Table 1
Racial-Group Differences (d), Stereotype Threat, and Test Score Means, Deviations, and Intercorrelations

Variable	1	2	3	*d*
1. Race	—			1.99
2. Threat	−.56**	—		−0.98
3. Raven	.31*	−.32*	—	0.66
M	0.48	8.67	17.36	
SD	0.50	3.31	7.12	

Note. Correlations in column 1 are estimated from the *d*s in the last column as reported by McKay et al. (2003, Table 2, p. 8) and differ slightly from their Table 1 correlations, presumably because of the effects of missing data in their Table 1. Whites ($n = 42$) were coded as ones and Blacks ($n = 45$) were coded as zeros in statistical analyses. Raven = Raven Advanced Progressive Matrices. Adapted from McKay et al. (2003).
* $p < .01$. ** $p < .001$.

struct that has no theoretical meaning in psychology, with stereotype threat, which, according to Sackett et al. (2004), has considerable consensually agreed upon theoretical meaning.

Sackett et al. (2004) required that stereotype threat, to be accepted as an explanatory mechanism, had to explain the "prototypical" Black–White standardized mean difference (*d*) between test scores of one standard deviation (Sackett et al., 2004, Footnote 1, p. 8). Effect sizes (e.g., *d*) may be converted to each other's metric (e.g., r^2; Thompson, 2002, p. 69). Thus, the prototypical *d* means that 20% of test-score variance is attributable to racial group; this variance may be replaced by stereotype threat if it consistently explains the same 20% of variance in test scores. In McKay et al.'s (2003) study, stereotype threat accounted for virtually all of the racial-group difference in their sample (i.e., 10%), but more studies are needed to determine the generalizability of its effect. Yet it does not seem unreasonable to suppose that, in subsequent studies, stereotype threat might explain even larger racial-group differences (i.e., 20% of test-score variance) when they occur.

In reacting to Sackett et al.'s (2004) minimization of the impact of stereotype threat on test scores, perhaps Steele and Aronson (2004) conceded the point too readily. Customarily, in psychology, researchers determine impact or effect size after testing hypotheses with data rather than by relying exclusively on their belief systems, no matter how benevolent their intentions. Future studies just might reveal that stereotype threat has a greater impact on racial-group mean test scores than either set of researchers presently believes and, in the process, provide empirically supported answers to the question of what really causes racial-group differences in cognitive ability test scores.

REFERENCES

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

McKay, P. F., Doverspike, D., Bowen-Hilton, D., & McKay, Q. D. (2003). The effects of demographic variables and stereotype threat on Black/White differences in cognitive ability test performance. *Journal of Business and Psychology, 18,* 1–14.

Sackett, P. R., Hardison, C. M., & Cullen, M. J. (2004). On interpreting stereotype threat as accounting for African American–White differences on cognitive tests. *American Psychologist, 59,* 7–13.

Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High-stakes testing in employment, credentialing, and higher education. *American Psychologist, 56,* 302–318.

Schmitt, N., Sackett, P. R., & Ellingson, J. E. (2002). No easy solution to subgroup differences. *American Psychologist, 57,* 305–306.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797–811.

Steele, C. M., & Aronson, J. A. (2004). Stereotype threat does not live by Steele and Aronson (1995) alone. *American Psychologist, 59,* 47–48.

Thompson, B. (2002). "Statistical," "practical," and "clinical": How many kinds of significance do counselors need to consider? *Journal of Counseling & Development, 80,* 64–71.

Correspondence concerning this comment should be addressed to Janet E. Helms, Institute for the Study and Promotion of Race and Culture, Boston College, 318 Campion Hall, Chestnut Hill, MA 02467. E-mail: helmsja@bc.edu

DOI: 10.1037/0003-066X.60.3.270

Stereotype Threat and the Social and Scientific Contexts of the Race Achievement Gap

Geoffrey L. Cohen
Yale University

David K. Sherman
University of California, Santa Barbara

In their article, Sackett, Hardison, and Cullen (January 2004) critiqued misrepresentations of the original stereotype threat findings presented by Steele and Aronson (1995). They criticized representations of the research that suggest that stereotype threat explains all the racial achievement gap in academic performance when, in fact, the original studies statistically equated the ability of Black students and White students by using SAT scores as a covariate. As Sackett et al. acknowledged, Steele and Aronson did not claim that stereotype threat explains all the racial achievement gap, though as they suggested in their critique, it may have been a claim made implicitly and even explicitly in some media and textbook coverage of the work.

We wish to make three points that Sackett and colleagues (2004) did not make. These points highlight the social and scientific contexts in which Sackett et al.'s critical commentary, and stereotype threat research in general, can be interpreted. The first point is that several studies have demonstrated large stereotype threat effects (reducing and even eliminating group-based differences in achievement) without covarying out participants' prior level of intellectual attainment (e.g., Blascovich, Spencer, Quinn, & Steele, 2001; Croizet & Claire, 1998; Good, Aronson, & Inzlicht, 2003). Although not all these demonstrations are in the domain of race, they are important to acknowledge in any characterization of the relevance of stereotype threat to real-world intellectual achievement.

The second point is that Steele, Aronson, and their colleagues have long emphasized that stereotype threat is not the sole contributor to the racial achievement gap. Poverty, parental style, socialization, and so on, also play a role. Indeed, Steele (1997) listed structural and cultural threats as the first factors to consider in understanding race differences in academic identification, and he presented stereotype

threat as a factor to consider "beyond these threats" (p. 616).

Most important, the focus of stereotype threat is to explain the *residual,* that portion of variance left over in the racial achievement gap after prior preparation and skills (as roughly assessed by prior indicators such as college board scores) have been controlled. As Bowen and Bok (1998) documented, even at the most selective universities, there is a large race gap disfavoring Black students in graduation rates, grade point average, and class rank even after controlling for SAT scores, socioeconomic status, and high school grades (see also Jensen, 1980). Likewise, there is a large racial gap in SAT performance—of about 150 points—at every level of socioeconomic status as measured by family income (Hacker, 1995). This residual gap is called the "overprediction" or "underachievement" phenomenon. It is a large and persistent difference in scholastic success between the races that occurs even when extraneous factors are controlled. It is this gap that stereotype threat is aimed at explaining (see Steele & Aronson, 1995). It is this gap that has garnered the attention of the social sciences more generally. And it is this gap that Steele and Aronson observed in the ability-diagnostic condition of their studies when, even after controlling for student SAT scores, a large difference in test performance was found between Black students and White students. Given the context of the problem, controlling for prior differences in SAT is a perfectly appropriate laboratory analog to the real-world issue at hand.

The third point not conveyed by Sackett et al. (2004) is the existence of a growing body of work suggesting that the theoretical insights offered by stereotype threat can be applied to close the racial achievement gap in real classroom settings (Steele, 1997; see also Aronson, Fried, & Good, 2002; Good et al., 2003). When the perceived relevance and salience of negative stereotypes are reduced, African American students have been found to perform significantly better in school, sometimes dramatically. The utility of stereotype threat is the strongest gauge of its relevance and validity vis-à-vis understanding real race differences in intellectual attainment.

REFERENCES

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38,* 113–125.

Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. (2001). African Americans and high blood pressure: The role of stereotype threat. *Psychological Science, 12,* 225–229.

Bowen, W. G., & Bok, D. (1998). *The shape of the river: Long-term consequences of considering race in college and university admissions.* Princeton, NJ: Princeton University Press.

Croizet, J. C., & Claire, T. (1998). Extending the concept of stereotype and threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin, 24,* 588–594.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology, 24,* 645–662.

Hacker, A. (1995). *Two nations: Black and white, separate, hostile, unequal.* New York: Ballantine.

Jensen, A. R. (1980). *Bias in mental testing.* New York: Free Press.

Sackett, P. R., Hardison, C. M., & Cullen, M. J. (2004). On interpreting stereotype threat as accounting for African American–White differences on cognitive tests. *American Psychologist, 59,* 7–13.

Steele, C. M. (1997). A threat in the air: How stereotypes shape the intellectual identities and performance of women and African Americans. *American Psychologist, 52,* 613–629.

Steele, C. M., & Aronson, J. A. (1995). Stereotype threat and intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797–811.

Correspondence concerning this comment should be addressed to Geoffrey Cohen, Department of Psychology, Yale University, 2 Hillhouse Avenue, Box 208205, New Haven, CT 06520-8205. E-mail: geoffrey.cohen@yale.edu

DOI: 10.1037/0003-066X.60.3.271

On Interpreting Research on Stereotype Threat and Test Performance

Paul R. Sackett and Chaitra M. Hardison
University of Minnesota, Twin Cities Campus

Michael J. Cullen
Personnel Decisions Research Institutes

We are gratified that our article (Sackett, Hardison, & Cullen, January 2004) prompted these thoughtful reactions. We offer comments on each. Wicherts (2005, this issue) focused on the assumptions underlying the use of analysis of covariance and noted possible violations of these assumptions in the use of prior test scores as covariates in stereotype threat research. We agree that violations are possible, though at this point the likelihood and impact of these violations are matters of speculation. We take this opportunity to note that including a prior test score as a covariate is not a critical element of testing stereotype threat theory. There are sound reasons for including a prior test score as a covariate, such as increasing the power to detect an effect via reduction of the error term. However, reporting results without the covariate would permit a straightforward examination of subgroup differences in threat versus nonthreat conditions.

A theme of both the Helms (2005, this issue) comment and the Cohen and Sherman (2005, this issue) comment is that studies other than the Steele and Aronson (1995) work that was the focus of our article are important for understanding the effects of stereotype threat. We fully agree that it is important to consider the full range of research on stereotype threat. We certainly do not believe that the question of the effects of stereotype threat on test scores in high-stakes settings is settled; in fact, our own work continues to explore this question (e.g., Cullen, Hardison, & Sackett, 2004). But we also urge careful examination of the research studies cited by Helms and by Cohen and Sherman.

Helms (2005) presented a reanalysis of data from McKay, Doverspike, Bowen-Hilton, and McKay (2003) and proposed a regression-based mediation analysis as a means of testing whether stereotype threat accounts for group differences in mean test scores. She reported data showing that race has no effect on test scores once measured stereotype threat is controlled, which leads her to the conclusion that in this data set, threat does explain the racial group difference. We raise two concerns here about the procedure and her interpretation of the McKay et al. (2003) data. First, readers may be confused because of an initial misstatement about the requirements for mediation. Helms stated that "for stereotype threat to account for racial-group differences in test scores, measures or manipulations of it only have to account for at least as much variance in racial group as racial group explains in test scores" (p. 269). In fact, the requirement for full mediation is not that race and threat covary to at least the same degree as race and test scores covary but rather that race accounts for the same variance in test scores as is accounted for by threat. In other words, the effect of race on test performance is transmitted via stereotype threat, such that race has no effect on test score once stereotype threat is controlled.