A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?

By THOMAS S. DEE*

The large and persistent achievement gaps separating minority and nonminority students are arguably the most important educational problem in the United States. In particular, reducing or eliminating these gaps by raising the achievement of minority students is widely seen as a critical component of promoting broader social equality with respect to a variety of outcomes like educational attainment and earnings as well as crime, health, and family structure (e.g., Christopher Jencks and Meredith Phillips, 1998). The more modest gender gaps in achievement are also viewed as a prominent policy concern, particularly with respect to the fields of science and mathematics (e.g., American Association of University Women, 1992). The recent federal legislation, the No Child Left Behind (NCLB) Act of 2001, clearly reflects these concerns, explicitly requiring that these demographic subgroups make "adequate yearly progress" toward proficiency on state tests.

NCLB also emphasizes that schools should meet these and other goals by implementing effective reforms grounded in "scientifically based" research. However, the available empirical evidence suggests that the determinants of the demographic achievement gaps are not very well understood. For example, Jencks and Phillips (1998) argue that traditional explanations for the black–white achievement gaps (i.e., those based on income inequality, differences in family structure, and school spending) actually have relatively little explanatory power. They also suggest that more successful future explanations of achievement gaps are likely to be based on credible data that assess more nuanced hypotheses about the dynamics within schools, classrooms and families.

The notion that the classroom dynamics between teachers and students make a substantive contribution to the demographic gaps in achievement already has a wide currency among educational researchers and commentators. For example, proposals for promoting racial and gender equity often emphasize the need for improvements in teacher training and professional development (e.g., American Association of University Women, 1992; Ronald F. Ferguson, 1998). Another frequently recurring proposal for increasing the relative achievement of minority students is to improve the recruitment and retention of minority teachers (e.g., Patricia A. Graham, 1987; Joint Center for Political Studies, 1989; National Commission on Teaching and America's Future, 1996; Beatriz Chu Clewell and Ana María Villegas, 1998). However, the evidence that the demographic interactions between students and teachers matter is surprisingly thin, sometimes contradictory, and usually based on small, localized samples.

In this study, I present new evidence on this issue by evaluating whether assignment to a demographically similar teacher influences the teacher's subjective evaluations of student behavior and performance. This analysis is based on a large, nationally representative survey, the National Education Longitudinal Study of 1988 (NELS:88). These data allow me to examine whether the effects of demographically similar teachers vary with key student traits or across Census regions. However, the key innovation of this analysis arguably involves how the identification strategy exploits a unique feature of the NELS:88 survey design to purge the potential biases created by the nonrandom sorting of students across and within schools. Specifically, for each of the 8th-grade students surveyed, NELS:88 solicited student-specific evaluations from teachers in two distinct academic subjects. This unusual feature of the data makes it pos-

* Department of Economics, 500 College Avenue, Swarthmore College, Swarthmore, PA 19081, and NBER (e-mail: dee@swarthmore.edu). I thank Brian Jacob and session participants for helpful comments. This research was supported by the Spencer Foundation and by a grant from the American Educational Research Association which receives funds for its "AERA Grants Program" from the National Science Foundation and the National Center for Education Statistics of the Institute of Education Sciences (U.S. Department of Education) under NSF grant no. REC-0310268. Opinions reflect those of the author and do not necessarily reflect those of the granting agencies.

sible to implement a fixed effect econometric specification that essentially identifies how two demographically different teachers evaluated the same student.

I. Teacher–Student Interactions

The extant literature suggests two general ways that the demographic matches between students and teachers could influence educational outcomes. One broad class of explanations involves what could be called "passive" teacher effects. These effects are simply triggered by a teacher's racial, ethnic, or gender identity, not by explicit teacher behaviors. The most widely discussed examples are "role-model" effects, which occur when the presence of a demographically similar teacher raises a student's academic motivation and expectations. A related type of passive teacher effect is the phenomenon known as "stereotype threat" (Claude M. Steele, 1997). Stereotype threat refers to the possibility that, in situations where students perceive stereotypes might attach (e.g., black students with white teachers, female students with male teachers), they experience an apprehension that retards their academic identification and subsequent achievement. A second class of explanations for the educational benefits of own-race teachers points to "active" teacher effects: unintended biases in their prior expectations of and interactions with students who have different demographic traits (e.g., Ferguson, 1998 p. 294).

The available evidence on whether any of these sorts of effects exist is limited but generally supportive. For example, a few small-scale experiments from the 1970s (see Jacqueline Jordan Irvine, 1990 table 3.1) suggest that white teachers do provide black students with less assistance and positive feedback. Similarly, the growing literature that began with a seminal study by Steele and Joshua Aronson (1995) suggests that stereotype threats by race and gender influence student achievement. Furthermore, a recent analysis of data from Tennessee's Project STAR class-size experiment (Dee, 2004) indicates that random assignment to a racially similar teacher improved the test scores of both black and white students. However, most of the relevant literature has focused on how pairings by race, ethnicity, and gender influence teachers' perceptions and expectations of students. A relative emphasis on such subjective outcomes is uncommon among economists studying education. Nonetheless, teacher perceptions clearly influence student access to future educational opportunities and may also shape the learning environment in meaningful ways. Therefore, analyses of these perceptions may provide a useful complement to conventional studies of student outcomes like performance on low-stakes tests.

Does assignment to a demographically similar teacher influence that teacher's perceptions of the student? The results from the available empirical literature are decidedly mixed. In an influential literature review, Ferguson (1998 p. 313) concludes that biases in teacher perceptions and expectations "help to sustain, and perhaps even to expand, the black–white test score gap." However, he also argues that this is a problem for both black and white teachers and dismisses recommendations to match students and teacher by race as "too simple a prescription." However, other reviewers (e.g., Irvine, 1988) have pointed to several small studies, which find that white teachers are more likely than black teachers to have negative perceptions and low expectations of black students (e.g., Albert R. Griffin and Clement B. G. London, 1979; Charles H. Beady and Stephen Hansell, 1981).

One possible explanation for conflicting interpretations of this limited evidence is the unintended bias that can be created by the nonrandom sorting of teachers and students both across and within schools. For example, Ronald G. Erhenberg et al. (1995), in one of the few studies based on a large, nationally representative survey and a rich set of background controls, find relatively limited evidence that black students are evaluated more positively by black teachers. However, if minority students with a propensity for poor outcomes were more likely to be assigned to minority teachers, the true effect of a demographically similar teacher would be understated by those results. In this study, I present new empirical evidence that attempts to address this identification problem directly by exploiting the unusual panel nature of the NELS:88 data. This evidence does not directly distinguish among the passive and active effects described here. Instead, it provides

reduced-form evidence on whether having a demographically similar teacher influences the teachers' perceptions of a student's performance and behavior.

II. Data and Specifications

The National Education Longitudinal Study of 1988 (NELS:88) is a nationally representative, longitudinal study that began in 1988 with a sample of 24,599 8th-grade students from 1,052 public and private schools (Steven J. Ingels et al., 1990). NELS:88 had a two-stage sampling design. Schools, the primary sampling unit, were selected with probabilities proportional to their 8th-grade enrollment. Approximately 26 students were then randomly chosen within each participating school. NELS:88 also fielded questionnaires to the teachers responsible for teaching each of the selected students in two of four academic subjects: mathematics, science, reading, and social studies. The surveyed teachers were chosen by randomly assigning each school to one of four subject-area groupings: mathematics/reading, mathematics/social studies, science/reading, and science/social studies. Two completed teacher surveys are available for 21,324 of the 8th-grade students because of some nonresponse and because some students did not have a class in one or both of their assigned academic subjects. The final data set consists of 42,648 observations since the unit of observation is each teacher–student pairing.

The teacher survey included several questions about how the teacher perceived the classroom performance and personal traits of individual sampled students. This analysis focuses on three pejorative teacher assessments (Table 1): whether the student was seen as frequently disruptive (DISRUPT), consistently inattentive (INATTEN), or rarely completed homework (NOHWK). One potentially important interpretative complication involves how these teacher perceptions relate to conventional measures of student achievement. Specifically, in this study, being perceived as disruptive or inattentive is viewed as a clearly undesirable outcome. However, it may be that a student becomes disruptive or inattentive to a particular teacher because (s)he has mastered the classroom material relative to their peers. Fortunately, this concern does not appear to be empirically relevant. More specifically, using these NELS:88 data, I found that, conditional on student and subject fixed effects, students performed significantly lower on subject tests when the teacher for that subject viewed them negatively. The students viewed negatively by teachers were also substantially less likely than other students in their school to take any Advanced Placement courses over the subsequent two years and more likely to have dropped out of high school. Furthermore, the results based on subjective teacher assessments that are not as subject to this ambiguity (e.g., not doing homework or being tardy to class) are similar to those based on these variables.

TABLE 1—DESCRIPTIVE STATISTICS, MATCHED STUDENT–TEACHER OBSERVATIONS, NELS:88

Variable	Description	Mean	Sample size
DISRUPT	Frequently disruptive	0.13	41,580
INATTEN	Consistently inattentive	0.21	41,536
HOMEWK	Rarely completes homework	0.20	41,627
OTHRACE	Teacher of other race or ethnicity	0.30	42,648
OTHSEX	Teacher of opposite gender	0.49	42,648

In order to identify the effects of demographically similar teachers, the students and teachers participating in NELS:88 were identified as belonging to one of four possible racial/ethnic categories: white (non-Hispanic), black (non-Hispanic), Hispanic, and all others. The binary indicators, OTHRACE and OTHSEX, identify students when observed by teachers who do not share their race or ethnicity or their gender (Table 1). The sample mean for OTHRACE is 30 percent. However, it should be noted that the prevalence of OTHRACE varied quite sharply across minority and nonminority students. Only 6 percent of white non-Hispanic student observations were with teachers who were not white non-Hispanic. In contrast, 67 percent of the black students and 89 percent of the Hispanic students were with a teacher who did not share their race or ethnicity.

The ability to make within-student comparisons with these data makes it possible to eliminate the biases generated by unobserved student traits. The results presented here also

condition on fixed effects for the subject in which the student is evaluated. However, race-specific patterns of unobserved teacher and classroom quality could quite reasonably bias these evaluations. For example, these evaluations would overstate the effects of OTHRACE on the performance of minority students if predominantly minority schools tended to attract relatively low-quality white teachers. A similar bias could occur if the white teachers within such schools were more likely to be assigned to classrooms with low-quality students or inadequate resources.

To assess the empirical relevance of these issues, I present the results of some specifications that introduce controls for several observed teacher and classroom traits. These include separate binary indicators for having attended graduate school, for having a major (graduate or undergraduate) in the academic subject they are teaching, and for having a minor in the academic subject they are teaching. There are also eight dummy variables that identify each teacher's years of experience (i.e., 1–3 years, 4–6 years, etc.). Class size is also introduced as a control variable. An additional set of four dummy variables indicates how the teacher compared the achievement of the surveyed student's class to that of the average 8th-grade student in the school (i.e., higher, average, lower, widely differing, no response). Prior research suggests that some of these controls have surprisingly weak effects on student achievement (e.g., Andrew J. Wayne and Peter Youngs, 2003). However, these variables are jointly significant determinants of all three of these dependent variables, with teacher experience and peer quality having particularly strong effects. Furthermore, since some teachers provided evaluations of multiple students, it is also possible to evaluate the effects of demographically similar teachers in linear probability models that include both student and teacher fixed effects. The results of that approach, though less precise, are quite similar to those reported here.

However, the results presented here are based on a fixed-effects logit that accommodates both the presence of student fixed effects and the binary nature of the dependent variables. A key feature of this approach is a *conditional* likelihood function that effectively removes student fixed effects from the estimation procedure (e.g., Badi H. Baltagi, 2001).[1] Jeffrey M. Wooldridge (2002 p. 492) cautions that it is misleading to state that this approach "conditions" on the unobserved fixed effects. Instead, this approach relies on describing a conditional density that allows one to identify the "structural" coefficients from the available data. Nonetheless, I found that ordinary least-squares (OLS) estimates based on linear probability models that do condition on student fixed-effects yield results quite similar to those based on the fixed-effects logit approach.

TABLE 2—ODDS RATIOS FROM FIXED-EFFECTS LOGIT MODELS FOR TEACHER PERCEPTIONS OF STUDENTS, NELS:88 8TH-GRADERS

Dependent variable	Teacher–class controls?	Independent variable		p value	Sample size
		OTHRACE	OTHSEX		
DISRUPT	no	1.36** (3.08)	1.38** (5.97)	0.0104	6,028
	yes	1.51** (3.84)	1.37** (5.49)	0.0000	5,744
INATTEN	no	1.33** (3.40)	1.19** (3.89)	0.0313	8,710
	yes	1.34** (3.35)	1.20** (3.93)	0.0000	8,286
HOMEWK	no	1.22* (2.33)	1.15** (2.83)	0.0030	7,528
	yes	1.29** (2.92)	1.15** (2.76)	0.0000	7,158

Notes: The cells contain the odds ratios from fixed-effects logit models. The absolute values of z statistics associated with the null hypothesis of no effect are reported in parentheses. All models include fixed effects for the academic subject of the student–teacher pairing. The additional teacher–class controls are: class size; three fixed effects which indicate whether the teacher has a major in the subject, a minor in the subject, or a graduate degree; and eight fixed effects that correspond to each teacher's experience level. The p values refer to Hausman tests comparing the fixed effect estimates to conventional logit results.

* Statistically significant at the 5-percent level.

** Statistically significant at the 1-percent level.

III. Results

Table 2 presents the key results from fixed-effect logit models based on the full sample and for each of the three binary outcomes. Since this estimation strategy does not actually generate

[1] One implication of this approach is that only the observations with within-student variation in the dependent variable contribute to the likelihood function. This explains why the sample sizes reported here are smaller than the total sample and vary across dependent variables.

estimates of the student fixed effects, I cannot directly evaluate the partial effects of the independent variables on the response probabilities. However, the magnitude of the estimated coefficients can be interpreted by converting them to odds ratios: the estimated factor by which an independent variable influences the odds of a particular outcome.[2] For example, the results in the first row of Table 2 indicate that the odds of a student being seen as disruptive by a teacher are 1.36 times as large when the teacher does not share the student's racial/ethnic designation. These odds ratios are quite consistent with the marginal effects implied by similarly specified linear probability models.[3]

The other results in Table 2 indicate that OTHRACE has similarly large and statistically significant effects on the other teacher perceptions. More specifically, having a teacher who does not share a student's racial/ethnic designation increases the odds of the student being seen as inattentive by at least 33 percent and the odds of rarely completing homework by at least 22 percent. The estimated effects of OTHSEX are similarly sized and statistically significant. The odds that a student was perceived as inattentive or disruptive are respectively at least 19- and 37-percent higher when the teacher is of the opposite gender. And the odds that a teacher will report that a student rarely completes homework are 15-percent higher when their genders do not match. Interestingly, Hausman tests comparing the results of all these fixed-effect logit models to those from conventional logit models that are otherwise similarly specified indicate that there are statistically significant differences. These comparisons suggest that acknowledging the unobserved student fixed effects does lead to more reliable inferences. The results in Table 2 also indicate that the estimated effects of OTHRACE and OTHSEX are very similar in specifications that introduce controls for teacher and classroom observables. In fact, including these controls generally leaves these estimated effects unchanged or increases them slightly. This pattern of results suggests that the effects identified here reflect racial, gender, and ethnic dynamics between students and teachers and are not biased by any demographic patterns in unobserved teacher and peer quality.

The results in Table 2 provide evidence that the racial/ethnic and gender dynamics within classrooms have quite large effects on how individual students are perceived by their teachers. One novel feature of these inferences is that they use the panel nature of the NELS:88 data to purge the inconsistency that could be imparted by unobserved student effects. However, as noted earlier, two other possibly valuable features of these data are that the sample size is relatively large and the data are nationally representative. These aspects of the data make it possible to examine how the effects of OTHRACE and OTHSEX might vary by student traits as well as across different regions of the United States. The results presented in Tables 3, 4, and 5 provide some evidence on the pattern of response heterogeneities for each of the three dependent variables.

For example, the results in Table 3 indicate that both white and minority (i.e., black and Hispanic) students are likely to be perceived as disruptive by a teacher who does not share their racial/ethnic designation. Similarly, both male and female students are more likely to be seen as disruptive by an OTHSEX teacher. However, the results in Table 4 indicate that minority and female students are particularly likely to be seen as inattentive by OTHRACE and OTHSEX teachers. The effects of OTHRACE teachers are more consistently different across students with high and low socioeconomic status. Among students with low socioeconomic status, the odds of being seen negatively are 35–57-percent higher when evaluated by an OTHRACE teacher. In contrast, these effects, though positive, are consistently smaller and statistically insignificant among students with high socioeconomic status.

The remaining results in Tables 3, 4, and 5 indicate how the estimated effects of OTHRACE and OTHSEX differ across the four Census regions. In each of these regions, teach-

[2] The odds ratio equals e^{β}, so the absence of an effect (i.e., $\beta = 0$) implies an odds ratio of 1 (i.e., $e^{\beta} = e^{0} = 1$). The z statistics reported in parentheses correspond to the null hypothesis of no effect.

[3] For example, the corresponding OLS-estimated effect of OTHRACE on DISRUPT is about 3.4 percentage points (t statistic = 4.08). If that assignment shifted the probability of DISRUPT from 0.100 to 0.134, the implied growth in the odds of DISRUPT would be 39 percent (i.e., from 0.111 to 0.155).

TABLE 3—FIXED-EFFECTS LOGIT MODELS FOR TEACHER PERCEPTIONS OF STUDENTS AS FREQUENTLY DISRUPTIVE BY SAMPLE TRAIT

Sample	Independent variables		Sample size
	OTHRACE	OTHSEX	
Full sample	1.51**	1.37**	5,744
	(3.84)	(5.49)	
White non-Hispanic students	1.64**	1.35**	3,728
	(3.25)	(4.15)	
Black and Hispanic students	1.47*	1.39**	1,710
	(2.46)	(3.03)	
Male students	1.48**	1.30**	3,854
	(2.95)	(3.47)	
Female students	1.57*	1.70**	1,890
	(2.43)	(4.83)	
Low-SES students	1.64**	1.35**	3,134
	(3.64)	(3.86)	
High-SES students	1.33	1.43**	2,606
	(1.59)	(4.11)	
Northeast region	0.94	1.51*	916
	(0.17)	(2.42)	
North-Central region	1.30	1.33*	1,442
	(1.04)	(2.49)	
South region	1.89**	1.44**	2,310
	(4.26)	(3.83)	
West region	1.17	1.30*	1,072
	(0.58)	(1.96)	

Notes: The cells contain the odds ratios from fixed-effects logit models. The absolute values of z statistics associated with the null hypothesis of no effect are reported in parentheses. All models include fixed effects for the academic subject of the student–teacher pairing and the additional teacher–class controls described in Table 2.

* Statistically significant at the 5-percent level.

** Statistically significant at the 1-percent level.

TABLE 4—FIXED-EFFECTS LOGIT MODELS FOR TEACHER PERCEPTIONS OF STUDENTS AS FREQUENTLY INATTENTIVE BY SAMPLE TRAIT

Sample	Independent variables		Sample size
	OTHRACE	OTHSEX	
Full sample	1.34**	1.20**	8,286
	(3.35)	(3.93)	
White non-Hispanic students	1.13	1.19**	5,422
	(1.00)	(2.99)	
Black and Hispanic students	1.71**	1.32**	2,360
	(4.02)	(3.12)	
Male students	1.13	0.99	4,720
	(1.05)	(0.01)	
Female students	1.64*	1.33**	3,566
	(3.71)	(3.64)	
Low-SES students	1.53**	1.17*	4,568
	(3.72)	(2.50)	
High-SES students	1.10	1.26**	3,716
	(0.70)	(3.20)	
Northeast region	0.74	1.15	1,434
	(1.13)	(1.15)	
North-Central region	1.00	1.27*	1,966
	(0.01)	(2.42)	
South region	1.61**	1.27**	3,408
	(4.03)	(3.24)	
West region	1.35	1.05	1,476
	(1.31)	(0.42)	

Notes: The cells contain the odds ratios from fixed-effects logit models. The absolute values of z statistics associated with the null hypothesis of no effect are reported in parentheses. All models include fixed effects for the academic subject of the student–teacher pairing and the additional teacher–class controls described in Table 2.

* Statistically significant at the 5-percent level.

** Statistically significant at the 1-percent level.

ers who did not share a student's gender were significantly more likely to view the student as disruptive. Similarly, the estimated effects of OTHSEX on the odds of being seen as inattentive were similarly large and statistically significant in two of the four regions. However, the estimated effect of OTHSEX on teacher perceptions of NOHWK was statistically significant only in the South (Table 5). In contrast, the geographic differences in the estimated effects of an OTHRACE teacher were more consistently heterogeneous. More specifically, the effects of OTHRACE on all three teacher perceptions were statistically significant only in the South. The estimated magnitudes of these effects are quite large. For example, OTHRACE increased the odds that a student would be seen as disruptive and inattentive by 89 percent and 61 percent, respectively. Evaluations that distinguish among the white and minority students in the South suggest that the OTHRACE on teacher perceptions are positive for both groups. However, the resulting loss of statistical precision qualifies these within-region comparisons.

IV. Conclusions

The results presented here indicate that the racial, ethnic, and gender dynamics between students and teachers have consistently large effects on teacher perceptions of student performance. However, the effects associated with race and ethnicity appear to be concentrated among students of low socioeconomic status and those in the South. Since these teacher perceptions are clearly likely to influence

TABLE 5—FIXED-EFFECTS LOGIT MODELS FOR TEACHER PERCEPTIONS OF STUDENTS AS RARELY COMPLETING HOMEWORK BY SAMPLE TRAIT

Sample	Independent variables		Sample size
	OTHRACE	OTHSEX	
Full sample	1.30**	1.15**	7,158
	(2.92)	(2.76)	
White non-Hispanic students	1.25†	1.16*	4,448
	(1.79)	(2.30)	
Black and Hispanic students	1.37*	1.24*	2,258
	(2.39)	(2.40)	
Male students	1.21	1.26**	4,288
	(1.61)	(3.27)	
Female students	1.46**	0.89	2,870
	(2.79)	(1.42)	
Low-SES students	1.37**	1.10	4,220
	(2.83)	(1.44)	
High-SES students	1.18	1.24**	2,934
	(1.07)	(2.66)	
Northeast region	0.99	1.22	1,218
	(0.02)	(1.47)	
North-Central region	0.98	1.10	1,648
	(0.18)	(0.86)	
South region	1.57**	1.25**	3,014
	(3.80)	(2.77)	
West region	0.92	1.02	1,276
	(0.36)	(0.36)	

Notes: The cells contain the odds ratios from fixed-effects logit models. The absolute values of z statistics associated with the null hypothesis of no effect are reported in parentheses. All models include fixed effects for the academic subject of the student–teacher pairing and the additional teacher–class controls described in Table 2.

† Statistically significant at the 10-percent level.

* Statistically significant at the 5-percent level.

** Statistically significant at the 1-percent level.

educational opportunities as well as the classroom environment, this evidence implies that these classroom interactions make important contributions to the observed demographic gaps in student achievement. The most widely recommended policy responses to these sorts of effects are arguably the ones that involve recruiting underrepresented teachers. One clear benefit of this approach is that it does not require a clear understanding of the extent to which the effects documented here are driven by passive responses (e.g., role-model effects and stereotype threat) or active biases in student or teacher behaviors. However, the results presented here also indicate that this approach could have the unintended and undesirable consequence of harming students who do not share the teacher's demographic traits.

This criticism suggests that alternative policies that improve the effectiveness of all teachers may be a relatively attractive way to close achievement gaps. For example, Ferguson (1998) recommends the implementation of more sophisticated programs of professional development for teachers as well as well-designed performance incentives. Steele (1997) recommends "wise" schooling practices and programs that negate the stereotype threats experienced by students through an optimistic emphasis on their potential and the offer of challenges instead of remediation. Policies of this broad nature do appear to offer a particularly promising way to promote the demographic neutrality of student–teacher interactions. However, the exact design and emphasis of such policies also require a clear understanding of the underlying structural mechanisms that make these student–teacher interactions relevant in the first place. For example, student-focused programs designed to negate stereotype threat would be relatively ineffective if the dominant problem involves biases in teacher behaviors. Similarly, if stereotype threat is the main source of the effects presented here, teacher training unrelated to those passive effects would also be ineffectual. Future research that illuminates the nature of these student–teacher dynamics will provide a particularly useful guide to sensible public policy.

REFERENCES

American Association of University Women. *How schools shortchange girls*. Washington, DC: AAUW Educational Foundation, 1992.

Baltagi, Badi H. *Econometric analysis of panel data*, 2nd Ed. Chichester, U.K.: Wiley, 2001.

Beady, Charles H. and Hansell, Stephen. "Teacher Race and Expectations for Student Achievement." *American Educational Research Journal*, 1981, *18*(2), pp. 191–206.

Clewell, Beatriz Chu and Villegas, Ana María. "Diversifying the Teaching Force to Improve Urban Schools: Meeting the Challenge—Introduction." *Education and Urban Society*, 1998, *31*(1), pp. 3–17.

Dee, Thomas S. "Teachers, Race and Student Achievement in a Randomized Experiment." *Review of Economics and Statistics*, 2004, *86*(1), pp. 195–210.

Ehrenberg, Ronald G.; Goldhaber, Dan D. and Brewer, Dominic J. "Do Teachers' Race, Gender and Ethnicity Matter? Evidence from the National Educational Longitudinal Study of 1988." *Industrial and Labor Relations Review*, 1995, *48*(3), pp. 547–61.

Ferguson, Ronald F. "Teachers' Perceptions and Expectations and the Black–White Test Score Gap," in C. Jencks and M. Phillips, eds., *The black–white test score gap*. Washington, DC: Brookings Institution Press, 1998, pp. 217–317.

Graham, Patricia A. "Black Teachers: A Drastically Scarce Resource." *Phi Delta Kappan*, 1987, *68*(8), pp. 598–605.

Griffin, Albert R. and London, Clement B. G. "Student Relations among Inner-City Teachers: A Comparative Study by Teacher Race." *Education*, 1979, *101*(2), pp. 139–47.

Ingels, Steven J.; Abraham, Sameer Y.; Karr, Rosemary; Spencer, Bruce D. and Frankel, Martin R. *National Education Longitudinal Study of 1988 base year: Student component data file user's manual*. Washington, DC: U.S. Department of Education, National Center for Education Statistics, 1990.

Irvine, Jacqueline Jordan. "An Analysis of the Problem of Disappearing Black Educators." *Elementary School Journal*, 1988, *88*(5), pp. 503–13.

Irvine, Jacqueline Jordan. *Black students and school failures: Policies, practices and prescriptions*. Westport, CT: Greenwood Press, 1990.

Jencks, Christopher and Philips, Meredith. "The Black–White Test Score Gap: An Introduction," in C. Jencks and M. Phillips, eds., *The black–white test score gap*. Washington, DC: Brookings Institution Press, 1998, pp. 1–51.

Joint Center for Political Studies. *Visions of a better way—A Black appraisal of public schooling*. Washington, DC: Joint Center for Political Studies, 1989.

National Commission on Teaching and America's Future. *What matters most: Teaching for America's future*. New York, NY: National Commission on Teaching and America's Future, 1996.

Steele, Claude M. "A Threat in the Air—How Stereotypes Shape Intellectual Identity and Performance." *American Psychologist*, 1997, *52*(6), pp. 613–29.

Steele, Claude M. and Aronson, Joshua. "Stereotype Threat and the Intellectual Test-Performance of African Americans." *Journal of Personality and Social Psychology*, 1995, *69*(5), pp. 797–811.

Wayne, Andrew J. and Youngs, Peter. "Teacher Characteristics and Student Achievement Gains: A Review." *Review of Educational Research*, 2003, *73*(1), pp. 89–122.

Wooldridge, Jeffrey M. *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press, 2002.