

Unleashing Latent Ability: Implications of Stereotype Threat for College Admissions

Christine R. Logel

*Renison University College
University of Waterloo*

Gregory M. Walton

*Department of Psychology
Stanford University*

Steven J. Spencer, Jennifer Peach, and Zanna P. Mark

*Department of Psychology
University of Waterloo*

Social-psychological research conducted over the past 15 years provides compelling evidence that pervasive psychological threats are present in common academic environments—especially threats that originate in negative intellectual stereotypes—and that these threats undermine the real-world academic performance of non-Asian ethnic minority students and of women in math and science. As a consequence, common measures of academic performance, including both grades and test scores, systematically underestimate the intellectual ability of ethnic minority students and of women in quantitative fields (Walton & Spencer, 2009). We review evidence for these psychological threats, discuss their implications for the meaning and interpretation of common performance measures used in important admissions decisions, and address their implications for the efforts of colleges and universities to create positive academic environments that allow all students to thrive.

Clearly something is missing from our understanding of black underachievement. Disadvantage contributes, yet blacks underachieve even when they have ample resources, strongly value education, and are prepared better than adequately in terms of knowledge and skills. Something else has to be involved. That something else could be of just modest importance—a barrier that simply adds its effect to that of other disadvantages—or it could be pivotal, such that were it corrected, other disadvantages would lose their effect. (Steele, 1992, p. 76)

In his seminal article, “Race and the Schooling of Black Americans,” Claude Steele (1992) argued that although the underachievement of African Americans in the U.S. educational system is ubiquitous, it can be overcome. As subsequent research has shown, one cause of the academic underachievement of African Americans and other negatively stereotyped groups, such as women in math and science,

involves the psychological environment in which learning and testing take place. Common academic environments require these students to perform well or risk confirming negative intellectual stereotypes about their group. Because of this “stereotype threat,” non-Asian ethnic minority students’ grades and standardized test scores, and women’s grades and test scores in math and science, systematically underestimate their intellectual ability and potential (Walton & Spencer, 2009). When academic environments are altered to reduce stereotype threat, negatively stereotyped students perform better than would be expected based on their prior level of academic performance. This phenomenon, termed *latent ability*, suggests that stereotyped students’ prior performances underestimate the full extent of their academic ability—that their ability is in part hidden on these common assessments.

An important implication of this research is that decisions about how students are admitted to selective schools—which candidates for admission are most qualified for, and most likely to succeed in, a given school—are fundamentally

Correspondence should be addressed to Christine R. Logel, Renison University College, 240 Westmount Road North, Waterloo, Ontario, N2L 3G4, Canada. E-mail: clogel@uwaterloo.ca

connected to the psychological environment present in those schools. Indeed, we suggest that the quality of a school environment is as much a selection parameter as the decision of what criteria to consider when admitting students. By building high-quality school environments that allow all students to achieve to their potential, selective schools can revise their admissions criteria and procedures to more accurately reflect students' abilities and potential. In this article, we review evidence for the effects of stereotype threat on students' intellectual performance and discuss implications for the efforts of colleges and universities to create positive academic environments and to make more appropriate admissions decisions.

STEREOTYPE THREAT AND THE ACADEMIC UNDERPERFORMANCE OF MEMBERS OF NEGATIVELY STEREOTYPED GROUPS

There is ample evidence that members of stereotyped groups do not perform as well as nonstereotyped students in school and on standardized tests. It is not just that non-Asian ethnic minorities and women in math and science perform poorly; they tend to perform worse than nonstereotyped students even when they have same level of prior performance. This underperformance has been demonstrated so many times, especially for ethnic minority students, that it is seen as near lawful (e.g., Bowen & Bok, 1998; Ramist, Lewis, & McCamley-Jenkins, 1994; S. J. Spencer, Steele, & Quinn, 1999; Steele, 1997).

Why does underperformance occur? Some have argued that such effects have at least a partial basis in underlying biological or genetic differences (Benbow & Stanley, 1980; Jensen, 1980). Others have suggested they are due to structural factors such as poverty, bad schools or teachers, and a culture of failure (Entwisle, Alexander, & Olson, 2005; Hanushek & Rivkin, 2009). Although we do not deny the importance of structural factors, evidence suggests that a pivotal factor involves a psychological threat that stereotyped students experience in common school settings.

Consider two high school juniors—Jennifer, a Black student, and Angie, a White student. Both students have many reasons to want to perform well in class and on the SATs—to get into a good college, to meet a personal goal, perhaps to make family members proud. But Jennifer, who is African American, faces an additional performance pressure that Angie does not face. She knows that there is a widely held stereotype that African Americans have less intellectual ability than others. Should she perform poorly, she knows that this poor performance could be taken by others as confirmation of this threatening stereotype about her group.

Jennifer is experiencing stereotype threat and, according to hundreds of studies conducted over the past 15 years, it can cause distraction and anxiety and undermine her performance on challenging intellectual tests, including tests like those used for admission to a selective college (Schmader, Johns, & Forbes, 2008; Steele, 2010; Steele, Spencer, & Aronson,

2002). The first published study to investigate stereotype threat manipulated the psychological circumstances under which African American and European American college students took a difficult section of the GRE English Literature exam (Steele & Aronson, 1995). One group of participants was told that the test was evaluative of verbal ability—much as tests like the SAT and ACT are purported and perceived to evaluate students' intellectual abilities. Among these participants, African American students performed worse than European American students, even controlling for their prior SAT scores. A second group of students took the same test, but they were told that the test was a laboratory exercise that did not evaluate intellectual ability. In this *stereotype-safe* condition, African American students performed far better, equalling European American students' scores (again controlling for SAT scores). These results suggest that the African American students were capable of performing as well as the European American students but that contending with a negative stereotype about their groups' intellectual ability while taking the test depressed their score.

A follow-up study provided more direct evidence of the role of intellectual stereotypes. African American students who anticipated taking an evaluative intellectual test exhibited increased accessibility of thoughts related to negative racial stereotypes and self-doubts than African American students who anticipated completing a laboratory puzzle task. The results show that intellectual tests can automatically activate negative intellectual stereotypes in people targeted by such stereotypes.

A second, early set of studies showed that stereotype threat also affects women's performance in math (S. J. Spencer et al., 1999). College women performed worse than men on a test based on the GRE Advanced Mathematics exam when this test was described as having shown gender differences in previous research. But when the same test was described as having not shown gender differences, women performed as well as men. A follow-up study showed that women performed worse than men even when no information was given about gender differences but the test was presented as evaluative of math ability. Again, women's underperformance was eliminated when the test was described as having not shown gender differences. Subsequent studies have shown that this stereotype threat can also affect women's performance on in-class tests in real classrooms (e.g., Good, Aronson, & Harder, 2008).

Stereotype threat is a general psychological phenomenon that can arise for anyone who contends with a negative intellectual stereotype in a performance setting. It can undermine intellectual performance among Latino students (Gonzales, Blanton, & Williams, 2002), Native American students (Kaufman, 2006), Arab immigrants in France (Chateignier, Dutrévis, Nugier, & Chekroun, 2009), students of low socioeconomic status (Croizet & Claire, 2008; Désert, Préaux, & Jund, 2009; B. Spencer & Castano, 2007), and even White men, for instance, when a math test is said to assess why Asians are so good at math (Aronson

et al., 1999). There are multiple mechanisms through which stereotype threat causes underperformance (see Spencer, Logel, & Davies, in press), a primary mechanism being that worries about being judged in light of a negative stereotype consume executive resources needed to accurately solve test problems (see Schmader et al., 2008).

Stereotype threat has the strongest effect on the performance of people who identify with the stereotyped domain, and for whom membership in the stereotyped group is central to their self-concept (see Spencer et al., 2011). Some situations can produce weaker stereotype threat effects (see Wicherts & de Haan, 2011): Black students show little evidence of stereotype threat at historically Black colleges, presumably because the risk of being stereotyped is low, as are race-based belonging concerns (Walton & Cohen, 2007). Stereotype threat effects are weak or nonexistent on easy tests (O'Brien & Crandall, 2003), most likely because stereotyped students can correctly solve easy test problems in spite of worries about the stereotype consuming executive resources (Schmader et al., 2008).

Stereotype threat is one of a number of psychological factors that can affect student performance. Learned helplessness (Seligman, 1975) and low self-efficacy (Bandura, 1977) are related concepts; however, stereotype threat is qualitatively different. It is a group-based threat, so it can affect students' academic performance even if they personally do not believe the stereotype, and even if they believe that they are skilled in the domain—in fact, it has the strongest effect on students who are highly identified with the domain (Steele et al., 2002). Repeated experiences of stereotype threat, and its accompanying underperformance, however, may trigger feelings of low self-efficacy and learned helplessness, leading to further decrements in performance (cf. Steele et al., 2002).

IMPLICATIONS OF STEREOTYPE THREAT FOR UNDERSTANDING AND INTERPRETING MEASURES OF ACADEMIC PERFORMANCE

Research on stereotype threat raises an important question: If stereotype threat depresses the grades and test scores of non-Asian ethnic minority students and of women in math and science, do these grades and test scores systematically underestimate the ability of such students and their likely performance in academic environments absent stereotype threat? Suppose that Jennifer and Angie, the students described in the previous section, both earn a 3.75 grade point average (GPA) in high school and a 1200 on their SAT Math and Reading tests. As a result, they are both accepted to the same selective college. Their grades and test scores are the same, but do they reflect the same level of ability? If Jennifer, as a Black student, had to contend with stereotype threat in high school, her grades and test scores might underestimate her academic ability—she earned those scores despite having performed with the burden of stereotype threat. If the college she and

Angie attends has an even greater degree of stereotype threat than she faced in high school—this may be common, as stereotype threat increases as the difficulty of academic work increases and as students become more anonymous in larger school settings, exacerbating worries about belonging (Walton & Spencer, 2009)—Jennifer might receive worse grades in college than Angie, even though they started college with the same high school grades and test scores.

But what if Angie and Jennifer's college has taken steps to reduce stereotype threat? If Jennifer no longer faces the extra pressure of being judged in light of negative stereotypes about her group, her college achievement should not be as good as Angie's—but in fact should be better. Like a high jumper who had competed in an earlier meet wearing ankle weights, stereotype threat depressed Jennifer's past performance. Her 3.75 GPA and 1200 SAT score reflected more ability than did Angie's identical GPA and test score. If Jennifer earned those scores despite having contended with stereotype threat, her ability is in part *latent*—not fully reflected in her level of performance in high school (Walton & Spencer, 2009). If so, once stereotype threat is removed, Jennifer should do better than this past performance would suggest. She should get better grades in college than Angie.

This would be the case if stereotype threat systematically undermined students' real-world academic performances. An alternative possibility proposed by Sackett and colleagues is that Jennifer should do no better than her past performance would suggest (Sackett, Borneman, & Connelly, 2008; Sackett, Hardison, & Cullen, 2004). These researchers have argued that stereotype threat is restricted to laboratory settings, where it can cause students to perform worse than expected, but that stereotype threat does not significantly affect real-world performance.

What, then, is the evidence for latent ability—for the hypothesis that stereotype threat causes real-world measures of academic performance to underestimate the ability of women and ethnic minority students and their likely performance in settings in which stereotype threat has been removed or reduced?

Walton and Spencer (2009) tested this question in two meta-analyses. The first meta-analysis was of 39 laboratory experiments involving 3,180 participants in five countries. Each study manipulated stereotype threat and assessed the performance of stereotyped and nonstereotyped students on an intellectual test administered in the laboratory as a function of students' level of prior real-world performance in the same domain (e.g., test scores or grades). Combining data across studies, in the stereotype threat conditions, stereotyped students performed worse than nonstereotyped students who had the same level of past performance. In the purified conditions of the laboratory, it seems, especially high levels of threat can be created that further depress stereotyped students' performance.

But critically, in the stereotype-safe conditions in which stereotype threat was experimentally reduced, the

achievement gap reversed—in these conditions, stereotyped students performed significantly *better than nonstereotyped students* at the same level of past performance.

This finding is analogous to Jennifer scoring better than Angie on an academic test in a stereotype-safe college environment, despite having the same prior academic record. It strongly supports the latent ability hypothesis. Further, the effect was not limited to any one group of students. It arose among ethnic minority students and girls and women; among students of diverse ages and nationalities; and among students who scored low, average, and high on prior measures of performance, and it was found in both published studies and unpublished studies.

Before addressing the implications of these results, an important question involves whether the latent ability effect would arise on outcomes beyond laboratory-based tests. If stereotype threat effects on performance are not restricted to the lab, as some have argued (Sackett et al., 2008; Sackett et al., 2004), but in fact affect students' real-world academic performance, there should be evidence of latent ability on grades that are achieved in stereotype-safe environments. Walton and Spencer's second meta-analysis combined data from three well-known intervention field experiments. Each sought to reduce stereotype threat among African American students. They did so using diverse methods. Together, these studies included 15,796 students.

In the first intervention, two cohorts of Black and White seventh-grade students completed an in-class writing assignment near the start of the school year (Cohen, Garcia, Apfel, & Master, 2006). Students in the treatment condition were asked to identify their most important value from a brief list (e.g., relationships with friends and family, being good at art, etc.) and to write a paragraph about why that value was important to them. Writing about self-relevant values was expected to alleviate the stress that arises from being targeted by a negative group stereotype, as shown in past laboratory research (see Frantz, Cuddy, Burnett, Ray, & Hart, 2004; Martens, Johns, Greenberg, & Schimel, 2006; Schimel, Arndt, Banko, & Cook, 2004; Sherman & Cohen, 2006). If so, this intervention might allow Black students to perform better in the face of stereotype threat. In the control condition, students wrote about why a value that was not important to them might be important to someone else.

The second and third interventions targeted college students. In the second intervention, Black and White 1st-year college students were invited to join a residential dormitory program described as designed for students with exceptional academic potential (Steele, 1997; Steele et al., 1998). The honorific nature of the program signalled that negative intellectual stereotypes were not seen as relevant in the setting. The program also involved weekly small-group discussions about the transition to college, which could help communicate to Black students that concerns and difficulties about the transition to college are not unique to them or to their racial group but are normal—experienced by all students—and

temporary, and thus are due to the difficulty of the transition to college and not to their personal or group identity (see Walton & Cohen, 2007). The academic performance of students in this program was compared to the performance of matched and randomized students not in the program.

The third intervention was designed to buttress a sense of social belonging in school among Black college students who, being negatively stereotyped and underrepresented in college, might otherwise harbor pervasive doubts about their belonging and about whether others would fully include and value them. In this intervention, Black and White 1st-year college students read the results of a survey of senior students, which indicated that many students, both ethnic minorities and nonminorities, wondered if they fit in and belonged in their 1st year of college, but with time came to feel confident in their belonging (Walton & Cohen, 2007). Participants then wrote an essay and delivered a speech to a video camera describing how many students worry at first about whether they belong in college but in time come to feel that they do. This treatment was designed to lead students to see negative social events and worries about belonging as a normal and temporary part of the transition to college, and not as indicative of a lack of belonging. Control participants followed the same procedures, but the materials described a process of change that was irrelevant to issues of belonging.

Meta-analyzing across the three interventions yielded the same pattern as in the laboratory studies (Walton & Spencer, 2009). In the control conditions, minority students earned worse grades than White students even when they had the same level of prior academic performance (i.e., the same prior grades and standardized test scores). But in the treatment conditions, evidence of latent ability emerged. Here Black students significantly outperformed White students with the same level of past performance. The latent ability effect replicated with real-world school performance as the outcome. This result is as though Jennifer, despite having earned the same grades in high school as Angie, earned a higher GPA in college than Angie once the college took steps to reduce the threat of negative stereotypes.

Walton and Spencer's (2009) findings of latent ability on real-world measures of academic performance are based on only three interventions. However, subsequent interventions by our own research teams (e.g., Logel, Walton, Peach, Spencer, & Zanna, 2010; Walton & Cohen, 2011a) and by other teams (Woolf, McManus, Gill, & Dacre, 2009) suggest that the findings are robust.

In both meta-analyses, the observed effect size for the latent ability effect represents just less than one fifth of a standard deviation. This effect size suggests that stereotype threat explains a meaningful proportion of group differences in intellectual performance. Applied to the SAT, which was used as the premeasure of performance in nearly two thirds of the laboratory experiments in the first meta-analysis, this effect size estimate suggests that the SAT Math test underestimates the math ability of women like those who participated

in the included studies by 19 to 21 points and that the SAT Math and SAT Reading tests underestimate the intellectual ability of African American and Hispanic Americans like those who participated in these studies by 39 to 41 points. This effect suggests that, for many promising students, the psychological context of common testing environments significantly undermines academic performance and contributes to achievement gaps.

These results fully address the critique that stereotype threat may undermine academic performance only in the laboratory (e.g., Sackett et al., 2008; Sackett et al., 2004). The latent ability meta-analysis of the interventions demonstrates that power of stereotype threat in real-world environments by demonstrating that removing stereotype threat in the real world dramatically improves Black students' performance.¹

IMPLICATIONS FOR CREATING STEREOTYPE-SAFE ENVIRONMENTS IN HIGHER EDUCATION

The finding of latent ability in common measures of academic performance raises significant questions for colleges and universities. If conventional measures used to make admissions decisions in selective schools systematically underestimate the ability and potential of negatively stereotyped students relative to other students (Walton & Spencer, 2009), it would be inappropriate to interpret such measures at face value in evaluating candidates for admission. To do so would be to discriminate against stereotyped students—to evaluate more highly and potentially to admit more nonstereotyped students over stereotyped students, even when the latter are more qualified and more likely to perform well (Walton, Spencer, & Erman, 2011). But attaining the intellectual potential of stereotyped students is not automatic; rather, realizing this potential depends critically on the nature of the psychological environment present in a school. If the environment, even if inadvertently or unintentionally, contains insidious psychological threats rooted in negative intellectual stereotypes, even talented stereotyped students may not perform

as well as they are capable. Thus, decisions about how to interpret measures used in admissions and decisions about how to structure educational environments are fundamentally linked. For a diverse student body to perform as well as it is capable, it is necessary for the school to create a stereotype-safe environment—an environment that mitigates stereotype-related psychological threats (Walton et al., 2011). Although a comprehensive review of ways to reduce stereotype threat is beyond the scope of this article, next we describe three strategies that colleges and universities may use. Each strategy is rooted in psychological theory and has been rigorously tested in experimental research, often in field settings. Finally, we address critical questions that arise for admissions procedures in school settings in which threat has been reduced.

One way that schools can create a stereotype-safe environment is by providing students with psychological strategies to cope more effectively with negative stereotypes. Such strategies can produce an immediate improvement in students' academic performance. One consequence of stereotype threat is that individuals activate thoughts about the stereotype and concern about performing poorly (Davies, Spencer, Quinn, & Gerhardstein, 2002; Steele & Aronson, 1995). This activation contributes to worse academic performance because, once a test or academic task begins, people actively suppress these thoughts. This effortful suppression leaves fewer mental resources available to focus on the academic task at hand (Logel, Iserman, Davies, Quinn, & Spencer, 2009). Laboratory research shows that one strategy that can prevent people from engaging in such costly thought suppression is to ask them to replace worries about their performance and the stereotype with thoughts of an important personal identity. In one series of experiments, when women were instructed to use this strategy ("Should you feel worried you don't know what to do while taking this test, please replace those thoughts with thoughts about the personal identity you [just] generated"), women performed as well as men on a difficult and evaluative math test. By contrast, women who were not given this strategy performed worse than men (Logel, Iserman, et al., 2009).

This strategy, however, may not prove practical across diverse situations or over long periods. Replacing unwanted thoughts may be impossible over months or years in a threatening school environment. Eventually students may even come to associate the replacement thought with the threat it was meant to replace. A strategy that may be more effective in the long term is the self-affirmation approach used by Cohen and colleagues (2006) to reduce stereotype threat among African American middle school students. This intervention is designed to buttress students' sense of self-worth by guiding them to reflect on personally important values, such as their social relationships or skills outside of academics in in-class writing exercises. Such "self-affirmations" alleviate the stress arising in threatening situations, thus improving performance (Frantz et al., 2004; Martens et al., 2006; Schimel et al., 2004; Sherman & Cohen, 2006; Steele, 1988).

¹An additional critique of Sackett and colleagues (Sackett et al., 2008; Sackett et al., 2004) comes from those who study psychometric modeling. Sackett and colleagues have argued that the SAT and ACT cannot be biased against stereotyped groups, because underprediction effects like latent ability are not found on stereotyped students' grades in typical college environments. Wicherts and Millsap (2009) showed that this argument is flawed. They demonstrated that standardized tests can be biased against lower scoring groups without any underprediction. In addition, Wicherts, Dolan, and Hessen (2005) have employed measurement bias analysis to demonstrate that stereotype threat in experiments can lead to measurement invariance (i.e., to measurement bias). This research is consistent with our theoretical interpretation that the underperformance of stereotyped groups can result from stereotype threat. Additional research can add to the large literature on measurement bias by seeking to understand the full extent to which stereotype threat contributes to measurement bias.

A 2-year follow-up of students in Cohen et al.'s (2006) original two cohorts, along with students in a third cohort ($N = 385$; Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009) found that the boost in grades earned by African American students in the treatment condition persisted through the final 2 years of middle school. The affirmation appeared to cut off a negative cycle that would have led to poorer achievement over time among African American students. The affirmation's long-term effects were strongest for initially low-performing students, for whom low performance could otherwise have exacerbated the experience of threat and thus caused worse performance. As a consequence, in the affirmation condition, African American students' preintervention academic performance was less predictive of their later academic performance. In addition, it is worth noting that affirmation improved African American students' academic performance independent of expectancy effects—students did not know that the affirmation was meant to help reduce stress and thus improve academic performance. Such an intervention may also prove useful for students who have low self-confidence for reasons unrelated to stereotypes (e.g., test anxiety, low self-efficacy) as it could improve their ability to cope with a threatening testing environment without activating their lower performance expectations.

Although important questions arise about how to effectively deliver affirmation interventions in diverse school settings, the results suggest that affirmation can effectively ameliorate threat in common academic environments, when designed appropriately and delivered early enough in students' academic experience to forestall negative recursive processes that would otherwise undermine performance over time.

A second way that schools can create a stereotype-safe environment is to help members of stereotyped groups construe academic situations in ways that defuse perceptions of threat. This approach depends on the fact that social reality is often ambiguous and amenable to multiple construals. Indeed, the very first manipulations of stereotype threat involved construal. African American students told that a verbal ability test was *not* diagnostic of their intelligence presumably saw the experimental situation as one in which they were not at risk of being judged by the negative intellectual stereotype about their group (Steele & Aronson, 1995). As a result, worrisome negative racial stereotypes did not come to mind and students performed well.

Other research illustrates how strategies that target construal can affect real-world academic situations. One study investigated how White mentors can provide critical feedback to Black students without appearing biased or undermining recipients' motivation (Cohen, Steele, & Ross, 1999). In this research, White and Black students received critical feedback on an essay they had written ostensibly for a White professor. As compared to White students, Black students saw the professor as more biased, reported less identifica-

tion with writing, and felt less motivated to improve their work. This was the case even when the critical feedback was buffered with praise ("Good job but . . ."). Only when the feedback was prefaced by a message that the professor was giving the student critical feedback *because* he was holding them to a high standard and was confident that he or she could meet this standard did Black students construe the feedback as unbiased. In this condition, Black students were as identified with the task and as motivated to revise their work as White students. The message of high standards and assurance of the mentor's confidence in their ability to meet these standards disambiguated the meaning of the critical feedback: It conveyed exactly why the mentor delivered the critical feedback—not as a result of the application of bias but because his confidence in the recipient to meet a higher standard. It is not hard to imagine how such psychologically "wise" feedback strategies could, if implemented widely in an academic environment, significantly improve the academic performance of negatively stereotyped students (see Cohen & Steele, 2002).

Consistent with this research, interventions that target students' construal of negative events in academic settings can powerfully improve stereotyped students' academic performance and reduce group differences. In the Walton and Cohen (2007) intervention, 1st-year college students were led to see worries about belonging and negative social events as normal in the transition to college, not as specific to their personal or racial identity, and as challenges that can be overcome. A 3-year follow-up of two cohorts of students who took part in this study found that the 1-hr intervention produced impressive benefits throughout students' college careers (Walton & Cohen, 2011a). In the treatment condition, Black students' GPA from sophomore through senior year increased significantly relative to multiple control groups, reducing the racial achievement gap by 52%. In addition, 3 years after its delivery (i.e., at the end of students' college careers), the intervention reduced the cognitive accessibility of negative racial stereotypes among Black students and improved their results on a series of additional outcomes linked to a secure sense of belonging, including better self-reported health and greater subjective happiness. The results underscore the importance of ensuring that students who confront worries and difficulties in the transition to a new school know that these challenges are not unique to them or to their group but are common and temporary (see also Wilson, Damiani, & Shelton, 2002). Understanding negative social experiences as normative changes their perceived meaning for students—no longer do these events carry a broad, symbolic meaning for students about their belonging and fit in the school (see Walton & Cohen, 2011a).

A third way to create a stereotype-safe environment is reduce the perceived likelihood for stereotyped students that they will be seen stereotypically. Changing who else is present in an academic environment, for instance, can have this effect. Several studies show that women's math

performance improves when they are exposed to successful female role models in math (Marx & Roman, 2002; McIntyre et al., 2005; McIntyre, Paulson, & Lord, 2003) or when they are not underrepresented in a testing environment relative to men (Inzlicht & Ben Zeev, 2000; Sekaquaptewa & Thompson, 2003). Similarly, African American students no longer underperform relative to European American students when a test is administered by an African American experimenter (Marx & Goff, 2005).

These studies suggest that one way to create a stereotype-safe environment is to ensure that stereotyped students have opportunities for positive contact with other members of their in-group. This analysis underscores how fostering diversity in educational environments can promote academic success. However, because in many colleges and universities, students who face negative stereotypes are also numeric minorities—for example, one of only a few African Americans or Latino Americans in an English class or one of just a few women in an engineering or math class—for many stereotyped students most of their daily interactions will be with members of the nonstereotyped majority group. Is there a way that intergroup contact can be structured to reduce threat? Evidence suggests several possibilities. We have already reviewed research on how mentors can provide feedback more effectively to mentees across group divides—that is, in ways that disambiguate the meaning of critical feedback, reduce perceptions of bias, and sustain recipients' motivation (Cohen et al., 1999). But evidence also suggests that peer-to-peer interactions can also either evoke or mitigate concerns about stereotypes. One line of studies investigated dyadic interactions between male and female engineering students and the effect these interactions had on women's performance on an engineering exam (Logel, Walton, et al., 2009). This research found that men who were high and low in sexism interacted very differently with women—the less sexist men behaved more equitably and less dominantly. Further, this behavior affected women's performance. Both when observed in naturally occurring dyads and when the men's behavior was experimentally manipulated (i.e., with a confederate posing as a participant), women who interacted with less sexist men evidenced less stereotype threat and performed better on the evaluative and otherwise threatening engineering test (Logel, Walton, et al., 2009). This study suggests the intriguing possibility that reducing prejudice among nonstereotyped students might help create a stereotype-safe environment for minority students and women. Of interest, one way to reduce prejudice among nonstereotyped students is through self-affirmation, which can reduce not only feelings of threat among targets of stereotypes (Cohen et al., 2006) but prejudice among majority group members (Fein & Spencer, 1997). If affirmation were provided widely to students, it could thus foster a stereotype-safe setting in two ways—by mitigating threat among stereotyped students and by reducing prejudice among nonstereotyped students.

IMPLICATIONS FOR DECISION MAKING IN ADMISSIONS IN HIGHER EDUCATION

The finding that students who face negative stereotypes in school possess academic ability that is not reflected in their academic performance carries important implications both for the efforts of schools to create positive academic environments and for how they make admissions decisions. If a school reduced the level of threat in its internal environment and observed the latent ability effect—if it found that scores on measures it uses to make admissions decisions underestimate the ability and predicted performance of stereotyped students in its environment—the school would face a significant problem. Using such biased measures without addressing the bias inherent in them would institutionalize discrimination against the stereotyped group and reproduce inequality. How should the school reform its admissions process? This is an exceedingly complex and difficult problem, which raises important questions about psychology, policy, and law (Walton et al., 2011b). Here we offer three preliminary suggestions.

First, the bias needs to be acknowledged. Organizations including colleges and universities should recognize that traditional admissions criteria may not reflect the full potential of all applicants and, in particular, that they may systematically underestimate the ability and potential of applicants from groups that are pervasively negatively stereotyped in academic settings.

Second, with this recognition in mind, colleges and universities should work to create stereotype-safe environments that allow members of stereotyped groups to perform up to their potential. Creating such environments will require a serious, extended commitment and will no doubt involve false starts and dead ends. But doing so addresses a central mission of a school—to ensure that all students have the opportunity to perform as well as they are capable. Moreover, creating stereotype-safe environments may benefit nonstereotyped students as well as stereotyped students (e.g., Aronson, Fried, & Good, 2002). Further, creating such positive environments may not be costly in time or money. The interventions described in detail here (Cohen et al., 2006; Cohen et al., 2009; Walton & Cohen, 2007, 2011a) lasted 1 hr or less but improved students' academic performance months and years later. Although delivering these and other psychological interventions persuasively and powerfully poses unique challenges, such interventions are low-hanging fruit relative to other reforms in education.

Third, as it is acknowledged that traditional admissions criteria underestimate stereotyped students' ability and predicted performance, and as colleges and universities reduce psychological threats, these schools should admit more stereotyped students than would be suggested by the consideration of traditional admissions criteria alone (see Walton et al., 2011). They should do so out of fairness, as traditional measures underestimate stereotyped students' ability, and

because doing so would benefit these students, who will be able to reach their full potential in a stereotype-safe college environment. In addition, admitting more stereotyped students would improve the overall level of performance in the school itself, as the newly admitted students will have greater potential than students they displace. Ways to mitigate bias in admissions decisions may include using performance measures only as minimum criteria, using a broader range of measures of students' ability, and correcting for bias in interpreting performance measures. Choosing how exactly to mitigate the bias is a complex question both legally and practically, but we believe that, with careful and thorough study, progress can be made to develop procedures that would be less biased and that would produce selection decisions that more accurately index students' abilities and potential.

CONCLUSIONS

In the past 15 years, social-psychological research has provided convincing evidence that one cause of the academic underperformance of non-Asian ethnic minority students and of women in math and science involves the fact that such students confront pervasive negative intellectual stereotypes in common academic environments. In these environments, students risk being judged or viewed in light of the stereotype rather than by their individual merits, experience high levels of stress and threat, and may worry persistently about their social belonging in school. When concern about such stereotypes is reduced, negatively stereotyped students show evidence of latent ability—they outperform nonstereotyped students on tests and in the classroom when both groups of students have the same prior academic records. By applying research investigating strategies to create stereotype-safe academic environments, colleges and universities can create settings that allow all students to reach their potential and then revise admissions criteria and procedures to take into students' full ability and potential in making selection decisions.

REFERENCES

- Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effect of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38*, 113–125.
- Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35*(1), 29–46.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191–215.
- Benbow, C. P., & Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or artifact? *Science, 210*, 1262–1264.
- 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.
- Chateignier, C., Dutrévis, M., Nugier, A., & Chekroun, P. (2009). French–Arab students and verbal intellectual performance: Do they really suffer from a negative intellectual stereotype? *European Journal of Psychology of Education, 24*, 219–234.
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science, 313*, 1307–1310.
- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science, 324*, 400–403.
- Cohen, G. L., & Steele, C. M. (2002). A barrier of mistrust: How stereotypes affect cross-race mentoring. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 305–331). Oxford, England: Academic Press.
- Cohen, G. L., Steele, C. M., & Ross, L. D. (1999). The mentor's dilemma: Providing critical feedback across the racial divide. *Personality and Social Psychology Bulletin, 25*, 1302–1318.
- Croizet, J., & 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.
- Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin, 28*, 1615–1628.
- Désert, M., Préaux, M., & Jund, R. (2009). So young and already victims of stereotype threat: Socio-economic status and performance of 6 to 9 years old children on Raven's progressive matrices. *European Journal of Psychology of Education, 24*, 207–218.
- Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2005). First grade and educational attainment by age 22: A new story. *American Journal of Sociology, 110*, 1458–1502.
- Fein, S., & Spencer, S. J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. *Journal of Personality and Social Psychology, 73*(1), 31–44.
- Frantz, C. M., Cuddy, A. J. C., Burnett, M., Ray, H., & Hart, A. (2004). A threat in the computer: The Race Implicit Association Test as a stereotype threat experience. *Personality and Social Psychology Bulletin, 30*, 1611–1624.
- Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and Social Psychology Bulletin, 28*, 659–670.
- Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology, 29*(1), 17–28.
- Hanushek, E. A., & Rivkin, S. G. (2009). Harming the best: How schools affect the black–white achievement gap. *Journal of Policy Analysis and Management, 28*, 366–393.
- Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11*, 365–371.
- Jensen, A. R. (1980). *Bias in mental testing*. New York, NY: Free Press.
- Kaufman, J. C. (2006). Self-reported differences in creativity by ethnicity and gender. *Applied Cognitive Psychology, 20*, 1065–1082.
- Logel, C., Iserman, E. C., Davies, P. G., Quinn, D. M., & Spencer, S. J. (2009). The perils of double consciousness: The role of thought suppression in stereotype threat. *Journal of Experimental Social Psychology, 45*, 299–312.
- Logel, C., Walton, G. M., Peach, J., Spencer, S. J., & Zanna, M. P. (2010, June). *An intervention to improve women's grades in masculine engineering domains*. Paper presented at the 8th biennial convention of the Society for the Psychological Study of Social Issues, New Orleans, LA.

- Logel, C., Walton, G. M., Spencer, S. J., Iserman, E. C., von Hippel, W., & Bell, A. (2009). Interacting with sexist men triggers social identity threat among female engineers. *Journal of Personality and Social Psychology*, *96*, 1089–1103.
- Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology*, *42*, 236–243.
- Marx, D. M., & Goff, P. A. (2005). Clearing the air: The effect of experimenter race on target's test performance and subjective experience. *British Journal of Social Psychology*, *44*, 645–657.
- Marx, D. M., & Roman, J. S. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin*, *28*, 1183–1193.
- McIntyre, R. B., Lord, C. G., Gresky, D. M., Ten Eyck, L. L., Jay Frye, G. D., & Bond, C. F. (2005). A social impact trend in the effects of role models on alleviating women's mathematics stereotype threat. *Current Research in Social Psychology*, *10*, 9.
- McIntyre, R. B., Paulson, R. M., & Lord, C. G. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology*, *39*(1), 83–90.
- O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin*, *29*, 782–789.
- Ramist, L., Lewis, C., & McCamley-Jenkins, L. (1994). *Student group differences in predicting college grades: Sex, language, and ethnic groups* (College Board Report No. 93–1). New York, NY: College Entrance Examination Board.
- Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). High stakes testing in higher education and employment: Appraising the evidence for validity and fairness. *American Psychologist*, *63*, 215–227.
- 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*(1), 7–13.
- Schimel, J., Arndt, J., Banko, K. M., & Cook, A. (2004). Not all self-affirmations were created equal: The cognitive and social benefit of affirming the intrinsic (vs. extrinsic) self. *Social Cognition*, *22*(1), 75–99.
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*, 336–356.
- Sekaquaptewa, D., & Thompson, M. (2003). Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology*, *39*(1), 68–74.
- Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. San Francisco, CA: W.H. Freeman.
- Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 38, pp. 183–242). San Diego, CA: Elsevier Academic Press.
- Spencer, B., & Castano, E. (2007). Social class is dead. long live social class! Stereotype threat among low socioeconomic status individuals. *Social Justice Research*, *20*, 418–432.
- Spencer, S. J., Logel, C. E. R., & Davies, P. (in press). Stereotype threat. *Annual Review of Psychology*.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*(1), 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). San Diego, CA: Academic Press.
- Steele, C. M. (1992, April). Race and the schooling of Black Americans. *Atlantic Monthly*, pp. 68–78.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*, 613–629.
- Steele, C. M. (2010). *Whistling Vivaldi and other clues to how stereotypes affect us*. New York, NY: Norton.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797–811.
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–440). San Diego, CA: Academic Press.
- Steele, C. M., Spencer, S., Nisbett, R., Hummel, M., Harber, K., Schoem, D., & Carter, K. (1998). *African American college achievement: A "wise" intervention*. Manuscript in preparation, Stanford University, Stanford, CA.
- Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, *92*(1), 82–96.
- Walton, G. M., & Cohen, G. L. (2011a). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science*, *331*, 1447–1451.
- Walton, G. M., & Cohen, G. L. (2011b, April). *A question of belonging*. Paper presented at Haas School of Business, University of California at Berkeley, Berkeley, CA.
- Walton, G. M., & Spencer, S. J. (2009). Latent ability: Grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science*, *20*, 1132–1139.
- Walton, G. M., Spencer, S. J., & Erman, S. (2011). *Affirmative meritocracy*. Manuscript under review.
- Wicherts, J. M., & de Haan, C. (2011). *A meta-analysis of the effects of stereotype threat on the cognitive test performance of African Americans*. Manuscript under review.
- Wicherts, J. M., Dolan, C. V., & Hessen, D. J. (2005). Stereotype threat and group differences in test performance: A question of measurement invariance. *Journal of Personality and Social Psychology*, *89*, 696–716.
- Wicherts, J. M., & Millsap, R. E. (2009). The absence of underprediction does not imply the absence of measurement bias. *American Psychologist*, *64*, 281–283.
- Wilson, T. D., Damiani, M., & Shelton, N. (2002). Improving the academic performance of college students with brief attributional interventions. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 89–108). Oxford, England: Academic Press.
- Woolf, K., McManus, C., Gill, D., & Dacre, J. (2009). The effect of a brief social intervention on the examination results of UK medical students: A cluster randomised control trial. *BMC Medical Education*, *9*, 35.