

When White Men Can't Do Math: Necessary and Sufficient Factors in Stereotype Threat

Joshua Aronson, Michael J. Lustina, Catherine Good, and Kelli Keough

University of Texas, Austin

and

Claude M. Steele and Joseph Brown

Stanford University

Received April 9, 1998; revised August 1, 1998; accepted August 4, 1998

Research on “stereotype threat” (Aronson, Quinn, & Spencer, 1998; Steele, 1997; Steele & Aronson, 1995) suggests that the social stigma of intellectual inferiority borne by certain cultural minorities can undermine the standardized test performance and school outcomes of members of these groups. This research tested two assumptions about the necessary conditions for stereotype threat to impair intellectual test performance. First, we tested the hypothesis that to interfere with performance, stereotype threat requires neither a history of stigmatization nor internalized feelings of intellectual inferiority, but can arise and become disruptive as a result of situational pressures alone. Two experiments tested this notion with participants for whom no stereotype of low ability exists in the domain we tested and who, in fact, were selected for high ability in that domain (math-proficient white males). In Study 1 we induced stereotype threat by invoking a comparison with a minority group stereotyped to excel at math (Asians). As predicted, these stereotype-threatened white males performed worse on a difficult math test than a nonstereotype-threatened control group. Study 2 replicated this effect and further tested the assumption that

Joshua Aronson and Michael Lustina, Department of Educational Psychology, University of Texas at Austin. Kelli Keough, Department of Psychology, University of Texas at Austin. Claude Steele and Joseph Brown, Department of Psychology Stanford University. This research was supported by a Spencer Fellowship and a James S. McDonnell Fellowship to Joshua Aronson and by National Institutes of Health Grant MH51977 to Claude Steele. We thank Julie Cantor, and Jayne-Ann Harder for their assistance with this research and Stacey Rosenkrantz, Paul Nail, and Jennifer Crocker for comments on an earlier draft of the manuscript.

Correspondence concerning this article and reprint requests should be addressed to Joshua Aronson, Department of Educational Psychology, University of Texas at Austin, Austin, Texas 78712-1288. E-mail: joshua.aronson@mail.utexas.edu.

stereotype threat is in part mediated by domain identification and, therefore, most likely to undermine the performances of individuals who are highly identified with the domain being tested. The results are discussed in terms of their implications for the development of stereotype threat theory as well as for standardized testing. © 1999 Academic Press

One's reputation, whether false or true, cannot be hammered, hammered, hammered into one's head without doing something to one's character.

—Allport (1954, p. 142)

For some reason I didn't score well on tests. Maybe I was just nervous. There's a lot of pressure on you, knowing that if you fail, you fail your race.

—Rodney Ellis, African-American State Senator (Texas) in a 1997 interview

Members of stereotyped groups often feel extra pressure in situations where their behavior can confirm the negative reputation that their group lacks a valued ability (see Aronson, Quinn, & Spencer, 1998b); Steele, 1997, for reviews). We call this pressure "stereotype threat" and argue that in the short term, it can undermine the intellectual performance of virtually anyone whose group is targeted by stereotypes alleging a lack of intellectual ability in some domain (Steele & Aronson, 1995). We have also argued that stereotype threat can prompt a long-term defense against the chronic exposure to ability impugning stereotypes and the low performance that it can provoke—a disengagement or "disidentification" from the threatened domain, a dropping of the domain as a basis of self-esteem (see Steele, 1992, 1997; Steele & Aronson, 1995). The current research focuses on the short-term effects of stereotype threat in an effort to better understand the conditions under which stereotypes impugning intellectual ability are likely to interfere with intellectual test performance.

Empirical support for our contention that stereotype threat can affect the member of nearly any stereotyped social group is now abundant. Steele and Aronson (1995) found, for example, that African-American college students were dramatically affected by stereotype threat conditions; they performed significantly worse than whites on a standardized test when the test was presented as a diagnosis of their intellectual abilities, but about as well as whites when the same test was presented as a nonevaluative problem solving task. When the test was framed as diagnostic, Steele and Aronson hypothesized, the possibility of confirming the well-known stereotype of African-American intellectual inferiority became salient, and thus disruptive. A number of studies have found that women, too, perform less well when the societal stereotype that they face—low math ability—is made relevant by experimental instructions (Aronson, Good, & Harder, 1998a; Shih, Pitinski, & Ambady, in press; Spencer, Steele, & Quinn, 1999). Aronson and Salinas (1997) have found virtually the same results with Latino students, who also face the stereotype that their group lacks scholastic ability, as have Croizet and Claire (1998) in a study involving participants of low socioeconomic status. Finally, Levy (1996) has demonstrated how the cognitive functioning of elderly individuals can be disrupted by stereotype threat. When the

elderly participants in her study were subtly primed with the stereotype regarding old age and senility, they performed worse on a test of short-term memory than when they were primed with the more positive “old-people-are-wise” stereotype instead.

That intellectual performance can be spoiled by conditions that make ability stereotypes relevant and improved by conditions that nullify them, and that this occurs across a range of social groups certainly encourages a situationist—or at least person–situation *interactionist*—explanation for the academic underperformance of stereotyped groups (Steele, 1997; Steele & Aronson, 1995). Our analysis, we believe, provides a more hopeful alternative to the standard accounts of minority underachievement, which cite such intractables as poverty (e.g., Bereiter & Engleman, 1966; White, 1982), genetic differences in intelligence (e.g., Benbow & Stanley, 1980; Herrnstein & Murray, 1994; Jensen, 1980), or cultural and societal barriers to skill acquisition (e.g., Hunt, 1969).

SITUATIONAL PRESSURE OR INTERNALIZED INFERIORITY?

But an important question remains. Must individuals belong to a minority group—or be chronically targeted by stereotypes—in order to experience stereotype threat and suffer its effects? Undoubtedly, the degree to which a person is exposed to stereotypes about his or her group breeds an awareness of stigma, and such “stigma-consciousness” has been linked with individual differences in responses to stereotype threat (e.g., Lustina & Aronson, 1998; Pinel, in press). Nonetheless, a stigmatized identity may not be necessary to suffer its effects because, in theory, stereotype threat derives its power from a motive common to all individuals, regardless of their race, gender, socioeconomic status, age, and so on—the motive to sustain a self-image of goodness or competence and of being able to secure important outcomes (e.g., Steele, 1988). This motive dramatically influences behavior in performance contexts, where people have been shown to do what they can to make themselves feel that their prospects are good and to project this image of competence to others (e.g., Jones, 1989). This is best accomplished, of course, by performing well, but even when performance is not optimal, individuals enjoy numerous ways of appearing competent or at least appearing not to care (Steele, 1992; Major & Schmader, 1998).

Stereotype threat arises when these performance motives are jeopardized by the awareness of an ability-impugning stereotype in a situation where that stereotype can be confirmed by low performance. Thus, because most people are motivated to feel and to appear competent, nearly anyone, we believe, can experience the pressure of stereotype threat in some situation and thus suffer the short-term consequence of impaired intellectual performance (Crocker, Major, & Steele, 1997).

By showing how subtle situational factors can dramatically affect the performance of minority students, the existing stereotype threat studies make a strong case for the environmental basis of their underperformance. Still, the current body of evidence does not rule out a plausible nonsituational explanation. Specifically,

there remains the possibility that there is something special about being Black, Latino, a woman, poor, or old that made these test takers underperform when confronted with stereotypes about their group. Could the performance debilitating effect of stereotypes conceivably depend upon long-term exposure to devaluing stereotypes, real feelings of inferiority that have been “hammered into one’s head” by persistent stigmatizing conditions? Various theorists have endorsed such a view, arguing that stigmatizing treatment can result in a stigmatized personality (e.g., Cooley, 1956; Mead, 1934). Indeed, the social scientists whose opinions served as a linchpin in the 1954 Supreme Court case that ended racial segregation in the schools (*Brown v. Board of Education*) saw internalized inferiority as a necessary consequence of prejudicial treatment (Allport, 1954; Cook, 1979; Gerard, 1983).

Applied to the results of the stereotype threat research, the internalized inferiority interpretation would suggest that in all of these studies, the testing situation merely brings to the surface deep-seated feelings of inferiority or low expectations that have become an unchanging part of the individual whose exposure to devaluing stereotypes has left inferiority as a permanent mark (e.g., Howard & Hammond, 1985; S. Steele, 1990). Thus, although research has shown that stereotypes can undermine the performance of ability-stigmatized groups like African-Americans, Latinos, and women, it is not clear whether belonging to a minority group is a necessary or merely sufficient factor in this underperformance.

THE PRESENT HYPOTHESIS

The present research, then, was aimed at testing the situationist hypothesis of stereotype threat phenomena. Specifically, we put to the test our contention that virtually anyone could be made to underperform on a difficult intellectual test if they were exposed to a stereotype that predicted underperformance for their group. In both Studies 1 and 2 we examined the intellectual test performance of the social group we deemed most unlikely to have internalized stereotype-based feelings of intellectual inferiority—white males selected on the basis of their high abilities. If these highly skilled majority-group members could be threatened by a stereotype alleging their relative inferiority, then it seems reasonable to assume that ingrained feelings of inferiority need not be involved in stereotype threat. In Study 2 we sought to replicate this effect, while at the same time testing a second assumption about the short-term effects of stereotype threat—that to underperform as a result of activating a stereotype alleging low ability, the individual must be self-invested in the ability domain being tested.

STUDY 1

Overview

White males with high scores on the mathematics section of the Scholastic Aptitude Test (SAT) took a very challenging math test. In one condition of this experiment we explicitly confronted them (before the test) with the stereotype that Asian students outperform Caucasian students in mathematical domains. In a

control condition no mention of the stereotype was made. The general prediction was that, compared to those not explicitly reminded of the Asian stereotype (control condition), the stereotype threatened test takers would perform less well.¹

Participants

Potential participants, drawn from the Stanford University student body, completed a questionnaire designed to identify students with strong math skills and who attached at least a moderate degree of importance to these skills. Students were eligible for the study if they: (1) indicated that their ethnicity or race was white/Caucasian or Jewish; (2) responded that they were neutral about, agreed with, or strongly agreed with the following two statements: “Math is important to me” and “I am good at math”; and, (3) scored 610 or above (of a possible 800) on the math section of the SAT. The average math SAT score of the participants was 712.17 ($SD = 60.6$), and the scores ranged from 610 to 800. Twenty-three qualifying male undergraduates participated in this study for pay or for course credit.

Materials and Procedure

Participants were greeted individually by a white female experimenter who explained that the study she was conducting had to do with differences in math ability. Those participants who were randomly assigned to the stereotype condition ($n = 12$) were given 2 min to skim over a packet of articles about the phenomenal math achievement of Asians and were told that the study was specifically concerned with understanding why Asians appear to outperform other students on tests of math ability. Taken from national newspapers and prominent journals, these articles emphasized a “growing gap in academic performance between Asian and white students.” The rationale of the experiment, coupled with the articles, we reasoned, would accomplish two things—make the participants feel targeted by a stereotype relevant to their math ability and give some plausibility to the stereotype. The titles strongly suggested that Asians are better at math than Caucasians, though they offered no explanation for the superiority.

At the end of the 2-min interval, the participants in the stereotype condition were further told that there seems to be a growing discrepancy between the academic performance of Asians and whites. “In math,” they were told, “it seems

¹ This study actually included two additional stereotype threat conditions in which we coupled the Asian Stereotype manipulation with some additional information we thought might moderate the effect of the stereotype on performance. In one condition, we attempted to nullify the stereotype by suggesting that the test at hand was not known to reveal ethnic differences in the past; in the other, we attempted to magnify the effect by presenting the test as having revealed them in the past (see Spicer, Steele, & Quinn, this issue). The performance results of these conditions were identical to those of the condition in which only the stereotype was presented. Why these additional statements failed to moderate the effect is unclear, but we suspect the reason may be that the manipulation of stereotype threat (i.e., the news articles) may have been too vivid and powerful to have been modified much by the addition of the more pallid verbal instructions. Because these conditions are not relevant to the central question posed by this study, they will not be further discussed.

to be the case that Asians outperform whites.” The experimenter then implied that research findings are inconclusive and that the purpose of the study was to learn more about the nature and scope of these differences. Participants assigned to the “control” condition ($n = 11$) did not read the articles or hear any mention of Asian–white ability differences; they were told only that the test was a measure of their math ability.

The participants were then given 20 min to take the test, which was administered on a computer. The test was composed of 18 questions derived from the Graduate Record Examination (GRE) mathematics subject test. The computer tracked the time each participant spent on each of the questions and recorded the participant’s answers.

At the end of the 20-min time period, participants first completed a measure of state anxiety, a modified form of the state form of the State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970). They then completed a questionnaire that asked how much effort they expended on the task, how difficult they found the problems, how much pressure they felt, how much confidence they had in their answers, and how many problems they thought they had solved correctly. The participants were then fully debriefed and compensated for their participation.

Results and Discussion

Test performance. The measure of test performance was simply the number of items correctly solved, which we submitted to a one-way analysis of variance (ANOVA). The results revealed the predicted effect of stereotype threat on test performance, $F(1, 21) = 5.51, p < .01$.² Participants solved fewer of the items in the stereotype threat condition ($M = 6.55$) than in the control condition ($M = 9.58$).

Follow-up questionnaire. What processes mediated the underperformance? No condition differences were found on measures of anxiety, time spent on items, or the self-reported difficulty of the items (All p ’s n.s.). But, the stereotype threat group did report expending more effort on the problems ($p < .05$), suggesting that the stereotype may have boosted their motivation, not undermined it. Thus, the underperformance could stem from trying too hard, rather than not hard enough. This finding is consistent with the analysis provided by Steele and Aronson (1995), who concluded that withdrawal of effort was not the cause of the underperformance of the African–American students in their studies. However, an analysis of covariance (ANCOVA) performed on participants test scores failed to support this reasoning; the difference between means is unchanged when we correct for self-reported effort. Thus, as in many studies of stereotype threat, the mediator between stereotype threat and performance is unclear.

The performance results are nonetheless quite supportive of our general hypothesis that making salient the Asian stereotype would depress the perfor-

² An analysis of covariance with participants’ quantitative SAT scores used as the covariate was also performed, yielding the same significant pattern of results.

mance of a group of nonstereotyped, high ability students. Thus, contrary to the prediction derived from the internalized inferiority view of group differences in performance (e.g., Allport, 1954; Howard & Hammond, 1985; S. Steele, 1990), stereotype-related underperformance does not appear to require the existence of doubts drummed in by chronically stigmatizing conditions or by minority status. Study 2 was undertaken to replicate this effect, but also to examine a proposition about what *is* required to induce stereotype threat.

STUDY 2

As we have noted elsewhere (Aronson et al., 1998b; Steele, 1997) our formulation has assumed from the outset that stereotype threat will have little if any effect on individuals who are not identified with the ability domain in question. To be threatened by the self-evaluative implications of a stereotype that alleges low ability of some kind, a person probably needs to either care about having the ability or at least care about the social consequences of being seen as lacking the ability (Brunstein & Gollwitzer, 1996; Steele, 1992, 1997). Thus, in an attempt to find evidence of stereotype threat we have selected students, at least in part, on the basis of their identification with some skill. Although internal analyses of past experiments and pilot studies suggest that stereotype threat has little effect on the unidentified (see Steele, 1997), no research has been conducted that examines domain identification as a factor independent of ability and confidence in that domain. Thus, no satisfactory test of our reasoning exists.

Directly examining the role of identification in stereotype threat processes is important for at least two reasons. First, it may help in making more accurate predictions about which individuals will be likely to most acutely experience stereotype threat, as well as the settings in which they will be at risk. Second, demonstrating the link between domain identification and the experience of stereotype threat will inform and strengthen our reasoning about the longer term self-protective consequences of stereotype threat on the identification process. Specifically, we (Aronson et al., 1998b; Steele, 1992, 1997; Steele & Aronson, 1995) have argued that to protect oneself against the chronic experience of stereotype threat in an academic domain, individuals *disidentify* with the domain, often with serious consequences for their motivation and achievement. In line with this reasoning, there is increasing evidence that ability-stigmatized groups (e.g., African-Americans) are more prone than their nonstigmatized counterparts to disidentify from academics (Aronson & Fried, submitted for publication; Major et al., 1997; Osbourne, 1995). Yet the link between this identification and stereotype threat is unclear. This reasoning would be much strengthened by finding that it is those most identified with a particular academic domain who are most prone to vulnerability to stereotypes alleging limited prospects therein.

Study 2 directly tested this hypothesis by conceptually replicating the previous study with the additional subject factor of students' degree of identification with mathematics. The prediction was straightforward: stereotype threat should be

most disruptive to the group selected for having high as opposed to moderate identification with mathematics.

Method

Overview. This experiment took the form of a 2×2 factorial design. The factors were math identification of the participant (high vs. moderate) and experimental condition (stereotype threat vs. control). As in Study 1, a math test was presented in the context of the Asian-math-superiority stereotype (stereotype threat condition) or without the stereotype (control condition). Test performance was the primary dependent measure.

Participants. Participants were white male students enrolled in the second semester of a rigorous year-long calculus course at the University of Texas at Austin. This course was selected for its high concentration of math-proficient students. Specifically, no student can enroll in the class with a QSAT score below 550. Furthermore, students are not allowed to enroll in the course's second semester without performing satisfactorily in the first.

Three weeks prior to the experimental session these students filled out a questionnaire during class regarding their math-related attitudes. There were 75 participants who filled out the questionnaire. These students were divided into three groups according to their responses to a question asking them to rate the importance of their math abilities to their self-concept. Scores could range from 1 (not at all important) to 15 (extremely important). Students responses to this item ranged from 4 to 15, ($M = 12.49$, $SD = 2.84$). Taking the top and bottom third of students on this measure produced a "high math-identified" group ($n = 26$, $M = 15$, $SD = 0$) and a "moderately identified" group ($n = 23$, $M = 10.27$, $SD = 2.12$), who had been randomly assigned to either the stereotype threat condition or the control condition of the experiment.³

Procedure. The testing occurred 3 weeks after the initial questionnaire, during students' weekly discussion section meetings, which were run by their calculus teaching assistant. Students were informed that they would receive extra credit on their homework grade for participating in this study, which was described as a national study of mathematics ability. The teacher and an experimenter handed out booklets that provided the manipulation of experimental condition, the test, and the dependent measures. The booklets were constructed so that the participant's experimental condition could not be seen by the teacher or experimenter.

Stereotype threat manipulation. Participants in the stereotype threat condition received an additional description of the study as an attempt to better understand

³ Because the bottom third of this sample was used it might be reasonable to think of the group as "low identified," but only in a relative sense. In absolute terms, both the group mean on the identification measure and the fact that only two of the participants had identification scores below the midpoint of the identification scale make it more appropriate to refer to these students as "moderately identified."

why Asians are superior to other groups in mathematics. Specifically, the participants read the following description:

As you probably know, math skills are crucial to performance in many important subjects in college. Yet surprisingly little is known about the mental processes underlying math ability. This research is aimed at better understanding what makes some people better at math than others. As you also may know, at some top schools, Asian students outnumber the white students in math majors and majors with math as a prerequisite, and there seems to be a growing gap in academic performance between these groups. A good deal of research indicates that Asians consistently score higher than whites on standardized tests of math. But thus far, there is not a good explanation for this. The research you are participating in is aimed at better understanding these differences. Your performance on the exam will be compared to other students from across the nation. One specific question is whether Asians are superior at all types of math problems or only certain types.

The control (no stereotype) group was run simultaneously. The only difference was in the paragraph that described the reason for the experiment. The control group read the following:

As you probably know, math skills are crucial to performance in many important subjects in college. Yet surprisingly little is known about the mental processes underlying math ability. This research is aimed at better understanding what makes some people better at math than others. Your performance on the exam will be compared to other students from across the nation.

Testing session and measures. We attempted to make the testing session feel similar to an actual standardized test administration such as one might encounter taking the SAT. After allowing participants time to read their test description, the teacher and the experimenter took the class through a sample problem and then allowed the students 20 min for the test.

The math test was developed specifically for use in this experiment by graduate students in the mathematics department. The questions were drawn from the math subject GRE test practice booklets that pilot testing revealed to be at the upper limit of these students' abilities. The test consisted of 15 calculus-related questions.

At the end of the time period, the students were told to stop the test and to proceed with the questionnaire packet. The packet contained the same questionnaire employed in Study 1.

Results and Discussion

Test performance. The primary hypothesis in this experiment was that participants would perform less well when confronted with the stereotype regarding Asian superiority in math, but that this would mainly be the case for those

⁴ As in the Study 1, an analysis of covariance was also performed and the pattern of the results and significance are nearly identical—indeed, they are more significant—when participants scores are corrected by their SAT scores.

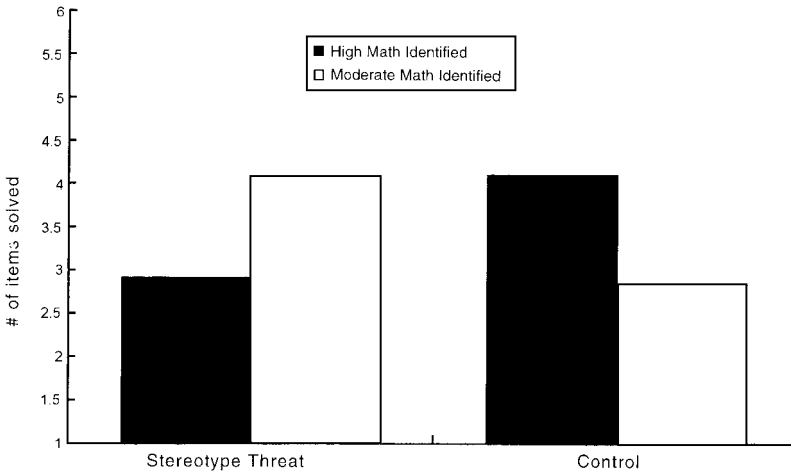


FIG. 1. Mean math test performance.

students who cared deeply about their math abilities—the high math-identified students. The ANOVA performed on their test scores offered strong support for this prediction. Only the math identification by experimental condition interaction was significant, $F(1, 45) = 9.66, p < .005$. As simple effects tests show, and as may be seen in Fig. 1, high math-identified participants performed less well on the test when the stereotype was mentioned ($M = 2.91$) than when it was not ($M = 4.10$), $t(45) = 2.08, p < .05$. Moderately identified participants showed precisely the opposite pattern, performing better when the stereotype was discussed ($M = 4.07$) than when it was not ($M = 2.83$), $t(45) = 2.325, p < .05$.

This pattern of data replicated the finding of study 1—and did so with a different procedure, participant population, and a different test—supporting our assumption that stigma is not necessary for stereotype threat to undermine performance. It also provides very clear evidence of the critically important role of domain identification in mediating stereotype threat.

Follow-up questionnaire. The ANOVA performed on participants' responses to the questions in the packet revealed only one significant difference, a significant interaction on the item measuring evaluation apprehension, $F(1, 45) = 4.96, p < .05$. Using a scale ranging from 1 ("I never had this thought") to 5 ("I had this thought very often"), high math-identified participants wondered more often what the experimenter would think of them in the stereotype threat condition ($M = 2.46$) than in the control condition ($M = 1.18$), $t(45) = 2.81, p < .01$. The stereotype threat manipulation had no apparent effect on evaluation apprehension among the moderately identified participants. They reported equal amounts regardless of whether they were in the stereotype threat condition ($M = 1.64$) or in the control condition ($M = 1.75$), $t < 1$.

There were no effects on any of the other distraction items, nor on the measures of state anxiety, effort, perceived performance, or confidence (all p 's, $> .2$). Thus,

lacking better evidence than participants' own self-reports, the best explanation for the effect of the stereotype on performance appears to be that the underperformance was caused by evaluation apprehension rather than by the withdrawal of effort or by measurable levels of anxiety. This finding is in contrast to that of past stereotype threat research (e.g., Aronson, 1998; Blascovich et al., 1998; Steele & Aronson, 1995), which points to the role of anxiety in undermining performance in stereotype threat conditions.

To further explore this evaluation apprehension interpretation, we submitted participants test scores to analyses of covariance (ANCOVA) using their self-reported evaluation apprehension ratings as the covariate. Although this weakened the effect somewhat, the interaction effect remained significant ($p = .02$), and the relationship between the adjusted means remained identical to that shown in Fig. 1. Thus, although there appears to be some mediational effect of evaluation apprehension, it does not appear to solely mediate the effect of stereotype threat on test performance.

As in previous studies, it is difficult to determine the precise mediation of stereotype threat effects. The difficulty stems in part from the self-report nature of the measures, but also from the fact that there are undoubtedly multiple mediational pathways through which psychological manipulations can affect performance. Even studies that have used direct measures such as blood pressure to show that anxiety accompanies stereotype threat (Blascovich et al., 1998) cannot confidently rule out the withdrawal of effort as a mediator of underperformance. Anxiety and effort withdrawal are not mutually exclusive. Indeed, they most likely work in tandem to undermine performance.

Looking within the control group, we find what one would reasonably expect on a mathematics test comparing students who are identified with math to students who are markedly less so. At the same level of ability and preparation, the high math-identified participants outperformed their low math-identified counterparts when not confronted with a stereotype alleging relative inferiority. But, the reverse pattern occurred when the stereotype was activated—moderately identified participants actually outscored high math-identified participants. This finding suggests that in stereotype threat testing situations, it may be an advantage to be moderately rather than extremely invested in a domain. The stereotype appeared to challenge these moderately identified students to do their best, but they were not so ego-involved that they were distracted, a finding quite consistent with much research on ego-involvement and performance (see Baumeister & Showers, 1986, for a review).

GENERAL DISCUSSION

The existence of negative stereotypes, we have argued, means that in situations where the stereotype is relevant, individuals who are targeted by stereotypes face the unpleasant predicament of confirming those stereotypes. One consequence of this predicament, many studies have shown, is interference with intellectual test performance. Taken together, the results of the present studies strongly suggest

that a person need not be chronically targeted by stereotypes to be impaired by them on tests that measure abilities of high personal importance. What appears to be necessary, on the other hand, is that an individual care enough about performing well to be bothered by a stereotype's implication that they may lack the ability to do so.

We can be fairly certain that our participants were not harboring deep-seated feelings of math inferiority—what Allport (1954) referred to as “traits due to victimization.” White males of high math ability clearly do not fit the profile of the “target of a stereotype” or “the disadvantaged minority student.” They are not normally considered at risk to be stereotyped, looked down upon because of their race, or targeted by low performance expectations. Indeed, in contrast to the cases of women and math, African-Americans and general intelligence, the elderly and forgetfulness, and so on, there is no stereotype directly targeting white males—alleging that “white men can't do math.” For example, consider a recent study in which male and female undergraduates (from numerous ethnic groups) were asked to list stereotypes about various groups (Aronson & Disko, 1998a). The study found that whereas 86% of the respondents mentioned the intellectual or academic prowess of Asians, and 30% listed the intellectual weakness of females, not a single participant listed stereotypes regarding the intellectual abilities of whites. Thus, the stereotype either does not exist or it exists only under such narrow circumstances that we would not expect it to be chronically “available” enough to become part of a stigmatized personality (Crocker et al., 1997).

Nonetheless, when placed in a situation where a minority group's relative superiority was made salient and relevant, highly skilled and identified white males experienced a decrement in intellectual performance—much like the members of groups for whom stereotypes regarding their intellectual abilities *do* exist and *are* widely known and cognitively available. Clearly, then, chronic feelings of stigmatization were not a necessary factor in their underperformance. Situational pressures alone—the stereotype about Asians coupled with the strong desire to perform well—were sufficient to interfere with performance.

“Direct” Versus “Indirect” Stereotype Targets

This by no means implies that the white males in these studies experienced the situation in *exactly* the same way or to the same degree as, say, women taking the same math test under stereotype threat conditions. Clearly there must be phenomenological differences that vary as a function of many factors. Otherwise, one would expect to see white males dropping out of math and science graduate programs—which are highly populated by Asian students—with the same frequency as women. According to the most authoritative study of attrition from math and science programs, this is simply not the case. Although white males do complain about the extra pressure engendered by the high concentration of Asian classmates, they are much less likely than women to disidentify and drop out of math and science fields (Seymour & Hewitt, 1997).

What makes the situation different for men than for women—or other stereo-

type threatened groups? Surely there are many reasons, many of which we believe to stem from the fact that the stereotype directly targets some groups and only indirectly targets others. In the context of math ability, for example, Asians and women are what we could call “direct” targets of a stereotype. The stereotype refers explicitly to them (e.g., “Asians are good at math, women are bad at math”) and corresponding expectations (e.g., “Asians will perform well on this math test, women will not”) may naturally arise whenever they are in a math-relevant situation. It is therefore easier to trigger and more difficult to suppress stereotype threat for them. White males, on the other hand, are “indirect” targets. The stereotype refers to them only by means of a *comparison* with the direct stereotype target. For this reason stereotype-related expectations may only arise in situations—like those of the present studies—where comparisons with direct targets are made explicit. In this, as in many contexts, white males are the implied standard—the norm from which direct stereotype targets are viewed as deviating (e.g., Miller, Taylor, & Buck, 1991).

So while indirect targets are certainly capable of feeling stereotype threatened in some circumstances, they are undoubtedly less likely than direct targets to experience its most pernicious effects (e.g., disidentification), for a number of reasons, a few of which we list here. First, direct targets are likely to have the stereotypes more cognitively available and thus be more easily threatened by them. Pinel (in press) refers to this heightened awareness of stereotypes as “stigma consciousness” and has shown that it predicts underperformance in intellectual testing situations (see also Lustina & Aronson, 1998). Second, in most cases, being a direct target means being distinctive—being readily identified by others as belonging to a particular group. Thus, direct stereotype targets may also tend to feel more identified with their groups (e.g., Brewer, 1991) and therefore more self-threatened than indirect targets by whatever the stereotype alleges. Past research has suggested that people who feel more identified with their group respond differently to stereotypes and prejudice than less identified individuals (e.g., Crocker & Major, 1989; Rosenkrantz, 1994). This greater sense of “groupness” may also make the direct target feel more responsible for representing their group and thus more keenly and chronically apprehensive about representing their group in situations where a group stereotype is relevant. And finally, as suggested by the looking-glass-self models of stigma, chronic exposure to stereotypes could engender self-doubts that are either chronically activated or easily brought to the fore in stereotype-relevant situations, making the direct target more vulnerable to stereotype threat.

It is certainly conceivable that indirect targets, such as white males in a math-intensive environment, could feel like direct targets. But even then, their vulnerability is likely to be mitigated by supports that direct targets do not have—namely, majority status. For instance, although white males pursuing degrees in math or science fields may feel threatened by the reputed superiority of Asians, their belongingness in the domain is nonetheless affirmed by the abundance of similar individuals in their ranks, a luxury that women in these fields do

not enjoy (Fulton, 1996). Moreover, as Crocker et al. (1997) point out, members of dominant groups may be stigmatized in one context, but they are buffered from that stigma because of their power position in the larger society. For these, and perhaps additional, reasons, it is understandable that indirect targets are quite capable of experiencing some of the short-term effects of stereotype threat without necessarily feeling the need to disidentify.

An intriguing implication of the direct–indirect distinction is that in performance situations, indirect targets may actually derive a benefit from comparisons with direct targets for whom stereotypes allege *inferior* ability. If stereotypes can cause white males to perform worse when they are made to focus on the abilities of Asians, might they also perform better when led to focus on the abilities of women? Spencer et al. (1999) have found some indirect support for this hypothesis in their studies of men and women taking math tests. The males in their study actually performed worse in conditions where the female stereotype was nullified by experimental instructions. Specifically, males tended to perform worse when told that the test was not expected to show gender differences, suggesting that their performance may be boosted by the implicit stereotype that they are superior to women in math.

Domain Identification and Test Performance

This research adds to the growing body of evidence that domain identification—the degree to which a person stakes their self-image on a given ability—is vitally linked to how people respond to failure (Brunstein & Gollwitzer, 1996) or to stereotypes about their abilities (Steele, 1997). But does this mean that stereotype threat will affect only those students who are highly identified with a domain? Steele (1997) has hinted at such an argument, suggesting that stereotype threat is likely to be felt most keenly among the “vanguard” of students targeted by stereotypes—those at the highest levels of ability and domain identification. The present research is certainly consistent with this reasoning—students who were more domain-identified tended to score less well under stereotype threat. But does this mean that targets who are not in the vanguard are invulnerable?

Not necessarily. It is quite possible that high degrees of domain identification are indeed necessary for stereotypes to become self-threatening in the relatively low stakes setting of the typical laboratory experiment, where the consequences of low performance are primarily self-imposed. After all, psychology experimenters lack “fate control.” They do not punish the low performing test taker with low course grades or denied access to college or graduate school. Lacking such consequences, the performance experiment is a high-stakes endeavor for only those participants who are so identified that *any* test of their ability is ego-involving.

In the real world, however, ability tests can be ego-involving even in cases where a person is not particularly identified with a domain because there are self-threatening consequences to underperformance. For example, a woman seeking an advanced degree in art history may be only marginally identified with

her math abilities, yet is nonetheless required to score well on the mathematics portion of the GRE to gain acceptance to graduate school. In such cases, low-identified students may be every bit as debilitated—and perhaps more so—by the relevance of a stereotype alleging low math ability. It therefore may be more correct to say that high motivation—a sense that something important is at stake—is the necessary factor in stereotype threat, not high identification per se. To be sure, the sense of high stakes is undoubtedly greatest among those who are most identified with the domain, but stereotype threat may nonetheless affect the test performances of average students, not just those in the vanguard (Aronson & Disko, 1998b).

The relationship between domain identification and test performance in stereotype threat situations may contribute to the poor predictive validity of college and graduate school admissions tests (e.g., Sternberg & Williams, 1997). Specifically, because extra pressure can cause a person to underperform, it is something of an advantage to be *moderately* identified rather than to be *extremely* so. But in most academic situations the relationship between domain identification and achievement will be linear; the more one cares, the harder one will study, the more regularly one will attend class, think through ideas, and so on. Thus, scores on high-stakes tests like the SAT may be rendered less accurate because, as in Study 2, test takers can be penalized for their devotion. Thus, we might well question the overreliance on standardized tests as gateways to higher education—certainly in the case of ability-stigmatized minority groups—especially when other measures exist which do not subtract points for caring.

What is Stereotype Threat?

We began this report with a quote from Rodney Ellis, a highly regarded African-American State Senator who “knew he was just as intelligent as the next guy,” but attributed his low performance on standardized tests to apprehensions about “failing his race.” The current evidence regarding stereotype threat does not allow us to know if Ellis explains or merely describes the process. Is stereotype threat self-threatening because it arouses a fear of being a bad ambassador of one’s group to mainstream society? Or is it more simply the apprehension about appearing incompetent—for the sake of one’s own reputation? Or alternatively, is it merely the result of worrying that one might lack ability? Or is it some combination of these concerns? These are important questions that will have to await the results of future research for answers.

What these studies do make clear is that, whether the fear is interpersonal or intrapersonal, motivated by staining one’s group or merely one’s self, it need not arise out of a chronic stigmatization. It is sufficient to be identified enough with a domain to be threatened by the possibility of limited prospects there and unlucky enough to be on the wrong end of a stereotype about an intellectual ability. And, clearly, if stereotype threat can be aroused in highly able, nonstereotyped students merely by making them aware of a stereotype that predicts lower performance for

their group relative to another, then it is not some exotic phenomenon felt only by the members of historically stigmatized groups.

As we have elsewhere argued (Aronson, Quinn, & Spencer, 1998; Steele, 1997; Steele & Aronson, 1995), this situationist view of minority underperformance is an encouraging one because it locates the problem not exclusively within the person, but within the social circumstances confronting the person. Stereotype threat research underscores how changing those circumstances, even subtly, can have dramatically positive effects on performance. The present studies carry this message a step further by suggesting that stereotype threat does not uniquely affect members of certain minority groups. Rather, it is a predicament that stems from quite normal responses to the low or demeaning expectations that come to the individual in the form of cultural stereotypes. Moreover, the critical role of domain identification suggests that these responses are based on self-protective processes that can be triggered, and perhaps intensified, by minority status and identity, but which do not depend upon them. This may prove to be useful knowledge because it may point toward ways of reducing stereotype threat that involve self-protective tactics (e.g., Aronson, 1998; Josephs and Scroeder, 1997) rather than the less tractable undertaking of eliminating cultural stereotypes.

REFERENCES

- Allport, G. (1954). *The nature of prejudice*. New York: Addison-Wesley.
- Aronson, J. (1998). *The effects of conceiving ability as fixed or improveable on responses to Stereotype Threat*. Manuscript submitted for publication.
- Aronson, J., & Disko, D. (1998a). *Frequency of stereotypes about blacks, whites, Asians and Latinos*. Unpublished raw data, Univ. of Texas, Austin.
- Aronson, J., & Disko, D. (1998b). *A test of the vanguard hypothesis in stereotype threat*. Unpublished raw data, Univ. of Texas, Austin.
- Aronson, J., & Fried, C. (1998). *Belief in the malleability of intellectual ability: An intervention to increase the performance of African-American college students*. Manuscript submitted for publication.
- Aronson, J., Good, C., & Harder, J. A. (1998a). *Stereotype threat and women's calculus performance*. Unpublished manuscript, Univ. of Texas, Austin.
- Aronson, J., & Salinas, M. F. (1997). *Stereotype threat, attributional ambiguity, and Latino underperformance*. Unpublished manuscript, Univ. of Texas.
- Aronson, J., Quinn, D., & Spencer, S. (1998b). Stereotype threat and the academic underperformance of minorities and women. In J. Swim and C. Stangor (Eds.), *Prejudice: The target's perspective*. Academic Press.
- Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology*, *16*(4), 361-383.
- Benbow, C. P., & Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or artifact? *Science*, *210*, 1262-1264.
- Bereiter, C., & Engleman, S. (1966). *Teaching disadvantaged children in the preschool*. Englewood Cliffs, NJ: Prentice Hall.
- Blascovich, J., Quinn, D. M., Spencer, S. J., & Steele, C. M. (1998). *The effect of stereotype threat on African Americans' blood pressure*. Unpublished manuscript, University of California, Santa Barbara.
- Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin*, *17*(5), 475-482.

- Brunstein, J. C., & Gollwitzer, P. M. (1996). Effects of failure on subsequent performance: The importance of self-defining goals. *Journal of Personality and Social Psychology*, **70**(2), 395–407.
- Cook, S. W. (1979). Social science and school desegregation: Did we mislead the Supreme Court? *Personality and Social Psychology Bulletin*, **5**(4), 420–437.
- Cooley, C. H. (1956). *Human nature and the social order*. New York: Free Press.
- Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, **96**, 608–630.
- Crocker, J., Major, B., & Steele, C. (1997). Social stigma. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (Vol. 4). Boston: McGraw–Hill.
- Croizet, J. C., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, **24**, 588–594.
- Fulton, J. D. (1996). AMS-IMS-MAA Annual Survey (First Report). *Notices of the American Mathematical Society*, **43**(12), 1493–1511.
- Gerard, H. (1983). School desegregation: The social science role. *American Psychologist*, **38**, 869–878.
- Herrnstein, R., & Murray, C. (1994). *The bell curve*. New York: Free Press.
- Howard, J., & Hammond, R. (1985, September 9). Rumors of inferiority. *New Republic*, **72**, 18–23.
- Hunt, J. McV. (1969). *The challenge of incompetence and poverty*. Urbana, IL: Univ. of Illinois Press.
- Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.
- Jones, E. E. (1989). The framing of competence. *Personality and Social Psychology Bulletin*, **15**, 477–492.
- Josephs, R., & Schroeder, D. (1997). *The self-protective function of the learning curve*. Unpublished manuscript, University of Texas at Austin.
- Levy, B. (1996). Improving memory in old age through implicit self-stereotyping. *Journal of Personality and Social Psychology*, **71**, 1092–1107.
- Lustina, M., & Aronson, J. (1998). *Measuring and predicting stereotype vulnerability*. Unpublished manuscript. Univ. of Texas at Austin.
- Major, B., & Schmader, T. (1998). Coping with stigma through psychological disengagement. In J. Swim & C. Stangor (Eds.), *Stigma: The target's perspective*. New York: Academic Press.
- Major, B., Spencer, S., Schmader, T., Wolf, C., & Crocker, J. (1997). Coping with negative stereotypes about intellectual performances: The role of psychological disengagement. *Personality and Social Psychology Bulletin*, **24**(1), 34–50.
- Mead, G. H. (1934). *Mind, self, and society*. Chicago: Univ. of Chicago Press.
- Miller, D. T., Taylor, B., & Buck, M. L. (1991). Gender gaps: Who needs to be explained? *Journal of Personality and Social Psychology*, **61**(1), 5–12.
- Osbourne, J. W. (1995). Academics, self-esteem, and race: A look at the underlying assumptions of the identification hypothesis. *Personality and Social Psychology Bulletin*, **21**, 449–455.
- Pinel, E. C. (in press). Stigma-consciousness: The psychological legacy of social stereotypes. *Journal of Personality and Social Psychology*.
- Rosenkrantz, S. L. (1994). *Attributional ambiguity among Mexican-Americans: The role of ethnic identity*. Unpublished doctoral dissertation, Stanford Univ.
- Sarason, I. G. (1980). Introduction to the study of test anxiety. In I. G. Sarason (Ed), *Test anxiety: Theory, research, and applications* (pp. 3–13). Hillsdale, NJ: Erlbaum.
- Seymour, E., & Hewitt, N. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.
- Shih, M., Pitinski, T. L., & Ambady, N. (in press). Shifts in women's quantitative performance in response to implicit sociocultural identification. *Psychological Science*.
- Speilberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the state-trait anxiety inventory*. Palo Alto, CA: Consulting Psychologist Press.
- Spencer, S. J., Steele, C. M., & Quinn, D. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, **35**, 4–28.

- Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261–302). Hillsdale, NJ: Erlbaum.
- Steele, C. M. (1992). Race and the schooling of black Americans. *The Atlantic Monthly*, April.
- Steele, C. M., & Aronson, J. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, **52**(6), 613–629.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, **69**(5), 797–811.
- Steele, S. (1990). *The content of our character*. New York: St. Martin's Press.
- Sternberg, R. J., & Williams, W. W. (1997). Does the Graduate Record Examination predict meaningful success in the graduate training of psychologists? *American Psychologist*, **52**(6), 630–641.
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, **91**(3), 461–481.