Gender Stereotype Threat among Women and Girls

Diana E. Betz, Laura R. Ramsey, and Denise Sekaquaptewa

INTRODUCTION

Can the mere existence of gender stereotypes undermine women and girls, even when women and men seem to be treated the same? Are women held back by subtle reminders of gendered expectations, despite comprising the majority of American college students? To address these questions, we examine the well-established and broad-reaching field of stereotype threat research.

Stereotype threat occurs when a situation raises concern that one will be judged in terms of group stereotypes (Steele, 2010). This can affect all kinds of people in many situations, but its impact on women’s underrepresentation in mathematics and science fields is of particular importance. Stereotype threat research makes us rethink the way we frame gender gaps in these fields. Rather than blaming women’s personal preferences, whether socialized or ‘hard-wired,’ threat suggests that responsibility may lie with environmental influences. It demonstrates how the most unsuspecting situations can impact women and girls’ performance, interests, and self-views. This focus on external rather than internal influences also paves the way for solutions, ways that we can shape environments to foster women’s and girls’ interests and talents. This chapter will benefit anyone who wonders why there are more men than women in science and engineering programs, anyone who wants to see more female students pursue non-traditional passions for math, computer science, military, or law enforcement careers, and anyone who questions whether existing gender divides really reflect some inevitable natural order.

Recognizing the relevance of stereotype threat to the lives of women and girls, this chapter aims to connect theory and laboratory experiments with real-world consequences. We first outline the theory, including its necessary and sufficient conditions. We then catalogue threat effects observed in women of different ages, considering how girls’ social-cognitive development maps onto the development of stereotype threat. We review evidence for threat’s presence in real evaluative situations and detail interventions that can reduce it. Finally, we discuss how stereotype
GENDER STEREOTYPE THREAT

threat research is currently being shared and applied in educational settings.

**ORIGINS OF STEREOTYPE THREAT THEORY**

Social psychologist Claude Steele was perplexed by a racial achievement gap persistent enough to reach elite halls of higher learning. Steele saw black, Latino, and Native American students being consistently outpaced by white students, even among cream-of-the-crop students who entered college with near-perfect Scholastic Aptitude Test (SAT) scores (Steele, 2010). Crucially, he also observed this pattern of underperformance among women, who tended to fall behind men in advanced mathematics classes. Noting that these groups shared in common stereotypes impugning their ability, Steele posited that merely knowing about such stereotypes was burdensome enough to harm academic performance. Steele’s hunch, which would form the crux of stereotype threat theory, was soon confirmed: Steele and Aronson (1995) demonstrated that performance gaps could be closed by changing testing situations to reduce the relevance of stereotypes.

Steele’s first demonstrations ushered in a huge body of research showing again and again that mere stereotype knowledge could affect all kinds of people in all kinds of situations. Of note, Spencer, Steele, and Quinn (1999) were the first to report empirical evidence of stereotype threat effects on women. When they described a math test as producing gender differences (i.e., the threat condition), women performed substantially worse than men; when they described the test as not producing gender differences, there was no difference between women’s and men’s performance. Although a recent meta-analysis demonstrates that sex differences in math and science aptitude are dropping at rates unexplainable by biological change (Hyde & Mertz, 2009), this stereotype remains prevalent, continuing to affect women’s performance and interests in the lab as well as outside of it. This chapter catalogues the conditions that engender and reduce such stereotype threat effects in the hope that disseminating this information can improve women’s and girls’ real-world outcomes.

**THEORY: NECESSARY AND SUFFICIENT CONDITIONS**

The most well documented outcome of experiencing stereotype threat is diminished test performance. According to stereotype threat theory, when tested on a domain in which one’s group is stereotyped as a poor performer (e.g., women testing in math or spatial ability), one becomes concerned about a possibly poor performance seeming stereotype-confirming in one’s own eyes or in the eyes of others. Paradoxically, test performance is thus diminished, perhaps as concerns about stereotypes disturb one’s concentration on the test or increase the test taker’s anxiety (see the ‘Moderators and Mediators’ section of this chapter for a full discussion of possible explanations of how stereotype threat causes diminished performance). When the threat of appearing stereotype-confirming is removed from the situation, members of stereotyped
groups score as well as non-stereotyped group members (e.g., women match men’s math test scores).

Although more than 15 years’ worth of research on stereotype threat has been published, some remaining misconceptions should be addressed up front. Perhaps chiefly, stereotype threat can be mistakenly perceived as stereotype internalization, or believing that stereotypes about one’s group are true. In fact, research shows that one does not have to endorse a given stereotype in order to be affected by stereotype threat (Huguet & Régner, 2009; Wheeler & Petty, 2001); one need only be aware that others may endorse it. Although personal stereotype endorsement can exacerbate the threat experience (Kiefer & Sekaquaptewa, 2007; Schmader, Johns, & Barquissau, 2004; Shapiro, 2011), it is not a core requirement.

Also largely inaccurate is the idea that threat makes people give up in the face of a challenge. To the contrary, individuals who suffer from threat are not inherently weak-willed, nor made so by situational reminders of negative stereotypes. Threatened individuals may indeed shy away from stereotype-relevant evaluative tasks (Davies, Spencer, & Steele, 2005) or expect to perform more poorly on them (Brodish & Devine, 2009; Kray, Galinsky, & Thompson, 2002; Kray, Thompson, & Galinsky, 2001; Stangor, Carr, & Kiang, 1998). Further, repeated threat experiences may decrease interest in or identification with the stereotyped domain (Steele, 1997); but in the moment, poor performance under threat is not for lack of trying (Huguet & Régner, 2007; Jamieson & Harkins, 2007, 2009; Keller & Dauenheimer, 2003).

That covers what stereotype threat is not; so what is stereotype threat? Researchers generally agree that a stereotype threat experience has three fundamental characteristics: one must be a member of a stereotyped group in a situation where the stereotype is salient while engaging in a task that can be evaluated. These conditions appear with some frequency for many women and girls learning and working in traditionally masculine fields, increasing their likelihood of experiencing threat.

The most basic criterion for experiencing stereotype threat is simply being a member of a stereotyped group. Further, that group identity must be currently salient, whether because of chronically strong identification (e.g., Schmader, 2002) or because something in the environment makes that identity stand out (e.g., being the only member of one’s group; Sekaquaptewa & Thompson, 2002).

Since all social groups can be targeted by a stereotype, theoretically anyone can experience stereotype threat. Indeed, stereotype threat effects have been demonstrated in a wide variety of groups, ranging from people of low socioeconomic status (e.g., Croizet & Claire, 1998) to white men (e.g., Stone, Lynch, Sjomeling, & Darley, 1999). However, much stereotype threat research focuses on low status groups because they are targeted by more numerous and negative stereotypes, and thus face particularly grim consequences. Stereotypes reinforce lower social status for non-dominant groups, such as women and racial and sexual minorities, in order to maintain the status quo (Berger, Rosenholtz, & Zelditch, 1980; Glick & Fiske, 2001; Jost & Banaji, 1994). For example, low-status groups are systematically stereotyped as less competent (Fiske, Cuddy, Glick, & Xu, 2002) but more communal (Conway, Pizzamiglio, & Mount, 1996) than high-status groups. This justifies the subordination of low-status groups as well as the economic and political power enjoyed by high-status groups. Furthermore, stereotypes of low status groups may be more persistent; Sekaquaptewa and Espinoza (2004) found that violations of stereotypes by low-status groups (but not high-status groups) lead to biased processing that maintains that stereotype.

The second primary criterion for the experience of stereotype threat is that the stereotype must be salient. This can happen in a variety of ways. In some studies, participants are verbally reminded of the stereotype prior to a test of their performance in that domain (e.g., Spencer et al., 1999). Of course, explicit
reminders should not be necessary, as threat occurs in everyday testing situations in which stereotypic group differences are well known (Steele, 2010). Thus, merely highlighting that a task measures a stereotyped skill (e.g., by explicitly labeling tasks ‘math tests’ or ‘intelligence tests’ or by emphasizing their diagnosticity) is enough to make stereotypes salient (e.g., Steele & Aronson, 1995).

Stereotype salience can also be manipulated to reduce the experience of threat, as when experimenters call spatial ability tests ‘puzzles’ or inform participants that a given math test reveals no gender differences. However, real-world evaluative tasks need little finessing to seem stereotype-relevant and thus may be harder to reframe in order to reduce threat: most students know very well what placement exams are meant to measure.

The boundaries between the criteria of identity and stereotype salience are often blurry, as stereotypes can also be made salient by emphasizing a stereotyped identity (e.g., Shih, Pittinsky, & Ambady, 1999). A striking example comes from Danaher and Crandall (2008), who reported that asking students for their demographic information, including gender, before an AP (Advanced Placement) test significantly reduced their performance compared to collecting demographic information after the test. Subtler still, Ambady and colleagues (Ambady, Paik, Steele, Owen-Smith, & Mitchell, 2004) induced stereotype threat among women taking a math test by subliminally priming women’s gender identity.

The final primary criterion for the experience of stereotype threat is the opportunity for evaluation of performance in the stereotyped domain, thus creating the opportunity to confirm or disconfirm the stereotype. Jamieson and Harkins (2010) found that women only underperformed on a math test when the gender stereotype was made salient and they were told that scores on the test would be compared between genders; stereotype threat effects were eliminated when women were told that test scores would be averaged across gender, even when the stereotype remained salient.

Often evaluation takes place in contexts where the individual is concerned about confirming the stereotype in the eyes of other people (Shapiro & Neuberg, 2007). However, stereotype threat effects have been shown even when performance is totally private (e.g., Inzlicht & Ben-Zeev, 2003; Shapiro, 2011; Wout, Jackson, Danso, & Spencer, 2008), demonstrating that it is the possibility of evaluation in the stereotyped domain, irrespective of audience, that is necessary for stereotype threat to occur. Some have even posited the existence of multiple types of threat: self-threat may be elicited by worries over private confirmation, whereas group-threat may arise from worries over public confirmation (e.g., Shapiro, 2011; Wout et al., 2008).

Akin to evaluation is task difficulty: there is some evidence that the performance task must be difficult in order to elicit stereotype threat effects, suggesting that threat may become more likely as students progress through school (O’Brien & Crandall, 2003; Neuville & Croizet, 2007; Shapiro et al., 1999; Walton & Spencer, 2009). The reasoning is that an easy test does not really provide the opportunity for evaluation because presumably everyone will do well, whereas distinctions (and thus stereotype confirmation) can be determined with a more difficult test. However, evidence that test difficulty is a necessary requirement for threat has not been wholly consistent (e.g., Stricker & Bejar, 2004).

Women and girls are particularly likely to encounter stereotype-threatening evaluative situations because gender and its attendant stereotypes are so frequently highlighted in our culture. Gender is pointed out frequently through linguistic labels (referring to people as ‘men’ or ‘women’, ‘guys’ or ‘girls’), gendered appearance norms, and segregation both subtle (tendencies for adults to have same-sex friends) and explicit (girls’ and boys’ lines in classrooms; Bigler & Liben, 2007). Further, gender’s central role as a social category (Gurin & Markus, 1989) fosters the dissemination of gender stereotypes,
such as those conveyed through advertising (Media Education Foundation, 2010). Even children quickly learn that gender matters and that they ought to pay attention to it, which encourages stereotyping from a young age (Bigler & Liben, 2007). Furthermore, women with multiple negatively stereotyped identities may face additional threats, whether because of additive stereotypes (e.g., mathematical intelligence stereotypes for Latina women; Gonzales, Blanton, & Williams, 2002) or stereotypes unique to certain groups (e.g., sexual stereotypes specific to African American women).

GENDER STEREOTYPE THREAT OUTCOMES AMONG WOMEN

Given these criteria, stereotype threat has the potential to affect women’s performance in any domain in which they are negatively stereotyped. For example, women are vulnerable to stereotype threat effects in financial decision-making (Carr & Steele, 2010), problem-solving (Carr & Steele, 2009), managerial tasks (Bergeron, Block, & Echtenkamp, 2006), a novel visuospatial task (Campbell & Collaer, 2009), driving (Yeung & von Hippel, 2008), leadership (Davies et al., 2005), and political knowledge (McGlone, Aronson, & Kobrynowicz, 2006). However, in keeping with the theory’s origins (Steele, 2010; Steele & Aronson, 1995) and befitting its serious consequences for academic and career-related domains, most gender stereotype threat research has focused on test performance in mathematics and science.

As previously mentioned, Spencer, Steele, and Quinn (1999) were the first to find that women underperformed on a math test when it was described as typically producing gender differences, but equaled men when told that the test did not produce gender differences. Stereotype threat effects among women taking math tests have since been replicated numerous times (e.g., Aronson & Steele, 2005; Davies, Spencer, Quinn, & Gerhardstein, 2002; Good, Aronson, & Harder, 2008; Martens, Johns, Greenberg, & Schimel, 2006; Walsh, Hickey, & Duffy, 1999).

In addition to direct effects on test-taking, stereotype threat has other consequences that could ultimately undermine girls’ and women’s academic outcomes. For example, women under threat are more likely to avoid stereotyped domains, as when they choose verbal over math test items (Davies et al., 2002) and show less interest in stereotype-relevant tasks (Smith, Sansone, & White, 2007) and math-related careers (Davies et al., 2002). Women under threat have difficulty producing new math problem-solving strategies (Quinn & Spencer, 2001), and stereotype threat also inhibits women from learning (Rydell, Rydell, & Boucher, 2010; Rydell, Shiffrin, Boucher, Van Loo, & Rydell, 2010; Sekaquaptewa & Thompson, 2002). The latter finding is particularly troubling because interventions aimed at promoting threat-free recall will be of little help to women prevented from learning in the first place.

Women under threat may experience a number of additional psychological consequences relevant to academic performance and identities, including increased anxiety (Ben-Zeev, Fein, & Inzlicht, 2005), internal attributions for their failure (Koch, Muller, & Sieverding, 2008), self-stereotyping (Gustafsson & Bjorklund, 2008), and rejection of feminine traits seen as incompatible with math success (Pronin, Steele, & Ross, 2004). Furthermore, Inzlicht, McKay and Aronson (2006) argue that effort devoted to dealing with threat depletes cognitive resources that could otherwise be devoted not only to the current math task but also to self-control in general. As a result, women may show ‘threat spillover effects’ beyond threatening evaluative environments, by acting more aggressively and eating less healthy foods (Inzlicht & Kang, 2010).

GENDER STEREOTYPE THREAT OUTCOMES AMONG GIRLS

The stereotype threat research emphasis on college students’ outcomes may underestimate
the prevalence of this phenomenon. The capacity to experience threat has been documented among high school (Keller & Dauenheimer, 2003), middle school (Good, Aronson, & Inzlicht, 2003; Huguet & Régner, 2007), and even elementary school girls (Ambady, Shih, Kim, & Pittinsky, 2001). Its effects on academic outcomes make stereotype threat uniquely relevant to educators. Thus, teachers should understand when and how these effects manifest for girls of different ages.

Much evidence demonstrates that stereotype threat effects can occur well before students enter college. Keller and Dauenheimer (2003) showed that high school girls performed less well than boys (average age 15.7 years) on a math test that was described as stereotype-relevant but matched boys when the test was described as gender-fair. Huguet and Régner (2007) demonstrated that 11- to 13-year-old girls recalled fewer details of a complex figure when they thought it was a test of geometry rather than memory or drawing, or when they performed alongside boys rather than other girls.

Evidence for stereotype-induced underperformance before middle school is somewhat more spotty. Muzzatti and Agnoli (2007) found stereotype threat effects in samples of fifth- and eighth-graders but not second-, third-, or fourth-graders. The authors argued that stereotype threat becomes more likely as girls gain awareness of gender stereotypes about math. Ambady and colleagues (2001) similarly found that subtly reminding sixth through eighth grade Asian American girls of their gender, lowered their standardized math test scores. In contrast to Muzzatti and Agnoli’s findings, however, girls in kindergarten, first, and second grade also performed less well after their gender was made salient (by coloring pictures of stereotypically female targets). Further, third- through fifth-grade girls actually did better after their female identities were highlighted. Importantly, the age groups harmed by threat were more likely to associate boys with greater math skills, whereas third- through fifth-grade girls tended to associate girls with math (Ambady et al., 2001). Future research may test stereotype awareness or endorsement as a mediating variable to better understand how threat may affect young children.

Developmental Insights

Although some evidence suggests that young girls may experience stereotype threat, this population is far less studied than college women. Reflecting on the necessary conditions for stereotype threat (stereotyped group membership, stereotype salience, potential for evaluation) alongside children’s social-cognitive development may shed light on when stereotype threat effects can occur for young girls.

With respect to the first criterion of group membership, children become aware of their gender very early. Infants can tell women from men as early as age six to nine months, and toddlers can verbally and nonverbally distinguish between genders by 18–24 months (Halim & Ruble, 2010). Further, children can identify themselves as girls or boys by 27 months or earlier (Halim & Ruble, 2010). Thus, girls as young as two years may arguably meet the criterion of stereotyped group membership. The development of a more multi-faceted connection between self and gender, or gender identity, however, continues on until age six or seven, as children grasp the concepts of gender stability (i.e., boys grow up into men and girls into women, not vice versa) and gender consistency (i.e., a girl has a ‘girl essence’ that does not go away even when she dresses like a boy; Halim & Ruble, 2010). Good and Aronson (2008) argue that children require a relatively sophisticated understanding of their gender identities before they can be said to experience stereotype threat, but the precise level of gender identification needed for stereotype threat to occur has yet to be demonstrated empirically.

Stereotype threat also depends on stereotype awareness, salience, and relevance; as it turns out, children learn to stereotype quickly.
Relying on social category cues to organize their social world (Bigler & Liben, 2007), children are particularly likely to rigidly expect specific traits or behaviors from so-called ‘natural’ categories such as gender and race (i.e., to engage in essentialist thinking; Bigler & Liben, 2007). Gender is particularly easy to stereotype, as gender is marked more frequently and with less taboo than other identities such as race (Ambady et al., 2001; Arthur, Bigler, Liben, Gelman, & Ruble, 2008; Bigler & Liben, 2007), and children directly learn gender stereotypes from their parents (e.g., feedback about gendered play) and the media (e.g., female characters portrayed as less active than male; Halim & Ruble, 2010). As a result, infants aged 18–26 months can match gender-typed toys, objects, and appearance features to boys’ and girls’ faces. By age five, children know that girls are gentle and boys are adventurous, among other traits. Importantly, when asked to draw students who are good at math, girls as young as six more often draw boys than girls, indicating early awareness of the math-gender stereotype (Beilock, Gunderson, Ramirez, & Levine, 2010). The criteria of stereotype salience and awareness, at least for the domain of math, may be fulfilled by the first grade or earlier.

Finally, to experience stereotype threat, one’s performance must have the potential to confirm a stereotype, whether privately or publicly. However, research has shed little light on either type of concern among children. For instance, little is known about children’s meta-perceptions of their gender. That is, research on children’s gender identities tends to focus on private regard (how I feel about being a boy or a girl) while ignoring public regard (how I think other people feel about boys and girls; Halim & Ruble, 2010). Moreover, we know more about how children feel about their race than about their gender (Halim & Ruble, 2010). To understand when threat emerges among children, research must examine children’s stereotype-confirmation concerns and awareness of others’ stereotypic judgments.

**IMPLICATIONS FOR FEMALE REPRESENTATION IN SCIENCE, TECHNOLOGY, ENGINEERING, AND MATHEMATICS (STEM) FIELDS**

Understanding and reducing stereotype threat is important because it impedes women’s entry into male-dominated fields. Although many researchers focus on math stereotypes and outcomes, women’s supposed ineptitude (and attendant threat effects) extends across STEM fields, targeting science (Nosek et al., 1999), engineering (Logel, Walton et al., 2009), and computer skills (Koch et al., 2008). Indeed, women remain outnumbered by men across STEM majors, graduate programs, faculty rosters, and non-academic jobs, particularly in the hard sciences, computer science, and engineering (AAUW, 2010; National Science Foundation, 2007). Women earn 38.4% of bachelor’s degrees awarded in these fields and hold just 24.2% of jobs (National Science Foundation, 2007).

Even more troubling, situational threat experiences present a particularly strong barrier to those women who are most interested and invested in STEM. Math-identified women are especially vulnerable to threat (Spencer et al., 1999), women in advanced college math courses report feeling threat (e.g., Good et al., 2008), and academically motivated women respond to threat conditions with increased performance-avoidance goals (Smith et al., 2007). In other words, accumulated threat experiences harm the ‘vanguard’ of the population (Steele, 1997; Steele, Spencer, & Aronson, 2002), thus robbing the field of the very women who might shatter stereotypes by persisting in STEM. Given the performance deficits, anxiety, and self-doubt engendered by threat, it is hardly surprising that women might want to avoid these experiences by leaving (or failing to enter) STEM fields.

Note that this loss of female talent actually perpetuates stereotypes and thus stereotype threat. Because women leave fields where they are vulnerable to stereotype threat, women
continue to be underrepresented in these fields. This further damages female students’ experiences in STEM fields, perpetuating a lack of female role models in male-dominated fields (a known buffer against threat effects; Marx & Roman, 2002) and contributing to a potentially unwelcoming or even hostile environment for new entrants. Further, because threat operates largely below public awareness, its role in the STEM gender gap remains hidden and underrepresentation is instead attributed to women’s intrinsic lack of interest or talent. Thus, stereotype threat experiences produce a cyclical process, both promoting the stereotype that women are not capable of success (leading women to avoid STEM) and contributing to the continuation of threat experiences (as lack of women in STEM promotes the stereotype).

MODERATORS AND MEDIATORS

Once stereotype threat had been demonstrated, researchers turned their attention to how, why, and for whom stereotype threat influences outcomes. This research has identified several factors that may exacerbate threat effects (i.e., moderating factors). Many can be interpreted as heightening one of the basic three criteria for stereotype threat effects. For example, women who are highly identified with their gender show increased vulnerability to stereotype threat (Schmader, 2002; Wout et al., 2008), as do women who find themselves to be the only woman in a group of men (Huguet & Régner, 2007; Inzlicht & Ben-Zeev, 2000; Sekaquaptewa & Thompson, 2002), presumably because their stereotyped group identity becomes more salient. Relatedly, women who individuate themselves (i.e., emphasize their individuality rather than their group membership) prior to taking a math test did not show the negative effects of stereotype threat (Ambady et al., 2004), perhaps because their gender identity became less salient.

Other moderators of the stereotype threat effect seem to enhance the salience or relevance of the stereotype. For example, women who are invested in the stereotyped domain (e.g., being highly motivated to pursue math) are more vulnerable to stereotype threat (Aronson et al., 1999; Spencer et al., 1999; Stone et al., 1999). Women who endorse gender stereotypes, both implicitly (Kiefer & Sekaquaptewa, 2007; Nosek et al., 2009; Ramsey & Sekaquaptewa, 2011) and explicitly (Schmader et al., 2004) are more vulnerable to stereotype threat effects as well, perhaps because the stereotype remains chronically salient for these women. Similarly, women who are high in stigma consciousness (i.e., the chronic self-awareness that one is a member of a stigmatized group) show larger performance decrements under threat than women who are low in stigma consciousness (Brown & Pinel, 2003). Interacting with others who endorse the stereotype may also increase its salience, as women showed increased stereotype threat effects after they interacted with sexist people (Logel, Walton et al., 2009).

Although person-level factors play a role (whether critical, such as stereotyped group membership, or intensifying, such as stigma consciousness), threat is a primarily situational phenomenon (Steele, 1997). Even women who care about math, value being female, and are chronically aware that they are the targets of stereotypes can do well in math provided that they perform in a non-threatening environment. Threat is fundamentally driven by something in the moment – rather than something in the person – that induces a focus on one’s stereotyped identity and one’s potential to confirm it (Good & Aronson, 2008), whether to oneself or to others (Shapiro, 2011).

Other moderators of stereotype threat do not fit cleanly into the three primary criteria: stereotype threat is more likely among women who are low self-monitors (Inzlicht, Aronson, Good, & McKay, 2006), have an internal locus of control (Cadinu, Maass, Lombardo, & Frigerio, 2006), lack a coping sense of humor (Ford, Ferguson, Brooks, & Hagadone, 2004), or have low working memory capacity (Régner et al., 2010). These criteria may instead speak to the processes underlying threat.
Over the past decade, researchers have focused on uncovering the mediating processes by which outcomes for members of stereotyped groups are affected under stereotype threat. This question is important because understanding why and how stereotype threat occurs reveals the cognitive or affective processes that must be targeted to protect women from threat. Researchers have proposed a variety of different mediators, including cognitive processes, physiological responses, emotional reactions, and changes in motivation.

Perhaps the strongest evidence points to distraction or cognitive resource depletion as a key mechanism underlying stereotype threat. Schmader and colleagues (Schmader & Johns, 2003; Schmader, Johns, & Forbes, 2008) found that stereotype threat leads to decreased working memory, which directs attention and delegates resources for cognitive tasks (see also Beilock, Rydell, & McConnell, 2007). Specifically, threat has been shown to reduce women’s ability to complete a taxing cognitive assignment (identifying math equations as true or false while remembering a list of words), and such reductions in working memory can mediate the effect of threat on math performance (Schmader & Johns, 2003). Similarly, Carr and Steele (2009) found that the suppression of stereotypes, another cognitively taxing task that may rely on working memory, mediated the relationship between stereotype threat and the pursuit of unsuccessful strategies while taking a math test, which contributes to lower performance (see also Logel, Iserman, Davies, Quinn, & Spencer, 2009). Smith and White (2002) also found that stereotype-threatened participants spent more time thinking about things extraneous to the test and less time thinking about actual test items.

The logic that women under stereotype threat are distracted or have their resources depleted in some way could be applied to a host of other mediators studied. For example, Keller and Dauenheimer (2003) found that women felt more dejected while taking a math test under stereotype threat, and they theorized that those feelings of dejection could be distracting while taking the test. Women under threat may also be distracted by negative thinking (Cadinu, Maass, Rosabianca, & Kiesner, 2005).

Other research focuses on physiology rather than cognition. Under stereotype threat, women experience increased physiological arousal that could lead to their lowered performance. Ben-Zeev and colleagues (2005) have theorized that this is a primary mechanism of stereotype threat. Arousal has been shown to have negative effects on performance for difficult tasks but positive effects on performance for easy tasks (e.g., Zajonc, 1965), which maps precisely onto stereotype threat findings (Ben-Zeev et al., 2005; O’Brien & Crandall, 2003). Furthermore, researchers have found empirical evidence that women under threat show evidence of physiological arousal through skin conductance, skin temperature, blood pressure measures (Osborne, 2007), and changes in heart rate (Croizet et al., 2004). Murphy, Steele, and Gross (2007) found that women have a physiological response (as measured through heart rate, skin conductance, and sympathetic activation of the cardiovascular system) to merely watching a video that portrays a math, science, and engineering conference with noticeably more male than female participants. Thus, physiological arousal appears to be an important mediator of stereotype threat effects.

The physiological arousal that accompanies stereotype threat could be evidence of anxiety produced when one becomes aware of being the target of a stereotype in a currently relevant domain. Indeed, several researchers have documented heightened anxiety while under threat (Brodish & Devine, 2009; Hoyt, Johnson, Murphy, & Skinnell, 2010; Osborne, 2001; Spencer et al., 1999). In a study by Hoyt and colleagues (2010), women under threat regarding their leadership abilities who were the only woman in a group of men reported increased feelings of anxiety and were rated as more anxious by independent observers (see Bosson, Haymovitz, & Pinel, 2004, for
GENDER STEREOTYPE THREAT

Evidence of observed rather than self-reported anxiety’s role in threat. Furthermore, some researchers have concluded that anxiety at least partially mediates the effect of stereotype threat on performance (Brodish & Devine, 2009; Osborne, 2001; Spencer et al., 1999).

Although anxiety may be common in threat, most researchers doubt that it is the only process driving its effects. Rather, attempts to control expressions of this anxiety may actually underlie stereotype threat processes, as evidenced by reaction time data (Johns, Inzlicht, & Schmader, 2008) as well as neural imaging studies (Krendl, Richeson, Kelley, & Heatherton, 2008; Wraga, Helt, Jacobs, & Sullivan, 2007). Suppression of anxiety depletes executive resources, leading to lowered performance in the task at hand. This interpretation of the role of anxiety supports the popular cognitive resource depletion account of stereotype threat. Croizet and colleagues (2004) even theorized that observed changes in heart rate variability under threat could signal a disruptive mental load.

A final class of mediators relate to women’s motivations under threat. Stereotype threat is said to be, at its core, a threat to the self: the possibility of giving a poor performance in a domain that you care about (i.e., that you personally identify with) is particularly threatening to the self-concept compared to performance in a less personally critical domain (Schmader, 2002). Therefore, when people enter threat-relevant settings (for example, women entering a math classroom), they appraise the situation and assess their capacity to overcome the threat. When people believe that they do not have the psychological resources or ability to succeed in the threatening situation, their motivation to perform well may wane. This can be seen in Keller’s (2002) finding that women self-handicap (i.e., make a task more difficult so as to provide a ready-made excuse for poor performance) while under threat (see also Sekaquaptewa, Waldman, & Thompson, 2007). Self-handicapping seems to be qualitatively different from arousal, anxiety, or cognitive distraction mediators, since it is tied to a motivation to preserve a positive view of the self.

Relatedly, Cadinu and colleagues (Cadinu, Maass, Frigerio, Impagliazzo, & Latinotii, 2003) found that expectations of performance mediated the relationship between stereotype threat and performance; these expectations may be closely related to various motivational strategies. Women under threat are less likely to adopt a promotion focus (i.e., eagerness toward the task; Seibt & Forster, 2004) as compared to a prevention focus (i.e., heightened vigilance against failure), which can lower performance. Similarly, women under threat are more likely to adopt performance-avoidance goals (Brodish & Devine, 2009). This pattern may be especially likely for women who are achievement-motivated; Smith and colleagues (2007) found that these women were more likely to adopt performance-avoidance goals, which led to lowered absorption and interest in the task.

Of course, it is unlikely that a single mechanism can explain stereotype threat; rather, it is likely to be a complicated process involving a host of mediators. It is worth noting that several theorists have attempted to develop models that involve more than one mediator (e.g., Ben-Zeev et al., 2005; Brodish & Devine, 2009; Chung, Ehrhart, Ehrhart, Hattrup, & Solamon, 2010; Shapiro, 2011). Understanding the process of stereotype threat can help researchers develop interventions that disrupt that process, thereby reducing the harmful effects of stereotype threat.

WAYS TO REDUCE OR PREVENT STEREOTYPE THREAT

Researchers examine when and how stereotype threat occurs so that they can uncover ways to prevent it. To that end, stereotype threat intervention research has been built upon what we know about its essential criteria, moderators, and mediators. For instance, identity-based interventions alter the criteria of stereotyped group membership and stereotype salience to reduce threat. Role model
Interventions reduce the salience or relevance of stereotypes by directly countering them. Interventions that focus on women’s attributions of either anxiety or achievement address physiological and motivational mediators. Finally, self-affirmation interventions address the criterion of performance evaluation, which may threaten self-esteem, and tap into motivational mechanisms driving threat.

**Identity-Based Interventions**

Strong identification with a stereotype-relevant identity can make people more vulnerable to threat (e.g., Schmader, 2002). However, individuals often have several important social identities. Some interventions thus emphasize people’s non-stigmatized identities (e.g., elite college students) over their stigmatized identities (e.g., women; McGlone & Aronson, 2007; Rydell & Boucher, 2010). Similarly, Shih and colleagues (1999) found that emphasizing Asian American women’s ethnic identity enhanced performance on a math test, whereas emphasizing gender identity diminished performance; these results were attributed to the positive expectations that Asian Americans are good at math (known as ‘stereotype lift’; Walton & Cohen, 2003). These findings imply that the particular intersection of one’s identities determines what kind of identity-based interventions may be effective.

Other identity-related threat reduction strategies blur boundaries between groups, which may weaken the salience of one’s stereotyped group identity or reduce expectations that one’s performance will be judged stereotypically (Rosenthal & Crisp, 2006). According to social identity theory (Tajfel & Turner, 1979), people categorize themselves and others into ‘ingroups’ (those to which we belong), and ‘outgroups’ (those to which we do not belong). Drawing the lines between ‘us’ and ‘them’ sets off a process of ingroup favoritism and outgroup derogation, giving rise to prejudice and stereotyping. Threat interventions that blur the lines between ingroups and outgroups make threat less likely by de-emphasizing the stereotyped ingroup and weakening the perception that one will be stereotyped by an outgroup. In one study, women under threat who reflected on ways that men and women are similar showed fewer negative outcomes than women who thought about how men and women differ, particularly when they generated academic similarities (Rosenthal, Crisp, & Suen, 2007). Thinking about intergroup similarities may have de-emphasized distinctions between the groups, which is known to reduce intergroup stereotyping (Hall, Crisp, & Suen, 2009).

The idea that interventions that reduce intergroup stereotyping might also reduce concerns about being stereotyped inspired an intervention based on the common ingroup identity model (Mowbray & Sekaquaptewa, 2011). Emphasizing a common identity shared by members of different groups can improve intergroup relations (Gaertner & Dovidio, 2000). Membership in a shared superordinate group reduces outgroup bias and increases favoritism toward the inclusive ingroup. Extending this logic to stereotype threat, the formation of a common ingroup may make stereotyped targets feel that their stigmatized identity is less salient or that others may be less likely to evaluate their performance according to stereotypes. Indeed, Mowbray and Sekaquaptewa (2011) showed that female science majors who believed they shared a common ‘cognitive style’ with mostly male attendees at a science conference (i.e., a common ingroup identity) reported fewer stereotype threat concerns relative to women who lacked this common ingroup identity. Seeing one’s gender underrepresented in a science context can induce stereotype threat for women (Murphy et al., 2007), and making them believe they shared something in common with the people in that context, prevented this. As when group boundaries are blurred, concern about stereotyping can be reduced when people think of a former outgroup as a new ingroup by virtue of a shared social identity.
Role Model Interventions

Stereotype threat occurs when stereotypes are salient and seem relevant to the task at hand. Role models who counter that stereotype may make it seem less relevant, thus buffering women from threat. Unfortunately for women and girls entering STEM fields of study, less than 10% of faculty in US mathematics and engineering departments are women (National Science Foundation, 2007). Female students are thus less likely to have same gender mentors than male students, a cause for concern considering the demonstrated benefits of having an ingroup mentor or role model (Lockwood, 2006; Lockwood & Kunda, 1997). Intervention researchers have highlighted the need for role models by demonstrating how exposure to successful women in STEM buffers threat among women and girls (e.g., McIntyre, Paulson, & Lord, 2003). To illustrate, Marx and Roman (2002) demonstrated that college women who learned that their female experimenter was successful in math performed better on a math test than women whose experimenter was a math-competent man or a woman said not to be skilled in math.

Role models may be effective not only because they make stereotypes seem less relevant, but also because they can weaken stereotypes. In particular, external forces such as role models may be especially adept at reshaping nonconscious or implicit stereotypes as opposed to explicit stereotypes (Stout, Dasgupta, Hunsinger, & McManus, 2011). Whereas explicit stereotypes are dictated by conscious efforts to appear egalitarian or – for women already in STEM – defend one’s identity, implicit stereotypes are theorized to stem from repeated pairings in our social environments (Karpinski & Hilton, 2001). For example, men are more often seen in prominent STEM roles than women, and stereotypic messages maligning women’s skills are prevalent. This explains why even women in STEM fields report implicit male–STEM associations (Nosek, Banaji, & Greenwald, 2002).

On a more positive note, it also suggests that female–STEM associations can be strengthened by changing environmental pairings (as Blair, Ma, and Lenton, 2001, demonstrated with ‘female–strong’ associations). Female role models in STEM are one way to promote such counterstereotypic pairings. In support of this possibility, Dasgupta and Asgari (2004) showed that female college students exposed to female leaders (in the form of a ‘famous women’ quiz) showed weaker implicit men–leadership associations (see also Dasgupta & Greenwald, 2001). To the extent that lowering implicit stereotypes reduces women’s vulnerability to stereotype threat (Kiefer & Sekaquaptewa, 2007; Nosek et al., 2009; Ramsey & Sekaquaptewa, 2011), exposing women to successful counterstereotypic female role models is a promising stereotype-threat intervention strategy.

In a preliminary demonstration that environments featuring role models can reduce implicit stereotyping and stereotype threat, Ramsey and colleagues (Ramsey, Betz, & Sekaquaptewa, in Press) tested an intervention meant to mimic a supportive academic environment for female STEM undergraduates. Female STEM students in the intervention condition took a quiz about famous women in STEM and were exposed to incidental environmental messages supporting women in STEM (i.e., a flyer announcing a predominantly female math faculty panel and a pencil listing encouraging statistics about female science majors). Compared to a control condition (students who were quizzed on flowers and saw flyers and pencils without STEM references), intervention students showed significantly reduced stereotype threat concerns and increased implicit identification with math. This study provided initial support for the idea that exposure to environmental messages promoting women’s success in STEM can reduce women’s concerns about confirming stereotypes, and thus potentially enhance their STEM outcomes.
Anxiety Attribution Interventions

Because threat leads to performance-dampening physiological outcomes such as stress and anxiety (e.g., Blascovich, Spencer, Quinn, & Steele, 2001), it is possible to reduce the negative outcomes of threat by changing the way that anxiety is attributed. Interventions thus seek to help students explain their stress in a way that is not based on concern about performance. An early demonstration of this misattribution principle (Stone et al., 1999) showed that attributions for performance-related anxiety could be redirected by pointing out the possible influence of external variables, such as the lighting and temperature of the room. Reframing the threat experience as a challenge also reduces threat effects, as stress and anxiety can be relabeled in positive terms, such as eagerness (Alter, Aronson, Darley, Rodriguez, & Ruble, 2010).

One of the most exciting interventions in recent years underscores the importance of sharing stereotype threat research with educators, policy-makers, and those vulnerable to stereotype threat. As in other threat studies, college women and men completed a math test described either as stereotype-relevant or -irrelevant. But when women under threat were first taught about stereotype threat research and encouraged to attribute any anxiety during the test to a common concern about stereotypes, rather than their personal ineptitude, they performed just as well as men and as women in the stereotype-irrelevant condition (Johns, Schmader, & Martens, 2005).

Achievement Attribution Interventions

From a motivational perspective, threat may be detrimental when students feel that they lack the ability to perform well. There is a long history of research showing that academic outcomes suffer when external attributions are made for one’s success and internal attributions are made for failure. This pattern implies that failure is due to lack of ability and thus stable, enduring, and likely to reoccur. In contrast, internal attributions for success and external attributions for failure are related to academic success because this pattern implies that failure is due to lack of effort, and thus unstable, fleeting, and unlikely to reoccur (Henry, Martinko, & Pierce, 1993; Weiner, 1979). Achievement attribution researchers have demonstrated that both children and college students who attribute academic failures internally (e.g., low ability) have more negative academic outcomes, whereas those who attribute failure externally (e.g., ‘not trying hard enough’) show improvement (Brownlow & Reasinger, 2000; Diener & Dweck, 1978; Dweck, 1975; Dweck & Bush, 1976; Hong, Chiu, Dweck, Lin, & Wan, 1999; Martinez & Sewell, 2000; Peterson & Barrett, 1987).

Of importance to stereotype threat research, gender stereotypes can influence patterns of achievement attributions. Observers (male and female) tend to attribute men’s success to high ability and women’s success to luck, reflecting the gender stereotype of higher competence among males than females (Deaux & Emswiller, 1974; Dweck & Reppucci, 1973; Etaugh & Brown, 1975; Reyna, 2000). Regarding self-attributions, female college students often show a more negative pattern (attributing their failures to low ability) than males, who tend to show more self-enhancing attribution styles; this difference is most evident within male-dominated academic domains (e.g., math; Beyer, 1998; Kiefer & Shih, 2006).

Intervention studies in schools (e.g., Good et al., 2003; Wilson & Linville, 1982, 1985) have shown that students can be taught more positive attribution styles (e.g., to take credit for successes and not to internalize failures; see Robertson, 2000, for a review). One short intervention wherein college students read peer testimonials or viewed videos of peers attributing their academic struggles to temporary, unstable factors immediately performed better on a test, and showed higher grade point averages and retention rates one year later (Wilson & Linville, 1982, 1985).
Furthermore, Good and colleagues (2003) later demonstrated that a similar attributional modeling intervention protected academic outcomes only among negatively stereotyped (ethnic minority and female) students. These targeted effects were presented as evidence of the ability of attributional style training not just to improve general academic performance but to specifically attenuate stereotype threat processes.

The attribution of failure to one’s stable identity can also be diminished by de-emphasizing the biological (i.e., ‘stable’) bases of race and gender. For instance, women were less susceptible to stereotype threat when they read experiential explanations of gender differences in math compared to genetic explanations (Dar-Nimrod & Heine, 2006), and ethnic minority students performed better on a quantitative test after learning about the social construction of race (Shih, Bonam, Sanchez, & Peck, 2007). The success of short-term interventions aimed at attribution training offers hope for the development of new interventions promoting positive attribution patterns among female STEM students.

**Self-Affirmation Interventions**

Stereotype threat is threatening to the self-concept, and even conceptions of one’s ingroup (Walton & Cohen, 2007). People must perceive that they have the ability to deal with threat in order to overcome it. ‘Self-affirmation’ increases people’s psychological resources, strength, and perceived ability to deal with adversity. Interventions that make use of self-affirmation strategies thus help people feel like they have the strength and motivation to deal with threat.

Using a simple writing exercise, Cohen and colleagues (Cohen & Garcia, 2008; Cohen, Garcia, Apfel, & Master, 2006) reduced threat appraisals among African American seventh-graders by having them complete a series of short writing assignments about their ‘personal values’ such as religion, family, and relationships with friends. Control groups wrote about value-neutral topics. The self-affirmation exercise bolstered their self-integrity, which increased their perceived ability to cope. As a result, these students earned significantly higher GPAs than control group students at the end of the academic year.

Women in physics who completed the self-affirmation writing exercise at the beginning of an academic term also showed improvements in performance compared to a control group (Miyake et al., 2010). This result emerged particularly among women who initially endorsed gender-STEM stereotypes, suggesting that this intervention technique may be particularly useful for women who are most vulnerable to stereotype threat: those with greatest awareness of the stereotype and concerns that it may be true (Shapiro, 2011).

**APPLICATIONS: GENDER STEREOTYPE THREAT BEYOND THE LAB**

Because stereotype threat effects and the methods to reduce them have been so reliably demonstrated in the lab, researchers, educators and others have a great interest in knowing its generalizability to real-world evaluative contexts. Does the threat of confirming stereotypes have a significant impact on important life outcomes such as performance on standardized tests in educational and vocational fields? This question is of vital importance because test scores often serve as important gateways into higher education, better jobs, and advancement not only in academia or business, but also within male-dominated institutions such as the military. If the threat of appearing stereotype-confirming impedes the test performance of stereotyped targets while those who are not stereotyped take tests free from this burden, this has immense real-world implications for social justice. Thus, researchers have been keen to assess the degree to which stereotype threat operates outside of the lab, in common testing situations.

At first glance, translating laboratory research methods on stereotype threat to applied settings may seem to be a relatively straightforward
task. However, there are several significant issues to be considered. First, many laboratory procedures used to induce or reduce stereotype threat are not feasible in applied contexts. For example, it is unrealistic to describe a standardized academic test such as the US SAT as ‘non-diagnostic’ of ability, or to describe a quantitative test as ‘gender fair’ (Spencer et al., 1999) when in fact men and women do historically score differently on the test. Second, real-world testing environments are likely fraught with sources of uncontrolled error variance that can reduce the detection of stereotype threat effects on performance. Such sources of error variance, including the gender and racial compositions of groups in the testing setting (e.g., Inzlicht & Ben-Zeev, 2000; Sekaquaptewa & Thompson, 2002), are often addressed and manipulated in laboratory settings, but are unlikely to be controlled or measured in most everyday test settings. These uncontrolled factors quite likely add unexplained variability to data in real-world settings, obscuring the actual effect of stereotype-confirmation concern. Finally, because real-world testing settings typically involve high-stakes outcomes, such as opportunities for education, training, or employment, an ethical question arises. How can some test-takers be randomly assigned to conditions that are predicted to diminish their outcomes?

Despite the complexities surrounding the task, researchers have focused efforts on testing stereotype threat predictions in real testing situations. Some of this work has used existing data to test models of what should emerge in the data if stereotype threat is indeed operating (e.g., Cullen, Hardison, & Sackett, 2004). Other research examined stereotype threat in real-world classrooms using lab-based materials and manipulations (e.g., Croizet & Claire, 1998; Keller & Dauenheimer, 2003). Finally, a small number of studies actually manipulated threat on real college placement tests, which determined meaningful outcomes for the test-takers. Namely, test-takers were randomly assigned to indicate their race and gender either prior to or after completing a real placement test (Stricker & Ward, 2004).

Using these various approaches, evidence has emerged supporting real-world effects of stereotype threat. First, experiments on stereotype threat interventions implemented in real classrooms (described above) provide important demonstrations of how stereotype threat is evinced in student experiences, as female and ethnic minority students performed worse than students not targeted by negative stereotypes in non-intervention control groups, and benefitted in groups receiving the interventions (Cohen & Garcia, 2008; Cohen et al., 2006; Good et al., 2003; Miyake et al., 2010; Wilson & Linville, 1982, 1985). Perhaps demonstrating the universality of stereotype threat effects, researchers have shown that stereotype threat lowered in-class test scores for German adolescents (Keller & Dauenheimer, 2003), and French undergraduates (Croizet & Claire, 1998) targeted by stereotypes (female students in math, and students from low income backgrounds, respectively). Finally, stereotype threat appears to influence non-academic real-world outcomes as well. When a soccer task was presented as irrelevant to stereotypes about women’s athletic skills, female soccer players’ performance improved (Chalabaev, Sarrazin, Stone, & Curry, 2008).

Although there is solid experimental evidence of how stereotype threat affects outcomes, results of other studies seemingly refute the applicability of stereotype threat effects outside of the lab. In the best-known demonstration of how threat seemingly failed to emerge in an actual testing situation, Stricker and Ward (2004) used a standard manipulation of stereotype threat, indicating one’s race and gender either prior to or after taking a test (Steele & Aronson, 1995), among-test takers in two settings: high school Advanced Placement (AP) calculus testing, and community college students taking Computerized Placement Tests (CPT). They compared groups in which test-takers reported their gender and race either before or after taking the test, and used two specific criteria.
to determine whether results indicated both ‘statistical and practical significance’. Specifically, results were considered ‘significant’ if the result met the statistical significance criteria of \( p < 0.05 \), and the associated effect size \( \eta \) was greater than or equal to 0.10 (following guidelines for effect size significance described by Cohen, 1988).

By these indicators, Stricker and Ward concluded that the stereotype threat manipulation had no significant effect on men’s and women’s AP calculus test scores and CPT scores. Other researchers have also failed to find significant effects of threat outside of the lab (e.g., Cullen et al., 2004; Gillespie, Converse, & Kriska, 2010). The lack of significant translation of stereotype threat findings (so robust in the lab) to real-world test-taking contexts is striking in light of intervention work showing that actions based on stereotype threat theory improve outcomes for stereotyped group members. To address this inconsistency, Danaher and Crandall (2008) reanalyzed Stricker and Ward’s (2004) data, providing a different picture of these results. Specifically, they applied an \( \eta \) criterion of 0.05 rather than 0.10 (retaining the \( p < 0.05 \) criterion). Using this small change, Danaher and Crandall argued that the effect of asking participants to report their race or gender either before or after the test had ‘very large practical effects, with real theoretical meaning’ (p. 1647), particularly regarding AP calculus scores. These authors proposed that the practical size of this effect shows that girls improved their performance from 67% to 88% of boys’ performance, simply by reporting their gender after completing the calculus test rather than before. Accordingly, the authors concluded that in 2004, 4,763 more female students would have received AP calculus credit (by virtue of their improved scores) as a result of delayed reporting of their gender.

Sackett, Hardinson, and Cullen (2004) have questioned the general interpretation of stereotype threat findings as ‘accounting for the racial gap’ (and presumably the gender gap) in test scores. They contend that because the original demonstrations of stereotype threat accounted for previous SAT scores, participants from different groups were essentially equalized in terms of previous performance. Thus, members of both groups would be expected to perform equally on the test – as occurs in ‘no threat’ conditions. The finding that group differences emerge only under ‘threat’ conditions is therefore interpreted to mean that reported group differences are an artifact of ‘adding threat’ in the lab. That is, threat manipulations are argued to produce test score differences rather than reflecting real-world testing experiences. Addressing this critique, Walton and Spencer (2009) made the case that women and racial/ethnic minorities who have had some academic success actually have greater ‘latent ability’ than non-minorities with similar test scores. In two meta-analyses, they show that when threat is removed (i.e., in ‘safe’ conditions), women and racial minorities actually perform better than non-minority men at the same level of previous performance. Stereotyped students are able to ‘catch up’ to non-targeted groups with similar prior performance only when threat is removed. This suggests that the test scores used to ‘equalize’ performance are already artificially depressed for groups targeted by negative stereotypes. This demonstrates stigmatized students’ strong underlying potential, which is too often dampened by concerns about confirming stereotypes – even on real-world tests that researchers use to ‘match performance.’

CONCLUSION

Stereotype threat research shows that awareness and concern about stereotypes in evaluative settings has a strong and negative effect on women and girls. These effects have important implications for women’s and girls’ life outcomes, such as representation in STEM fields. An important take-home point of our review is that the experience of stereotype threat is situational; the research does not imply that women underperform relative
to men in all settings, or that girls will always be less interested in mathematics than boys. Instead, stereotype threat research points out that women and girls have excitement for and can excel in domains in which they are negatively stereotyped when the environment signals that stereotypes are irrelevant. In this sense stereotype threat research is heartening; it suggests that solutions to gender gaps in educational outcomes may be addressed, perhaps in large-scale ways, by educators and policy-makers who are capable of shaping environments to reduce threat. In order to effect such change, they must first be informed about this burgeoning body of research findings.

For people looking to institute threat-free environments, stereotype threat intervention research provides some promising leads. For example, self-affirmation interventions are simple techniques that can have significant and relatively long-term benefits for stereotype-vulnerable students. In addition, providing role models, especially ingroup members modeling attainable success, can help. Although institutional change may entail considerable time and effort, there is a lot that individual teachers can do to change the climate in their classroom to reduce the relevance of stereotypes. Examples include emphasizing commonalities between gender groups (Rosenthal et al., 2007), avoiding unnecessary divisions by gender in the classroom (Bigler & Liben, 2007), and introducing counterstereotypic role models (Marx & Roman, 2002), perhaps by incorporating examples of successful women into the general course curriculum (Rios, Stewart, & Winter, 2010).

Perhaps most importantly, the negative effects of stereotype threat can be significantly reduced by simply learning about stereotype threat research in order to develop interventions, but teachers can also help their students by letting them know about the demonstrated effects of social stereotypes on women and girls. Finally, the more people know about the influence of stereotype threat, the less the gender gap may be attributed to intrinsic gender differences in ability or interests. In many ways our society has a lot to gain by increasing the general population’s awareness of stereotype threat research.

**NOTE**

For this reason, researchers, including Steele and colleagues, have introduced the term ‘social identity threat’ which includes socially privileged groups, such as whites and men, as potential targets of stereotype threat under some circumstances. However, in this chapter we retain the term ‘stereotype threat’ to connect this research clearly to societal stereotypes, in light of our focus on external influences on academic outcomes for women and girls.

**REFERENCES**


GENDER STEREOTYPE THREAT

childhood through adulthood (pp. 66–86). Oxford: Oxford University Press.


Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereo-


Carr, B. P., & Steele, C. M. (2009). Stereotype threat and inflex-


Crozet, J. C., & Claire, T. (1998). Extending the concept of stereo-
type and threat to social class: The intellectual underper-


Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardtstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academi-


Smith, J. L., Sansone, C., & White, P. H. (2007). The stereotyped task engagement process: The role of interest and


