, ethnic stereotypes or individuating information? Table 4 clearly shows that all three teacher perceptions were based far more on students' grades, standardized test scores, and self-concept of ability than they were based on student ethnicity.

### **Discussion**

Study 2 provided some clear insights into why the results of Study 1 showed such minimal evidence of bias. With a few notable exceptions (discussed below), teachers seemed to be basing their perceptions of students on those students' actual performance and motivation. Neither student social class nor ethnicity influenced teacher perceptions, after controlling for students' actual achievement and motivation. Similarly, student sex had no influence on teacher perceptions of talent, after controlling for students' actual achievement and motivation.

These results clearly rule out one possible explanation for the results showing accuracy in Study 1. Those results did not occur because teachers were relying on an accurate stereotype. Instead, these Study 2 results showed that teachers judged students almost exclusively on the basis of their actual performance and motivation. Thus, either teachers were oblivious to sex, class, and ethnic stereotypes, or they did not apply their stereotypes when evaluating their students.

The likelihood that teachers were oblivious to three of the major stereotypes in American culture seems vanishingly small. The cumulative wisdom of years of social psychological research on stereotypes points to the second explanation, that teachers did not apply their stereotypes. Abundant research in the laboratory and field shows that perceivers generally evaluate targets far more on the basis of targets' personal characteristics than on targets' membership in social groups (Jussim, 1990, 1991,

1993, for reviews). In general, the more individuating information perceivers have, the less they rely on stereotypes (Eagly, Ashmore, Makhijani, & Longo, 1991; Krueger & Rothbart, 1988; Locksley et al. 1980). Of course, teachers interacting with students over the first month of the school year generally have considerably more (and probably more objective) individuating information about students than do subjects in even the most ecologically valid laboratory experiment. Thus, it should come as no surprise that, in general, teachers did not rely much on their stereotypes when evaluating students.

There were a few exceptions to this pattern. In the case of student sex, teachers did indeed seem to be relying on an accurate stereotype regarding performance: Teachers apparently evaluated students' performance on the basis of their sex, independent of their actual achievement. However, the extent to which they did so corresponded reasonably well with actual prior sex differences in achievement. It is important to highlight just what this means. Because even an accurate stereotype does not apply equally well to all members of the stereotyped group, it seems likely that teachers misperceived some boys and girls. However, it also means that there was no tendency to systematically over- or underestimate the performance of girls.

In contrast, however, teachers seemed to be relying on an inaccurate stereotype in evaluating boys' and girls' effort. Teachers' more favorable impressions of girls' effort probably occur because, on average, girls are often more cooperative and pleasant than are boys and because teachers prefer more cooperative and pleasant students (e.g., Brophy & Good, 1974; Bye, 1994; Wentzel, 1989). The results regarding effort are consistent with a growing body of literature showing that school is often a hostile place for boys. For example, at least some teachers believe that boys suffer from inferior verbal skills—and this belief may become self-fulfilling (Palardy, 1969). Similarly, boys are referred for psychological evaluations far more than are girls, even when the teachers themselves do not rate boys as any more aggressive or in need of psychological services than girls (Bye, 1994). Similarly, one usually finds far more boys than girls in special education classes (Bye, 1994; Ravitch, 1993). And boys often receive lower grades

than do girls, even when their performances on standardized achievement tests are similar (Kimball, 1989).

The finding that teachers seemed to be relying on sex stereotypes more than ethnic or social class stereotypes is broadly consistent with other research suggesting something uniquely powerful about sex stereotypes. Specifically, after 25 years of research, there is currently excellent converging evidence from both laboratory and field studies in a variety of contexts, showing that sex stereotypes are often self-fulfilling. In contrast, the evidence regarding the self-fulfilling effects of ethnic or social class stereotypes is either extremely limited or nonexistent. Precisely identifying why sex stereotypes may be unique remains an important question for future research.

## CONCLUSION

Given the extent to which sexism, classism, and racism supposedly pervade American society, at first glance, the results of this study may appear surprising indeed. There was no evidence of teachers being biased against girls, students from lower-class backgrounds, or African American students. When teachers evaluated the students from one group more favorably than students from another group, those perceptions usually corresponded reasonably well to reality. The few biases and errors teachers seem to have committed were in the direction of evaluating students from traditionally disfavored groups more favorably than they deserved (girls' effort; in the African American, segregated district, performance and talent), or of seeing a slightly smaller difference than really existed (social class differences in performance and talent).

Of course, the current studies are mute on the question of the genesis of those real differences among groups. Furthermore, we are not claiming that stereotypes, prejudice, and discrimination do not exist or that they are unimportant. Perhaps sexism, classism, and racism contributed to the real differences among students. Nonetheless, claims about the power of stereotypes are rarely based on scientific, empirical data collected under naturalistic conditions. The current studies redress this limitation and pro-

vide little evidence of powerful or pervasive biases produced by stereotypes.

Social problems associated with gender, social class, and race undoubtedly exist, and they are terribly important. But economics, cultural differences among groups, socialization, after-effects of a history of discrimination, and a host of other factors probably play major roles in those social problems. The role of individuals' stereotypes in creating those problems is less clear.

## REFERENCES

- American Psychological Association. (1991). In the Supreme Court of the United States: *Price Waterhouse v. Ann B. Hopkins: Amicus curiae brief for the American Psychological Association*. *American Psychologist*, 46, 1061-1070.
- Anastasi, A. (1982). *Psychological testing*. New York: Macmillan.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Bulletin*, 84, 191-215.
- Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, 66, 5-20.
- Bodenhausen, G. V. (1988). Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of Personality and Social Psychology*, 55, 726-737.
- Brophy, J. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology*, 75, 631-661.
- Brophy, J., & Good, T. (1974). *Teacher-student relationships: Causes and consequences*. New York: Holt, Rinehart & Winston.
- Bye, L. (1994). *Referral of elementary age students for social skill training: Student and teacher characteristics*. Unpublished doctoral dissertation, Rutgers University, New Brunswick, NJ.
- Campbell, D. T. (1967). Stereotypes and the perception of group differences. *American Psychologist*, 22, 817-829.
- Darley, J. M., & Fazio, R. H. (1980). Expectancy-confirmation processes arising in the social interaction sequence. *American Psychologist*, 35, 867-881.
- Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, 44, 20-33.
- Devine, P. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5-18.
- Dusek, J., & Joseph, G. (1983). The bases of teacher expectancies: A meta-analysis. *Journal of Educational Psychology*, 75, 327-346.
- Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but . . . : A meta-analysis of research on the physical attractiveness stereotype. *Psychological Bulletin*, 110, 109-128.
- Eccles, J. (1988). *Achievement beliefs and environment*. Final report to the National Institute of Child Health and Development. Ann Arbor: University of Michigan.
- Eccles (Parsons), J., Adler, T., & Meece, J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology*, 46, 26-43.
- Eccles, J., & Wigfield, A. (1985). Teacher expectations and student motivation. In J. Dusek (Ed.), *Teacher expectancies* (pp. 185-226). Hillsdale, NJ: Erlbaum.
- Elashoff, J. D., & Snow, R. E. (1971). *Pygmalion reconsidered*. Worthington, OH: Jones.
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, 23, 1-74.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). Reading, MA: Addison-Wesley.
- Hamilton, D. L., Sherman, S. J., & Ruvolo, C. M. (1990). Stereotype-based expectancies: Effects on information processing and social behavior. *Journal of Social Issues*, 46, 35-60.
- Jones, E. E. (1990). *Interpersonal perception*. New York: W. H. Freeman.
- Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review*, 100, 109-128.
- Jussim, L. (1987). Interpersonal expectations in social interaction: Self-fulfilling prophecies, confirmatory biases, and accuracy (Doctoral dissertation, University of Michigan, 1987). *Dissertation Abstracts International*, 48, 1845B.
- Jussim, L. (1989). Teacher expectations: Self-fulfilling prophecies, perceptual biases, and accuracy. *Journal of Personality and Social Psychology*, 57, 469-480.
- Jussim, L. (1990). Social reality and social problems: The role of expectancies. *Journal of Social Issues*, 46, 9-34.
- Jussim, L. (1991). Social perception and social reality: A reflection-construction model. *Psychological Review*, 98, 54-73.
- Jussim, L. (1993). Accuracy in interpersonal expectations: A reflection-construction analysis of current and classic research. *Journal of Personality*, 61, 637-668.

- Jussim, L., & Eccles, J. (1992). Teacher expectations: II. Reflection and construction of student achievement. *Journal of Personality and Social Psychology*, *63*, 947-961.
- Jussim, L., & Eccles, J. (1995). Naturalistic studies of interpersonal expectancies. *Review of Personality and Social Psychology*, *15*, 74-108.
- Jussim, L., Madon, S., & Chatman, C. (1994). Teacher expectations and student achievement: Self-fulfilling prophecies, biases, and accuracy. In L. Heath et al. (Eds.), *Applications of heuristics and biases to social issues* (pp. 303-334). New York: Plenum.
- Kimball, M. M. (1989). A new perspective on women's math achievement. *Psychological Bulletin*, *105*, 198-214.
- Krueger, J., & Rothbart, M. (1988). Use of categorical and individuating information in making inferences about personality. *Journal of Personality and Social Psychology*, *55*, 187-195.
- Linville, P. (1982). The complexity-extremity effect and age-based stereotyping. *Journal of Personality and Social Psychology*, *42*, 193-211.
- Locksley, A., Borgida, E., Brekke, N., & Hepburn, C. (1980). Sex stereotypes and social judgment. *Journal of Personality and Social Psychology*, *39*, 821-831.
- Marger, M. N. (1991). *Race and ethnic relations* (2nd ed.). Belmont, CA: Wadsworth.
- McCauley, C., Stitt, C. L., & Segal, M. (1980). Stereotyping: From prejudice to prediction. *Psychological Bulletin*, *87*, 195-208.
- Miller, D. T., & Turnbull, W. (1986). Expectancies and interpersonal processes. *Annual Review of Psychology*, *37*, 233-256.
- Palardy, J. (1969). What teachers believe—What students achieve. *Elementary School Journal*, *69*, 370-374.
- Parsons, J. E. (1980). *Final Report to the National Institute of Education*. Washington, DC: National Institute of Education. (ERIC Document Reproduction Service No. ED 186 477)
- Ravitch, D. (1993, November 26). Gender bias in the schools? It isn't true. *Newsday*, p. 151.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and student intellectual development*. New York: Holt, Rinehart and Winston.
- Sedikides, C., & Skowronski, J. J. (1991). The law of cognitive structure activation. *Psychological Inquiry*, *2*, 169-184.
- Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology*, *71*, 3-25.

- Wentzel, K. R. (1989). Adolescent classroom goals, standards for performance, and academic achievement: An interactionist perspective. *Journal of Educational Psychology*, *81*, 131-142.
- West, C., & Anderson, T. (1976). The question of preponderant causation in teacher expectancy research. *Review of Educational Research*, *46*, 613-630.
- Williams, T. (1976). Teacher prophecies and the inheritance of inequality. *Sociology of Education*, *49*, 223-236.
- Wineburg, S. S. (1987). The self-fulfillment of the self-fulfilling prophecy. *Educational Researcher*, *16*, 28-37.