<CN>7</CN>
<CT>The Myth of Intelligence</CT>
<CST>Smartness Isn't Like Height</CST>
<CA>Gregory M. Walton</CA>

<A>Introduction</A>
Formal education is the primary mechanism by which modern societies prepare citizens to take part in civic life and the workforce. Given the enormous social, political, and economic advantages that arise from education, the way educational opportunities are distributed raises important questions of justice and equal opportunity (e.g., Brighouse and Swift, this volume; Reich, this volume). Yet the way we distribute educational opportunities depends on assumptions about the nature of intelligence and its development. Who is educable and how and when? For instance, if intelligence is assumed to be fixed early in life through genetics, the state may have a responsibility only to ensure that individuals reach their inborn potential or some minimum level of that potential and, perhaps additionally, to identify individuals' level of intelligence and track them into appropriate educational or career paths. If intelligence is thought to be largely the product of early childhood experiences, the state's responsibility may be to ensure an equal or an adequate educational experience early in life (see Reich, this volume). Drawing on modern social science, I argue in contrast to both assumptions that intelligence is in part *socially situated*—that it is an interaction between persons and situations, between potentials afforded by the individual and how those potentials are elicited or denied in social situations. If this is the case, the demands of justice are larger in scope: they include ensuring that educational situations are structured in ways that allow all people to learn and to perform up to their potential. Moreover, this approach suggests that efforts to assess enduring differences in individuals' intelligence and differential treatment based on these assessments are fundamentally misguided and inappropriate.

When Lewis Terman brought the IQ test to America, he helped create the modern concept of intelligence. An early version of the test had been created by the French psychologist Alfred Binet as a means of identifying the intellectual strengths and weaknesses of poor French schoolchildren. Binet's goal was a hopeful one that emphasized children's malleability and educability. In developing an IQ test, Binet hoped to help schools identify and build on individual students' strengths and identify and address their weaknesses. The approach of Terman, a Stanford psychologist, was radically different. Terman revamped the test, called it the Stanford-Binet Intelligence Scale, and aimed to use it to identify smart people. Terman delivered his test to thousands of California schoolchildren in the 1920s, identified those with the very highest scores, and admitted them to a still-running longitudinal study to track what he anticipated would be their impressive successes in life.

Terman's research institutionalized "IQ" in American culture. Embedded in his approach was a specific conception of intelligence as a stable property of the self, relatively unchanged over the life span, experience, and situations, in which there are broad individual differences that originate largely in inheritance and genetics. Further, Terman suggested, intelligence could be accurately assessed even in childhood using brief standardized tests that would predict important outcomes like educational and career

success years later. In short, Terman believed that intelligence is something like height, only more important. It is a fact of who we are: some people are tall, some people are short, and, aside from wearing platform shoes, there is not much we can do about it.

There are those who still endorse a Termanlike conception of intelligence with large, stable individual differences rooted primarily in genetics (e.g., Jensen 1973, 1980; Herrnstein and Murray 1994). Richard Nisbett's (2009) impressive review of social and environmental factors that shape intelligence shows how a strong hereditary view of intelligence persists as well in recent books like Judith Rich Harris's *The Nurture Assumption* (1998), Steven Pinker's *The Blank Slate* (2002), and Steven Levitt and Stephen Dunber's *Freakonomics* (2006). An approach to intelligence that aims primarily at identifying individuals' level of smartness is also evident in the continued use of standardized testing in educational admissions, with tests like the SAT and LSAT shaping admissions decisions in higher education and new tests being used to diagnose the ability of four- and five-year-olds and determine admission to selective kindergarten programs. Some even suggest that society should give up on providing higher educational opportunities to all and restrict these opportunities to people with high intelligence, who are presumably best equipped to take advantage of them (e.g., Herrnstein and Murray 1994).

A view of intelligence as largely stable across much of the life span can also be found in very different quarters, even in research that emphasizes the malleability of intelligence early in life. Scholars of early childhood emphasize the profound importance of early educational experiences in setting children on the right academic track. Such experiences can help children establish the foundation of academic skills and abilities that undergirds later academic and professional success. As a consequence, developmental scholars emphasize the importance and cost effectiveness of early childhood interventions (e.g., high-quality preschool programs) compared with later educational interventions (e.g., job training, adult education) (see Heckman 2006; Diamond et al. 2007). In this approach, just as height is disproportionately shaped by early nutritional experiences, so intelligence is disproportionately shaped by early educational experiences.

There is no doubt that early childhood education lays the critical foundation for academic growth and development. But in some forms this view too can be taken to imply that later in life intelligence is largely fixed, albeit by early experience, not genetics, and relatively immune to local contextual factors. As I will argue in this chapter, modern research from social psychology suggests a different approach. In this approach, intelligence is not like height at all. The problem is not just that our rulers for measuring intelligence are bad—that they are inaccurate—although they often are. The problem is more fundamental: intelligence is not located simply inside an individual. If I am five feet eleven, I am five feet eleven at home and in the office, by myself and with friends. Whether I am "tall" or "short" may depend on whether I am at a jockey convention or in an NBA locker room. But I am always five feet eleven. It may seem common sense that intelligence too is a part of me that does not change in different circumstances; this intuition testifies to the enduring power of Terman's concept of intelligence as a stable property of the self. But modern theorizing of human psychology in general, and research on the impact of situational factors on intellectual performance in particular, suggests a different view.

If intelligence is not like height, what is it like? Consider people's assessments of their own traits and attitudes. Suppose I think I am an assertive person—I speak up and say my mind. Suppose I am a chocoholic—there is nothing I love more than a piece of dark chocolate. These self-perceptions, especially personally important, well-elaborated ones, are things we carry with us from place to place, situation to situation. They are stable aspects of self. And, when asked, we can accurately report on them at will. Right?

In fact, a broad range of research in modern psychology suggests a different view. This research finds that, far more than we commonly believe, our assessment of our own traits and attitudes are constructions of the moment whose expression is the product of situational factors, even subtle ones.

In a classic demonstration, Norbert Schwarz and his colleagues (Schwartz et al. 1991) asked some people to list six examples of situations in which they had behaved assertively and asked other people to list twelve examples. They were interested in the information people glean from the ease or difficulty they experience calling such examples to mind. They predicted that if people pay attention to how difficult it is to list twelve examples of assertiveness, they might infer that perhaps they are not so assertive. This is just what they found: when asked to list twelve examples of their assertiveness, people rated themselves *less* assertive.

Another line of research I conducted with Mahzarin Banaji illustrates how subtle cues shape attitudes. We asked people to list their favorite drink, food, and so forth (Walton and Banaji 2004). Later, under cover of a study on "handwriting," we assigned some people to describe the preferences they had listed using a noun phrase—to write, for instance, "I am a chocolate-eater," a grammatical construction that conveys a strong and enduring preference. Other people were asked to use a weaker verb phrase, for instance, to write "I eat chocolate a lot" (cf. Gelman and Heyman 1999). People were then asked how much they liked chocolate. Would people attend to the linguistic cues embedded in their self-descriptions and thus report stronger attitudes in the noun phrase condition than in the verb phrase condition? They did. People rated their love for chocolate, for example, as stronger and more stable when they had been assigned to write "I am a chocolate-eater" rather than "I eat chocolate a lot."

Perhaps intelligence works something like this: it is a property we commonly see as characteristic or even as defining of the self, as stable in different situations and over time, but one whose quality and expression in fact vary significantly in different social settings. I call this view of intelligence *socially situated*. In this view, intelligence is the product of an interaction between potentials afforded by aspects of the individual and how those potentials are elicited or denied in social situations. This term borrows from scholars who emphasize that intelligence is "situated"—that it depends on context and culture and arises to help people solve the particular problems they face in daily life (e.g., Rogoff and Lave 1984). Complementing this work, the term socially situated emphasizes how common social situations give rise to even subtle psychological processes that affect intelligence and intellectual performance.

Asking how we understand "intelligence" is no mere ivory tower exercise. On the contrary, different understandings of intelligence have profound consequences for educational and social policy and for students and society. A view of intelligence as stable and fixed leads to an emphasis on assessing students' abilities rather than on improving abilities—on testing over teaching and performing over learning. This

approach can yield pessimism about the prospects of improving people's academic performance (Nisbett 2009). For instance, a view of intelligence as originating primarily from early childhood experiences can lead to an emphasis on high-quality early school experiences but the neglect of later educational opportunities. If early childhood education is all-important, why improve later educational situations? Further, an understanding of intelligence as fixed can undermine students' effort and motivation in the face of academic setbacks and challenges (Dweck 1999, 2006; Dweck and Leggett 1988). If ability is fixed, setbacks may seem evidence of its lack. If ability is lacking, why try?

In addition, a view of intelligence as stable and fixed implies that grades and standardized test scores can measure a person's inherent intelligence and thus imply that admissions and hiring decisions based on such indicators of merit are fair and just (Bobocel et al. 1998). But what if pervasive situational factors systematically bias these scores for specific groups—for instance, depressing the scores of students from disadvantaged groups and underestimating their ability and potential relative to other students' (Walton and Spencer 2009)? If so, standard selection systems would disadvantage already disadvantaged students and advantage others. Under the auspices of fairness they would launder advantage, turning a psychological edge into an increased chance of admission to a selective educational program and do it in a way that appears fair and just (see Walton, Spencer, and Erman, in press; see also Guinier 2011; Guinier and Sturm 2001). This system reproduces inequality. With so much at stake, what is the evidence that intellectual performance is socially situated?

Before proceeding, I note that the present analysis is consistent with scholarship exploring the limitations of traditional indicators of intelligence. For instance, such measures may be only weakly predictive of important criterion variables (e.g., school and job performance). They can also be overly narrow and so fail to capture the full range of relevant skills and abilities. Illustrating this approach, Sternberg and colleagues (2006) find that assessing college applicants' practical and creative skills in addition to their SAT scores and high school grades can shrink group differences and raise correlations with college grades (see also Gardner 1999; Oswald et al. 2004; Sedlacek 2004). In examining how performance on even traditional indexes of intelligence is inextricably shaped by social situations, this chapter complements this past work, provides new insights into the nature of intelligence, and draws novel implications for ways to improve educational outcomes and expand equal opportunity.

<A>A Social-Psychological Approach to Understanding Persons and Situations</A>
Since its founding, social psychology has powerfully emphasized how the immediate social and psychological situation affects people's behavior and experience (Ross and Nisbett 1991). Classic studies demonstrate that even brief interactions with strangers can reshape people's very perception of physical reality (e.g., through conformity processes: Asch 1952; Sherif 1935; see also Berns et al. 2005) and cause people to behave in seemingly extraordinary ways, for instance, to harm others as a consequence of obedience to authority (Milgram 1974) or powerful social roles (e.g., as a guard in a mock prison; Haney, Banks, and Zimbardo 1973). Further illustrating the power of social situations is the research described above showing how the subjective ease or difficulty

of listing instances of assertiveness affects people's assessment of their own assertiveness.

The social-psychological approach to human nature provides a broad critique of a fixed view of intelligence on two levels. First, in emphasizing the power of situations, it implies that in explaining behavior we give undue weight to persons compared with situations. Lee Ross termed the tendency to explain behavior in terms of the internal attributes and characteristics of others "the fundamental attribution error" (Ross 1977). In a classic study, Ross, Amabile, and Steinmetz (1977) randomly assigned participants to be either the questioner or the contestant in a mock "quiz show." The situational dynamics in this quiz show seem obvious: the questioner can select topics he or she knows something about. Yet when third-party observers are asked at the end of the study to rate the questioner and contestant on their "general knowledge," questioners were consistently rated more highly. Even contestants showed the bias. People did not appropriately weigh even a seemingly obvious feature of the situation—that the questioner got to choose the questions—and as a result they saw questioners as more knowledgeable than contestants.

When we see a student perform well or poorly in school, are we subject to the same fundamental attribution error? Especially when the situational forces at work are far more subtle, do we too readily attribute students' performance to their internal characteristics—their inherent smartness or dumbness—and neglect situational factors that might affect their performance?

Second, the social-psychological approach to human nature challenges the conception of the self implied by a fixed and stable view of intelligence. If pressed, many laypeople view different aspects of people's selves—like their attitudes, beliefs, memories, and personality, as well as their skills and abilities—as basically stable properties that they can access or call forth at will. In contrast to this view is research investigating the effects of context on cognition, self-perceptions, and attitudes (Schwarz et al. 1991; Walton and Banaji 2004; see also Schwarz 2000; Tversky and Kahneman 1981). This research suggests that diverse aspects of self should be thought of as constructions whose content is shaped in part by the context and manner in which they are assessed. In many ways, for instance, memory functions as a reconstruction whose form and content are shaped by the context in which the memory is recalled. This quality gives rise to the possibility of "false memories"—memories that people experience as accurate representations of the past but that are substantially influenced by or even created by subtle suggestions (Loftus 2003).

Consider even personality. Although historically research has examined broad dispositional factors such as individual differences in extraversion or conscientiousness as predictors of behavior, much modern personality research instead emphasizes interactions between people's potentials and how situations elicit those potentials (Mischel and Morf 2003). As an example, a person who has low self-esteem may behave much like a person with high self-esteem in many contexts. But in a relationship, people with low self-esteem may be more apt to doubt their partners' love. Implicitly or explicitly, they think, "I'm not such a great person. How could my partner love me?" As a consequence, they may perceive ambiguous relational events—like a romantic partner's having a friendly conversation with an attractive store clerk of the opposite sex—as evidence of a lack of love and respond by lashing out or withdrawing from the

relationship (Murray et al. 2002). Here the behavioral disposition is not automatic or inevitable. It is a potential elicited in specific kinds of situations but not others.

Does intelligence work like this? Is it inextricably bound to social contexts, affected by the manner of assessment, and unstable? Is intelligence socially situated?

<A>Socially Situated Intelligence</A>

In some ways, the notion that intellectual performance is affected by situational factors is obvious and uncontroversial. If a construction crew is taking down a wall outside a classroom, of course students will perform worse. If students enroll in an effective test-prep class and thus take a standardized test with a better understanding of its form and content, they may do better. But these examples seem to be idiosyncratic—sources of random variance in students' performance rather than pervasive factors having broad effects that could systematically disadvantage some groups relative to others and cause us to think differently about the nature of intelligence. Do subtle social situations have large and pervasive effects on the intellectual performance of broad swaths of students? Here I review two examples of such powerful situational factors.

<B>Implicit Theories of Intelligence</B>

One important determinant of intellectual performance is whether students persist and work hard in the face of academic challenges or give up. And whether students persist depends importantly on their implicit theories of intelligence. Some students view intelligence as a fixed quantity that a person either possesses or does not (an "entity" theory). Others view it as a malleable quantity that increases with effort and learning (an "incremental" theory; Dweck 1999, 2006; Dweck and Leggett 1988). Notably, even traditionally measured forms of intelligence are strikingly malleable, with specific kinds of cognitive training programs (e.g., to expand working memory capacity) yielding large improvements on tests of fluid intelligence, a common measures of IQ (Jaeggi et al. 2008).

Here, however, the emphasis is on students' *theories* of intelligence—as fixed or malleable. Students who hold one theory versus the other may be similarly motivated when they feel confident in their abilities or after an academic success. But after an academic failure, their profiles diverge sharply. As shown in diverse laboratory experiments and longitudinal studies, students with a fixed (entity) theory of intelligence tend to perceive academic failure as evidence of their inability—proof that they do not have what it takes to succeed. In response, they tend to withdraw and thus show declines in academic performance over time. By contrast, students with a malleable (incremental) theory of intelligence tend to perceive academic failure as evidence that they have not yet mastered the material. In response, they may increase effort and show improved academic performance over time. For instance, one longitudinal study found that students' implicit theories of intelligence predicted their academic performance over two years in the transition to junior high school (Blackwell, Trzesniewski, and Dweck 2007, study 1). Although students with the fixed and malleable theories entered junior high school with identical past achievement test scores, as the work became harder their math grades diverged. This divergence occurred because students with the two theories reacted differently to challenges and setbacks, with the malleable theory predicting more resilient, effort-based responses.

Where do students learn an entity theory of intelligence? Perhaps indirectly from Terman and others, who advance a fixed concept of intelligence. But an entity theory of intelligence can also be passed on directly to children, for instance through the kind of praise they receive. In one set of studies, Mueller and Dweck (1998) praised fifth-graders for completing a moderately difficult set of logic problems. Some children received entity praise ("That's a really high score, you must be very smart at these problems"). Others received effort praise ("That's a really high score, you must have worked hard at these problems"). A third group received neutral praise ("That's a really high score"). The children were then given a very difficult set of problems, on which they all performed poorly. Finally, children were given a critical second set of moderately difficult problems. The key question involved their resilience—how well would they bounce back after failing very difficult problems?

Children who received neutral praise performed equally well on the first and second sets of moderately difficult problems. But consistent with the effects of implicit theories, children who had received entity praise solved 30 percent fewer problems in the second set. They were also less willing to do more challenging problems in the future— they preferred easy problems on which they could score well. By contrast, students who had received effort praise performed well on the final set—even better than on the first set—and requested more challenging problems in the future—problems they could learn from. Just a single sentence of praise changed the situation and the meaning of failure for students, and shaped their subsequent performance.

If a malleable (or incremental) theory of intelligence predicts more adaptive student behaviors, does teaching students this theory improve academic outcomes? The answer is yes. Before describing the implicit theory of intelligence interventions, I emphasize that this intervention and other social-psychological interventions described below are not silver bullets, even though they can have large and impressive effects (Yeager and Walton 2011). Rather, in testing psychological interventions, researchers necessarily hold constant other important factors, such as the nature of the schooling environment, teachers, and pedagogy, while randomly assigning students to treatment and control conditions. This allows a rigorous test of the causal effect of the tested intervention, but it does so within a particular educational context, which, while not part of the intervention, may be critical to its success. Such studies can demonstrate the potential impact of an intervention in common academic environments, but they do not mean that the intervention will always be effective or that it can compensate for other important factors (such as bad teaching).

This said, in common field settings interventions to a malleable (or incremental) theory of intelligence can significantly raise students' academic achievement. In one experiment, Blackwell and colleagues randomly assigned seventh-grade students to one of two workshops (Blackwell, Trzesniewski, and Dweck 2007, study 2). In the incremental theory workshop, students learned how intelligence can "grow" through effort and learning. They also learned study skills with which to put their malleable theory into practice. In the control workshop, students learned only the study skills. Both groups had shown declining math grades before the workshop, and the control group continued to decline. Although they had learned effective study skills, they were apparently not motivated to put them to work. In contrast, students in the malleable

theory workshop condition showed a sharp rebound in their math grades so that they earned significantly higher final grades than students in the control group.

Students who are underrepresented and negatively stereotyped in school may be at heightened risk from the view that intelligence is fixed. The stereotypes that "women are bad at math" and that "black and Latino students are not smart" allege fixed inability. If so, the view that intelligence is malleable may help sustain motivation and achievement among negatively stereotyped students. Consistent with this theorizing, incremental theories interventions have been shown to improve black college students' grades and school attitudes (Aronson, Fried, and Good 2002) and middle school girls' scores on state standardized math tests, eliminating gender differences (Good, Aronson, and Inzlicht 2003).

If seemingly small cues like a sentence of praise or seemingly small situations like participating in a school workshop can substantially affect students' academic engagement and performance over time, it calls into question what we mean by intelligence. Is it just within a student's head? Or is intelligence a quality that emerges as interactions between persons and situations? Are there other examples of subtle situational factors that profoundly shape students' academic performance?

<B>Stereotype Threat</B>
A second example involves stereotype threat (Steele 2010; Steele, Spencer, and Aronson 2002). When African American students perform on intellectual tasks, or when women perform in quantitative fields, they are often aware of negative stereotypes about the ability of their group. This awareness can cause *stereotype threat*—the worry that, should they perform poorly, others could view their poor performance as evidence for the truth of the stereotype. This worry is distracting and causes anxiety; it can prevent students from becoming fully absorbed in the task at hand and thus undermine their performance.

In a classic laboratory study, African American and European American college students took a test composed of difficult verbal problems from the Graduate Record Examinations (Steele and Aronson 1995). When students were told the test was intended to evaluate their verbal reasoning ability—a description that made the negative stereotype about African Americans' intellectual ability relevant—African Americans performed far worse than European Americans. But when students were told the test was a verbal puzzle task, African American students' performance soared and, in fact, equaled European Americans' (controlling for Scholastic Aptitude Test scores) (Steele and Aronson 1995). Similarly, a study of math performance found that women performed worse than men when students were told that the test investigated gender differences (Spencer, Steele, and Quinn 1999). But when they were told that the test was fair across gender groups, women and men performed equally well.

Notably, negative intellectual stereotypes also affect the performance of students who are not targeted by negative stereotypes, such as whites and men, but here the effect is a positive one. In a meta-analysis (or statistical summary) of forty-three experiments, Geoffrey Cohen and I found that nonstereotyped students experience a performance boost called *stereotype lift* when they perform in situations in which a *different* group is subject to the negative scrutiny and suspicion of a stereotype (Walton and Cohen 2003).

Stereotype threat can also undermine learning—the acquisition of new academic knowledge (Taylor and Walton 2011; see also Rydell et al. 2010). In one study, black and

white students were given rare words to learn. For some students the task was described in a way designed to elicit stereotype threat—it was said to evaluate students' ability to learn, thus making negative stereotypes about African Americans relevant. For other students the task was described as investigating "learning styles," which does not evoke the stereotype. A week later, students returned to the laboratory and completed a "warm-up" task in which they were asked to define some of the words they had studied a week earlier. All participants completed the "warm-up" in a nonthreatening, nonevaluative setting, so that performance would reflect how well they had learned the words a week earlier, not psychological threats that could undermine how well they performed. White students performed well regardless of the condition in which they had studied the words. But black students performed far worse when they had studied in the threatening environment than when they had studied in the nonthreatening environment (Taylor and Walton 2011). Those who had studied in the threatening learning environment never fully acquired the academic material. Notably, when students were subsequently asked to define the rest of the words in a threatening situation—a "test" said to evaluate students' verbal ability—even black students who had studied in the nonthreatening environment performed poorly. The results show that black students experience a form of double jeopardy in academic environments as a consequence of stereotype threat—threat can both prevent them from fully acquiring academic knowledge and prevent them from performing well on material they have learned well.

Given the effects of stereotype threat on academic performance in the laboratory, can interventions to reduce stereotype threat improve academic outcomes in real-world school settings? They can. Here I review two such interventions.

First, one antidote to stereotype threat involves "self-affirmation." The self-affirmation intervention is premised on the idea that stereotype threat can cause students to experience high levels of stress and anxiety in school, which undermines academic performance (Cohen et al. 2006). To alleviate this stress and anxiety, in the affirmation exercise students are given a list of important values (e.g., religion, relationships with friends and family, athletics) and asked to select their most important value and write about why this value is important to them. The theory underlying this intervention is that when a situation threatens an important identity, reminding people of other valued identities can bolster their feeling of self-integrity—their perception of themselves as good, efficacious, and moral—and thereby help people perform at a high level despite threat (see Sherman and Cohen 2006).

In one randomized field experiment, white and black seventh-graders in a middle-income school completed an in-class self-affirmation near the beginning of the school year, soon before a stressful experience (taking a test) (Cohen et al. 2006). The exercise was presented as an in-class writing assignment, although, to keep teachers blind to students' condition assignment, teachers passed out materials in sealed envelopes and returned these to the researchers, never having seen students' essays. In the self-affirmation condition, students wrote about a value that was personally important to them. Control students wrote about why a value that was not important to them might matter to someone else. In total, the affirmation exercise took fifteen to twenty minutes; it was administered once, twice, or several times depending on the cohort.

The affirmation had no effect on white students, but it reduced the cognitive accessibility of racial stereotypes among black students and raised their end-of-term

course grades by one-third of a grade point, narrowing the racial achievement gap by 40 percent. The number of black students who received a D or an F grade in the course was reduced by half. Long-term follow-ups with three cohorts of students showed that the boost in grades for black students extended to their grade-point average (GPA) across all core academic courses and persisted over the last two years of middle school. It appeared to do so by interrupting a negative recursive cycle whereby poor performance begat worse performance over time (Cohen et al. 2009). The affirmation's effects were thus most evident among initially low-performing black students.

A second intervention to address stereotype threat involves bolstering students' sense of social belonging in school. One consequence of stereotype threat is to cause students to question whether others in school will fully include, value, and respect them (Walton and Cohen 2007). The need to belong—to form positive social relationships with others—is fundamental to humans (Baumeister and Leary 1995) and forms a critical basis of motivation and achievement (Walton, Cohen, Cwir, and Spencer 2012.) If students feel uncertain of their social belonging in school, they may monitor the school environment for cues that indicate whether they belong. In this state of uncertainty, negative social events—like critical feedback from a teacher or loneliness in class—may seem to be diagnostic of a global lack of belonging in the school. In turn, such global attributions may erode students' engagement in school and academic performance (see Mendoza-Denton et al. 2002).

To prevent such deleterious attributions, the "social-belonging" intervention provides students a nonthreatening explanation for negative social events in school (Walton and Cohen 2007, 2011). The intervention was first tested in a sample of black and white first-year college students. Students read a survey of upper-year students at their school. The survey indicated that negative social events and feelings of not belonging are normal at first in college (e.g., experienced by students of all ethnicities) and dissipate with time. The treatment was designed to lead students to attribute such events and feelings to the difficulty of the transition to college rather than to a lack of belonging on their part or on the part of their racial group (cf. Wilson, Damiani, and Shelton 2002). To encourage students to internalize the treatment message, they were asked to write a speech describing how their own experiences had changed over time in college and to deliver the speech for a video that could be shown to incoming students the following year, ostensibly to ease their transition to college. This procedure takes advantage of the "saying-is-believing" effect—leading people to advocate for a persuasive message to a receptive audience is a potent means of persuasion (Aronson, Fried, and Good 2002). In total, the intervention lasted about an hour. By securing students' sense of social belonging in school, the intervention was designed to change the trajectory of their experience in the school over time—to help them build the positive experiences and social relationships that form the foundation for lasting academic success.

The results suggest that the intervention had this effect. For white students, who have little cause to doubt their belonging in school on account of their race, the treatment had little impact. But for black students it had large and long-lasting benefits. In two cohorts of students, the intervention delivered in the spring of students' freshman year raised black students' GPA from sophomore through senior year compared with multiple randomized control groups. Consistent with the idea that the intervention changed the

trajectory of students' experience, the intervention led to a term-by-term rise in African American students' GPA over this three-year period. Overall, from sophomore year through senior year, the intervention reduced the achievement gap between African American and white American students by 52 percent (Walton and Cohen 2011). Moreover, the effects were statistically mediated by measures of the underlying psychological process targeted by the intervention: daily diaries completed in the week after the intervention showed, as predicted, that the intervention made black students' sense of belonging less vulnerable to daily adversity, which no longer seemed to carry a global or symbolic meaning for them. This effect mediated the long-term gain in black students' GPA.

Research on stereotype threat carries an important implication. Insofar as school is a pervasively evaluative environment in which negative stereotypes routinely come to mind, the academic performance of negatively stereotyped students may be systematically depressed by stereotype threat. If this is the case, they would be like runners facing a stiff headwind, whose times underestimate their ability and potential to perform well in a neutral environment. To test this idea, Steve Spencer and I conducted two meta-analyses, one summarizing data from thirty-seven laboratory experiments and the other summarizing data from three field experiments. Together these studies included 18,976 students in five countries (Walton and Spencer 2009). Both meta-analyses compared the performance of negatively stereotyped students (e.g., women, African Americans) in academic environments in which threat had been reduced with the performance of nonstereotyped students (e.g., men, whites). In addition, we compared students who had the same prior level of academic performance—the same prior grades or test scores. If stereotype threat systematically depresses the performances of negatively stereotyped students, these performances should underestimate their performance in settings in which stereotype threat has been removed. If so, in these circumstances, stereotyped students should outperform nonstereotyped students who had the same level of prior academic performance.

This was just what we found. In both meta-analyses, when threat was reduced stereotyped students performed better than nonstereotyped students with the same prior scores. This "latent ability effect" was found for students with all levels of academic performance, and both for women and girls and for ethnic minority students. To illustrate the magnitude of the effect using the SAT as the metric, the size of the observed effect suggests that the SAT Math test underpredicts the math ability of women like those in the surveyed studies by nineteen to twenty-one points and that the SAT Math and Reading tests underpredict the intellectual ability of African and Hispanic Americans by thirty-nine to forty-one points for each group. These differences are substantial and could make the difference between rejection and admission by a selective college or university. Further, they suggest that a broad range of promising stereotyped students have significant academic potential that is hidden by their scores on common performance metrics like grades and test scores.

<A>Implications</A>
This chapter has questioned how useful and valid it is to maintain a view of intelligence as something stable lodged within individuals. The research reviewed here shows that intellectual performance can vary widely depending on how people think about

intelligence, on the kind of praise they receive, and on whether subtle cues in school settings trigger concern about negative intellectual stereotypes. At the same time, such stereotypes can positively affect other students not targeted by the stereotype and raise their performance. And, importantly, research finds that even brief interventions—a few school workshops, an in-class writing exercise, or reviewing a survey of older students' experiences in the transition to one's school—can lead to large improvements in students' academic performance, raising grades months and years later (Yeager and Walton 2011).

This research suggests that intellectual performance is not simply lodged within individuals—high test scores and straight As are not owned by a person alone, nor are poor scores or bad grades only the student's responsibility. Rather, intellectual performance is an emergent property of persons and social situations—an interaction between the two. These findings suggest that conceptualizing "intelligence" as a stable property of individuals and a reliable way of distinguishing between them may be inappropriate.

For schools and organizations, understanding intellectual performance as socially situated carries important implications. It suggests that an important task for schools and educators is to sculpt educational environments to eliminate negative social and psychological dynamics that undermine students' experience and performance and contribute to group disparities. The effectiveness of existing psychological interventions suggests the promise of this approach to improve students' outcomes. These psychological strategies may be one way schools, potentially even schools that operate in socially, economically, and racially disadvantaged environments, can effectively raise student outcomes (see Rothstein, this volume).

Insofar as mitigating negative psychological dynamics contributes to the core academic mission of schools, how well schools accomplish this should be an important component of measures of school quality. Determining how to assess school quality is exceedingly difficult, as Ladd and Loeb discuss in this volume. But creating sensitive measures of subtle psychological processes poses further challenges. Although a full discussion of how to assess these processes is beyond the scope of the current chapter, an experimental approach may be of use. For instance, if interventions aimed at reducing stereotype-related threat in a school improved students' outcomes in a randomized field experiment, it would suggest that stereotype threat was present in the school—that it was there to be remedied—and, further, that the school is making progress in mitigating this dynamic. Such evidence would be further strengthened if the intervention was shown to affect important intervening psychological processes.

Relatedly, the socially situated nature of intelligence points to a new mission for the state in general and for selective schools in particular: instead of assessing individuals' intelligence with an eye toward identifying the "best" candidates for admission and promotion, academic environments should be structured to ensure that all students can perform as well as possible (see Guinier 2011).[A] However, when selective schools do decide whom to admit and whom to reject, it is important to consider psychological processes like stereotype threat that can systematically disadvantage members of devalued groups, causing them to perform less well than they are capable of (Walton and Spencer 2009). Insofar as stereotyped students like African Americans and

---

[A] Out yet? Date still OK?

other ethnic minorities often face other important disadvantages, such as a greater incidence of poverty (Phillips et al. 1998; Rothstein 2004), this compounds disadvantage with disadvantage. When indicators of merit are systematically biased, it is unfair and discriminatory to interpret them at face value (see Walton, Spencer, and Erman, in press). Instead, schools and organizations need to take psychological dynamics into account—to understand intellectual performance as a product of both situations and persons—to interpret indicators of merit and to evaluate candidates fairly and accurately.

<EOF>