

Stereotype Threat: Implications for the Organizational Sciences

by

Chantal Van Esch

WP-12-04

Copyright
Department of Organizational Behavior
Weatherhead School of Management
Case Western Reserve University
Cleveland OH 44106-7235
e-mail: ler6@case.edu

Stereotype Threat: Implications for the Organizational Sciences

Chantal van Esch

Case Western Reserve University

Department of Organizational Behavior

ABSTRACT

Stereotype threat, “being at risk of confirming, as self-characteristic, a negative stereotype about one’s group” (Steele & Aronson, 1995: 797) is a hot topic in psychology, but its impact on individuals in organizations has not been thoroughly considered. This paper reviews the construct of stereotype threat by reviewing its definition and origins, summarizing current conflicts in the literature, the methodology currently used, and common intervening variables which have been considered in this area. This information is then used to suggest future research directions in the organizational sciences and examine the obstacles and implications of such research.

Despite years of progress toward gender equality in America, women continue to be hampered professionally on the basis of gender, and are rarely able to obtain the same powerful and prestigious corporate positions commonly offered to their male counterparts (Catalyst, 2012). Despite boasting equal opportunity in the workplace, only four percent of Fortune 500 companies have women as CEO's (Keyishian, 2012). To make matters worse, women who do attain such high ranking positions are often the subject of gender-based criticism. Marissa Mayer, for instance, made front page news after being named the CEO of internet search giant *Yahoo!* while she was six months pregnant, a fact that, despite having little bearing on her competency to lead the company, nonetheless generated a firestorm of criticism (R. Martin, 2012). Mayer's case, unfortunately, is far from uncommon. Once women reach high ranking positions, they are often subject to speculation that they are bound to fail simply because of the preconceived stereotypes surrounding their gender, stereotypes which, sadly, often become self-fulfilling due to the phenomenon known as stereotype threat.

The purpose of this paper is to define the phenomenon of stereotype threat and introduce the research done on the subject by social psychologists. I then present the current state of the research stream, areas of disagreement within the academic conversation, the way in which stereotype threat has been tested, and the intervening variables which have shown promise in their interaction with stereotype threat. Finally, I discuss the reasons why stereotype threat should be considered in an organizational context and propose future directions.

STEREOTYPE THREAT

Definition

Stereotype threat (ST) is defined by Steele and Aronson as the "risk of confirming, as self-characteristic, a negative stereotype about one's group" (Steele & Aronson, 1995: 797), and

has been shown to cause behaviors which parallel the relevant stereotype. To clarify, Steele (Steele, 1997) wrote that stereotype threat is “the event of a negative stereotype about a group to which one belongs becoming self-relevant, usually as a plausible interpretation for something one is doing, for an experience one is having, or for a situation one is in, that has relevance to one’s self-definition” (Steele, 1997: 616).

Some researchers began framing stereotype threat by its effect on individuals. Specifically, they started looking less at the threat and more at the effect of ST but maintained the same term. Therefore, a new definition of ST formed where, “stereotype threat is usually seen as a reduction in performance induced by an anxiety of confirming a pre-existing, culturally known negative stereotype” (Nadler & Clark, 2011). The idea that stereotype threat includes the outcome of this risk is wide spread (Koenig & Eagly, 2005) but not universally accepted (Smith, 2004; Steele, Spencer, & Aronson, 2002). This fact is a definite cause for concern among academics interested in studying the phenomenon (Shapiro & Neuberg, 2003). One method in which some have chosen to address this concern is to come up with terms to describe the effect of stereotype-consistent behavior due to ST. The terms stereotype assimilation effect (Wheeler & Petty, 2001), stereotype vulnerability (Hoyt, Johnson, S. E. Murphy, & Skinnell, 2010), stereotype susceptibility (Shih, Pittinsky, & Ambady, 1999), and stereotype confirmation (Kray, Thompson, & Galinsky, 2001) all refer to such an effect. For clarity in this paper, stereotype threat (ST) will refer to the original definition and meaning as presented in Steele and Aronson (1995) work. Stereotype susceptibility will be used to refer to the effect of people acting in accordance with a stereotype because of ST.

Initial Findings

In Steele and Aronson's (1995) seminal article "Stereotype Threat and the Intellectual Test Performance of African Americans" the authors presented four studies to establish the theory of ST by testing the negative academic effects of the negative stereotype that African Americans have poor verbal ability. This was a response to finding the existing explanations of socioeconomic disenfranchisement, discrimination, and segregation insufficient to explain racial gaps in educational achievement.

The first study looked at whether or not ST would trigger a racial gap in performance on GRE verbal exam questions. Participants were randomly assigned to three conditions; diagnostic/ST, non-diagnostic, and non-diagnostic-challenge. Participants in the ST condition were informed that they were to take a verbal test diagnostic of their ability, while non-diagnostic participants were told that the study was looking at psychological factors, and finally non-diagnostic-challenge participants were told the same thing as the non-diagnostic participants, but were additionally told that the task was difficult and they should take the challenge seriously. The hypotheses – that African Americans in the ST condition would be hindered by stereotype susceptibility and that this effect would cause their scores to be lower than participants in non-ST conditions (and white participants in the diagnostic condition when controlling for SAT scores) – were supported, however, results of the whole design were found to be only marginally significant. For this reason, the authors chose to run additional studies to verify that their results were not perhaps due to chance, since this was the first study to look at this phenomenon.

The second study replicated the results from the first study and also extended it by attempting to see if anxiety contributed to the ST effect. Using the diagnostic/ST and non-diagnostic conditions from the first study, the second study additionally asked participants to

indicate their level of anxiety after completing the verbal exam. This study found significant effects for race and race by condition when controlling for SAT scores (in the direction hypothesized) with African American participants performing significantly worse under ST, while white participants' results remained unchanged across all conditions. Levels of anxiety and other measured factors – such as self-reported effort and cognitive interference – did not explain the relationship between ST and verbal performance.

The third study aimed to show that ST was causing these results and that they were not due to some other unmeasured inconsistency in the conditions. This study used the diagnostic/ST and non-diagnostic condition from the previous studies, as well as a control condition in which participants were not informed about any impending verbal exam. Participants in the diagnostic/ST and non-diagnostic conditions were told about the exam, but were first asked to take a measure which claimed to test lexical access processing but were actually measures that tested for stereotype activation (the extent to which participants were thinking about stereotypes) and stereotype avoidance (the extent to which participants avoided stereotypically African American activities). Control participants were asked to take the lexical processing task without being primed with the stereotype. First, participants filled out word fragments (some of which could have been filled out with racial or self-doubt words), then indicated their interest in certain activities (some of which had stereotypical African American associations), completed a demographic questionnaire, and finally completed a self-handicapping measure. The results supported the hypothesized interaction, that when ST is triggered in individuals the stereotype in question becomes salient in their cognitions in several ways. First, African American participants who were in the diagnostic/ST condition filled in more stereotypical racial and self-doubting words than both African Americans in the non-diagnostic and control conditions, and their white

counterparts. Second, this group avoided stereotypically African American activities (like basketball and jazz) on the questionnaire. Third, this group was less likely to indicate their race on their demographic forms. Finally, this group was more likely to indicate things that would be seen as self-handicapping (not having gotten enough sleep the night before, not being able to focus, etc). These results were interpreted to mean that a sense of ST was aroused in African American participants in the diagnostic/ST condition, and that it was causing stereotype salience and a desire to distance themselves from such stereotypes.

Finally, the last study was designed to replicate the finding that ST caused stereotype-consistent behavior from studies one and two without having an explicit prime (e.g., such as verbal announcements by the experimenter that the test measured verbal ability, the stereotyped measure) to trigger ST. Using a design similar to the non-diagnostic design from study two, in this study the ST and control conditions only differed in the timing in which participants were asked to indicate their race on the demographic form. Even without an explicit stereotype prime, the effect of making group identity salient (by designating race before taking the verbal test) still induced stereotype susceptibility – African Americans underperformed when compared to whites in studies where they were asked to indicate their race while, in studies in which race was not primed, African Americans performed similarly to their white counterparts.

In sum, the authors found that, by simply making group identity salient, African Americans would underperform in traditionally stereotyped areas. This result indicates that there is a stereotype susceptibility effect that leads individuals from the stereotyped group to act in accordance with stereotypes when ST is activated.

Construct Refinement

Subsequent research further clarified and added caveats to the concept of ST, such as when stereotype susceptibility would apply and what the effect of it would be. First, ST and stereotype susceptibility are applicable across groups and stereotypes, and are not just limited to African Americans on verbal tests (Aronson et al., 1999; Spencer, Steele, & D. M. Quinn, 1999). The concept has been applied to multiple stereotypes affecting different groups and in different domains. Stereotype susceptibility causes men to be less socially sensitive (Koenig & Eagly, 2005), white men to be less athletic (Stone, Lynch, Sjomeling, & Darley, 1999), and student athletes to be less academic (Stone et al., 1999). Second, as cultural stereotypes change, so too does the impact of ST. In France, stereotype susceptibility was seen in low socioeconomic students on academic tests (Croizet & Claire, 1998) and in Canada, findings showed that, without similar stereotypes to the US, stereotype susceptibility disappears (Shih et al., 1999). ST was also found in female Asian students as early as kindergarten, with those coloring pictures of girls with dolls significantly underperforming on math tests when compared to those who colored pictures of landscapes or children eating with chopsticks (Ambady, Shih, Kim, & Pittinsky, 2001). Third, in order for stereotype susceptibility to occur, an individual must first identify with and be in the upper echelons of the stereotyped domain (Aronson et al., 1999; Spencer et al., 1999; Steele, 1997). For example, females are only susceptible to the negative stereotype of women and math if they are proficient in math and it is important to their self-concept (Nguyen & Ryan, 2008). Fourth, there is a difference between ST and internalized feelings of group inferiority, supported by the fact that even white men fell victim to stereotype susceptibility when the stereotype of Asian superiority in math was made salient (Aronson et al., 1999; Steele, 1997). Fifth, students do not have to believe a stereotype for ST to occur, simply the knowledge that the stereotype is prevalent in society is enough to trigger ST and cause

stereotype susceptibility (Steele, 1997). Finally, ST is an additional pressure, on top of whatever pressure already exists in the situation for everyone, such that people in a stigmatized group face an additional burden (Steele, 1998).

Disagreement

Despite the multitude of studies, there are areas of disagreement. As mentioned before, even the breadth of ST is not universally accepted, as some take the theory to include the effect of the threat, while others look solely at the threat individuals face due to being confronted with a stereotype. Many dialogues have opened between Steele and colleagues with other authors who have dissenting opinions on ST. In 1998, the first exchange was about misconceptions regarding the idea that ST implied there was no truth behind individuals being judged by stereotypes and questions regarding the generalizability of the concept (Steele, 1998; Whaley, 1998). Since then, multiple dialogues between Steele, colleagues and others have occurred in different journals to clarify what effect controlling for SAT scores had on the original results, and how ST works or fails to work outside of the laboratory (Sackett, 2003; Sackett, Schmitt, Ellingson, & Kabin, 2001; Steele & Aronson, 2004; Steele & Davies, 2003). In sum, differing opinions represent the state of theory development, as there is still a lack of consensus.

One of the issues brought up by Sackett (2003) is the idea that ST might be meaningless in real world settings. While many authors disagree with this assessment (Cohen, Garcia, Apfel, & Master, 2006; Gillespie, Converse, & Kriska, 2010), Sackett is not the only one to have made such an argument (Cullen, Hardison, & Sackett, 2004; Stricker & Ward, 2004). Among the dissenting, the general consensus is that ST is not a strong enough force to cause differences in performance when there is an overarching threat of failing at a significant life event for everyone. When tested on students taking the AP Calculus exam, Stricker and Ward (2004)

found there to be no “statistically and practically significant” results of lowering ST by inquiring about demographic information either before or after the exam. When the data was reinterpreted with a lower criterion for what should be considered practically significant, Danaher and Crandall (2008) found that there were significant results. They stated that an additional 2,837 (or 17%) of young women would receive credit for the AP Calculus exam yearly. This means that almost three thousand women would start college with calculus credit and be able to start more difficult math and science classes which require calculus as a prerequisite fall semester of their freshman year. Additionally a special edition of *Human Performance* centered on stereotype susceptibility effects on personnel selection tests (Farr, 2003). Most authors interpreted their results to indicate that ST did not affect people (at least to the same extent) in real world settings. However, Steele and Davies (2003) argued that the results actually showed ST across all conditions due to researchers’ control conditions not effectively mitigating ST. Thus the debate continues on multiple aspects of ST and how it applies in real world settings.

Other Terms

Another contested aspect of ST is the effect that stereotypes can have, beyond stereotype susceptibility. In some instances, stereotypes have shown to improve an individual’s performance, and sometimes it seems that the priming increases stereotype-relevant behavior in individuals who are *not* in the stereotyped group. For instance, academics have noted multiple ways in which a stereotype can actually boost performance. The two most common theories associated with this boost are ‘stereotype lift’ and ‘stereotype reactance’ (Hoyt et al., 2010; Kray et al., 2001; Walton & Cohen, 2003). *Stereotype lift* was defined when a meta-analysis (Walton & Cohen, 2003) found that when ST was activated, the group to whom the stereotype did not apply (i.e. men and math; Caucasians and verbal ability) actually performed better in comparison

to the stereotyped group in the condition where ST had been mitigated. Therefore, negative stereotypes can provide an increase in performance when made salient to those not targeted by the stereotype. Another way in which individuals increase performance when faced with a negative stereotype is that, when ST is blatantly induced – as when the stereotypical expectation is stated outright by the experimenter or written in the materials presented – and stereotyped individuals are in the present environment in the minority (Hoyt et al., 2010; Kray et al., 2001). In such situations, individuals respond with ‘stereotype reactance’, where they respond in astereotypical manners. Priming effects of stereotypes have also been found in individuals not in the stereotyped group. College students, for example, were found to walk slower when primed with elderly stereotypical words (Bargh, Chen, & Burrows, 1996; Logel, Iserman, Davies, D. M. Quinn, & Spencer, 2009).

Meta-Analysis

Three meta-analyses of ST have been conducted, each looking to clarify different aspects of the theory. The first meta-analysis tested whether there was actually a stereotype lift effect occurring in what had, to this point, been the comparison population (Walton & Cohen, 2003). A similar idea had been put forth by Shih, Pittinsky, and Ambady (1999) who showed that Asian-American women actually over-perform when their Asian identity is made salient, as opposed to underperforming when their gender identity was made salient. Walton and Cohen (2003) found that a significant effect of increased performance of the ingroup when a negative stereotype of the outgroup was made salient. For example, men reminded of the stereotype that women are bad at math outperformed men who had been told that the exam showed no gender differences. Thus, the first meta-analysis of ST studies actually looked at side-effects of stereotype activation on other groups, rather than looking at ST specifically. Nguyen and Ryan (2008) point out this

issue, and therefore conducted their own meta-analysis. They looked specifically at ST activation in the outgroup, as well as possible moderators of this effect. Based on a meta-analysis of 116 studies, Nguyen and Ryan (2008) found that the overall mean effect size (1.261) (Nguyen & Ryan, 2008: 1314) was attributed to ST. They also found that the type of stereotype (gender or race) was a significant moderator, indicating that there may be a different type of ST activated for women as opposed to minorities. In the third meta-analysis, Nadler and Clark (2011) compared studies which tested the ST effect on African Americans compared to those on Hispanic Americans. They found an overall significant effect size of stereotype susceptibility and looked at possible moderators, but found no significant variations for race (African American or Hispanic American), being tested alone or with members of the ingroup (European Americans), or how ST was activated (explicitly or implicitly).

These studies helped to frame the literature in a more succinct way and allowed the results of multiple studies to be considered all at once, thereby providing more specific conclusions about this construct. They also provide some evidence that ST is a universal phenomenon and bring our attention to the fact that moderators of ST have yet to be determined. Therefore, the literature stream in this area still needs development.

METHODS

There is no consensus on the methodology used to investigate ST; however, there are only four major ways in which the methods diverge. These include (1) how to measure stereotype susceptibility, (2) how to measure ST, (3) how to manipulate ST, and (4) how to mitigate ST. Even the first article to present the phenomena (Steele & Aronson, 1995) dealt with these issues. The first three studies presented used a manipulation where the experimenters presented the test as either a test of a stereotyped domain or a nondiagnostic task. The last study

claimed the task to be nondiagnostic but induced a state of ST by asking half the participants to fill out demographic information before the exam, thus confounding the methods of manipulation used to induce ST (Steele & Aronson, 1995). In studies 1, 2, and 4, stereotype susceptibility was measured by performance on a verbal test, while in study 3 ST was measured by a word completion and self-description task (Steele & Aronson, 1995). The use of multiple methods for testing ST and stereotype susceptibility in the same set of studies, as well as a general lack of distinction, set the tone for the following research streams within the area of ST.

Measuring Stereotype Susceptibility

As mentioned, the first issue is there is not yet universal agreement on how to measure stereotype susceptibility. Since stereotypes differ greatly in the domains they target, each new stereotype studied comes with the difficulty of figuring out another way to test its effects. Since most of the previous research focuses on stereotypes in academic domains, there have been many applicable tests which researchers have focused on such as math (Cadinu, Maass, Rosabianca, & Kiesner, 2005) or verbal questions (Croizet & Claire, 1998) taken from the GRE, and general IQ tests such as the Raven Advanced Progressive Matrices Test (Croizet et al., 2004). Usually the decision is made by how well the stereotyped group performs relative to the non-stereotyped group in both the ST and the control condition (although it is problematic that some studies have not used a non-stereotyped group or control condition).

Measuring ST

As noted above, the second issue is the measurement of ST. Because measuring ST is as tenuous as measuring stereotype susceptibility, psychologists have often been forced to quantify the phenomenon by using a multitude of different tools. As mentioned above, the first study of ST used word fragment and self-description tasks to see if participants in the ST condition were

more likely to use negatively stereotyped words and distance themselves from stereotypical activities (Steele & Aronson, 1995), while others have attempted to use unseen forces (such as anxiety) to explain the effect. Although it has not been expressly determined to be related to ST (Blascovich, Spencer, D. Quinn, & Steele, 2001), anxiety has frequently been used to measure the effect of ST (Steele et al., 2002). However, because of the dissenting opinions surrounding anxiety as a measure, many psychologists prefer to use physiological signs (like cardiovascular reactivity) to measure ST (Mendes, Blascovich, Lickel, & Hunter, 2002) (although it should be noted that self-reported measures of anxiety do not always correlate with physiological measures of anxiety; Martin, 1961). Additionally, others have attempted to create scales to measure ST (though the study in question divided ST into general and specific threat rather than analyzing them together) (Ployhart, Ziegert, & McFarland, 2003). The lack of a reliable and universally agreed-upon ST measure has been a noted point of concern (Shapiro & Neuberg, 2007) that has impeded consensus in this area.

Manipulating ST

The third methodological issue is how to manipulate ST. Multiple mechanisms have been found to be successful in activating ST to elicit stereotype susceptibility. Originally, presenting diagnostic stereotyped exams or reporting stereotyped demographic information before the measure of stereotype susceptibility was used was the primary mechanism. Although many researchers since then have attempted to stick close to this procedure (Sackett, 2003), other researchers have attempted to find different manipulations which cause the same effect. For example, placing women in a room with two men or showing gender stereotypical characters in a commercial have both been found to trigger ST (Davies, Spencer, & Steele, 2005; Inzlicht & Ben-Zeev, 2000).

Another debate about methodological issues in manipulating ST is the difference between implicit and explicit manipulations. All the manipulations mentioned thus far were implicit manipulations of ST. However, ST can be manipulated explicitly when the experimenter or test materials include descriptions of what the stereotype is and how it is expected to impact performance. The argument has been made that explicit ST activation may lead to stereotype reactance instead of stereotype susceptibility (Hoyt et al., 2010; Kray et al., 2001). Two studies found such results in women during studies on stereotypes about female leadership and negotiations (Hoyt et al., 2010; Kray et al., 2001). A meta-analysis which ran statistics to see if there was a difference in stereotype susceptibility when ST was manipulated implicitly or explicitly, however, did not find a difference in stereotype susceptibility by manipulation across the studies (Nadler & Clark, 2011).

Mitigating ST

While there is a concern about the differences in how to manipulate ST (Shapiro & Neuberg, 2007), for the field as a whole, the final (and perhaps most important) concern is how to mitigate the ST that is generally inherent in the situation (Steele, 2010; Steele & Davies, 2003). Simply by presenting an evaluation in a stereotyped domain, one is triggering ST in stereotyped individuals (Steele & Aronson, 1995). Therefore, the main goal is to be able to remove the threat in one condition, not to cause it (Steele, 2010). Those who have failed to mitigate this threat have been critiqued, since both their ST and control conditions seem to show stereotype susceptibility (Steele & Davies, 2003).

Intervening Variables

Since the discovery of ST in 1995, many intervening variables have been analyzed with the goal of determining what effect, if any, they have although most results have proved

inconclusive. Shapiro and Neuberg (2007) explain the inconsistencies using the idea that ST is actually an umbrella theory of multiple threats which differ based on who is the target and the source of the particular threat in question. Steele et al. (2002) claim the inconsistencies are due to differences in individual factors, situational factors, and the specific stereotype which is made salient, but they do not argue that this implies different threats. In the following section, I look at some of the most common and interesting intervening variables related to ST. While this is by no means a comprehensive list, it nevertheless sheds light on some of the most important variables researchers are considering.

Mediators

Anxiety, self-efficacy, negative thinking, effort, ego-depletion, and the possibility of a mediator chain have all been considered as possible mediators of ST. Each of these is reviewed below.

Anxiety. Starting what will be a pattern for this literature stream (Blascovich et al., 2001; Spencer et al., 1999), the mediator of anxiety has shown conflicting results. The timing of asking about anxiety, method of testing, and other factors have all been listed as reasons for why these results are not congruent (Steele et al., 2002). The first study which used anxiety as a test factor found no implications of mediation when measured before a math exam on women, but cautioned against generalizing these results (Spencer et al., 1999). Two other studies found anxiety not to be a mediator – both measured pretest levels of anxiety and both were self-reports, but one looked at African American test performance, while the other looked at white athletes' miniature golf performance (Mayer & Hanges, 2003; Stone et al., 1999).

Three additional studies have found that anxiety at least partially mediates the relationship between ST and stereotype susceptibility. Of these, two were self-reports (both

measured posttest, one on African Americans on a cognitive ability task; the other on police), and the other study tested African American police and firefighters on occupational advancement exams (Blascovich et al., 2001; Chung-Herrera, Ehrhart, Ehrhart, Hattrup, & Solamon, 2005; Ployhart et al., 2003). The sole study which looked at physiological signs of anxiety during ST arousal did not test for effects of mediation, yet the results showed the typical effects of ST. It also showed that, when under ST, African Americans (who were negatively stereotyped) had increased blood pressure but European Americans did not (who were not stereotyped) (Blascovich et al., 2001). The conclusions which can be drawn based on these results are twofold. First, these studies show that self-reported pretest anxiety measures tend not to mediate the relationship between ST and stereotype susceptibility, while those measured posttest do; and second, they show that there may be a difference between self-reported anxiety levels and physiological anxiety levels in response to ST. Future research should clarify these distinctions.

Self-efficacy. As with anxiety, self-efficacy has been found to have mixed results as a mediator (Steele et al., 2002). Self-efficacy was found as a mediator between ST and selection test performance (Chung-Herrera et al., 2005), and self-esteem and self-efficacy were significantly correlated even though mediation was not tested in leadership and athletic tasks (Burnette, Pollack, & Hoyt, 2010; Stone et al., 1999). In other studies, self-efficacy was not found as a mediator between ST and stereotype susceptibility (Mayer & Hanges, 2003; Spencer et al., 1999).

Negative thinking. In contrast to positive self-efficacy, negative thoughts have also been tested as mediators (Cadinu et al., 2005). When asked to list negative thoughts during the task (i.e., in between each question), women taking math exams showed signs that negative thoughts mediate the relationship between ST activation and stereotype susceptibility (Cadinu et al.,

2005). Specifically, when the test was divided into halves, negative thought listing in the first half positively correlated with ST activation and predicted performance on the second half of the exam (Cadinu et al., 2005).

Effort. With regard to effort, the idea is that ST might cause individuals to lower their effort and therefore cause stereotype susceptibility (Steele et al., 2002). This variable has been consistently considered since the original article by Steele and Aronson (1995) specifically looked at both number of problems completed and amount of time spent on each problem. Results of these measurements of effort vary though, and while a consensus has not been reached, most researchers lean towards assuming that effort has a mediating role in the ST-stereotype susceptibility relationship (Steele, 2010).

Ego-depletion. Ego-depletion is the theory that individuals have limited stores of self-control, so using self-control in one instance (for example, when dieting) would limit one's ability to use it in another instance (for example, when controlling what one says at work) (Baumeister, Bratslavsky, Muraven, & Tice, 1998). It has been proposed that when individuals are under ST, something about the situation (hypothesized as stereotype or general thought suppression) puts the individual under cognitive load, and thereby renders him/her unable to perform the task at hand as well as s/he would otherwise. Thought suppression has been found to cause use of self-control, and is therefore relevant to ST (Baumeister et al., 1998). Surprisingly, results on ego-depletion have been comparatively consistent. Not every study tested for mediation, for example some studies found that ST leads to ego-depletion and, specifically, that ST leads to stereotype suppression as a method of inducing ego-depletion (Carr & Steele, 2009; Inzlicht, McKay, & Aronson, 2006) but these studies did not test for mediation. There have been studies that have found mediation effects though – in particular, ego-depletion has been shown to

mediate the relationship between ST and financial decision making for women, and suppression of stereotypes (a method of inducing ego-depletion) was shown to mediate the relationship between ST and women's performance on math exams (Carr & Steele, 2010; Logel et al., 2009).

Mediator chain. Smith (2004) argues that the reason these mediators have not had consistent results is because there is actually a mediator chain which explains the relationship between ST and stereotype susceptibility. According to him/her, any one mediator by itself will not consistently have enough clout to show results. Further, Smith (2004) proposes the Stereotyped Task Engagement Process (STEP) model, which suggests that ST is mediated by a chain of mediators, and that the performance goal adopted affects task behaviors and phenomenological experiences in a self-regulatory loop (Smith, 2004).

Moderators

Moderators such as domain identification, group identity, ST triggers, education, and misattribution are considered next as possible moderators of the relationship between ST and performance.

Domain identification. Domain identification was proposed as a moderator relatively early in the development of ST theory, suggesting that only those who self-identified with the stereotyped domain would be affected by stereotype susceptibility (Steele, 1997). Shortly thereafter, Aronson et al. (1999) showed that only white males who were highly identified with the domain of mathematics were affected when ST was triggered by the stereotype that Asians' are better at math than whites. Meta-analyses have found contradicting results on the moderator of domain identification. For instance, Walton and Cohen (2003) found that there was a larger stereotype susceptibility effect for ST and a larger positive effect of stereotype lift when studies chose highly identified individuals. In contrast, Nguyen and Ryan's meta-analytic findings

(2008) showed that females who were moderately identified with math were more prone to stereotype susceptibility effects than highly-math identified women.

Group identity. Although not originally considered part of ST theory, misconceptions about the original use of identification brought the focus to identification with the stereotyped group as a moderator (Carr & Steele, 2010). Not much research has tested the implications of group identification on the ST-stereotype susceptibility relationship as most just assume that having the identity is enough. However, the one study that did test this relationship found significant results with gender identification found to moderate the relationship between ST and performance in gender-stereotyped managerial tasks (Bergeron, Block, & Echtenkamp, 2006). Davies and Simmons (2009) support similar implications for racial identity as a moderating test factor for ST due to racial stereotypes.

ST triggers (*implicit vs. explicit*). Another moderator that has been suggested is how ST is triggered, and whether the trigger is implicit (not stated or written) or explicit (conveyed verbally or in writing to the participants) (Nguyen & Ryan, 2008). The argument has been made that different ST triggers cause stereotype susceptibility or stereotype reactance, but to date the results are contradictory (Hoyt et al., 2010; Kray et al., 2001; Nguyen & Ryan, 2008). Because this issue is addressed in other sections, further emphasis will not be given at this point.

Education of ST. Research has shown that when women are educated on the effects of ST, ST activation no longer leads to stereotype susceptibility (Hoyt et al., 2010). Although future research must be done as to the longevity of this effect as well as on relevant contextual factors, education as a real-world moderator that can mitigate stereotype susceptibility does show potential.

Misattribution. Whatever cognitive or emotional reactions are triggered by ST seem to cause stereotype susceptibility, but when participants have been given an alternate explanation for these feelings, the stereotype susceptibility is mitigated (Ben-Zeev, Fein, & Inzlicht, 2005; Stone et al., 1999). For example, when white males were informed that emotional arousal may be due to a newly renovated laboratory, they did not underperform in a miniature golfing task presented as a measure of athletic ability. However, when no such explanation was given, they did underperform (Stone et al., 1999). In another study, women who were taking math exams did not show signs of stereotype susceptibility (even when ST was triggered) when the experimenter explained that white noise in the room may cause feelings of anxiety. However, when white noise was explained as not having an effect on previous participants, signs of stereotype susceptibility were shown (Ben-Zeev et al., 2005).

Other Intervening Variables

Two additional intervening variables, identity salience and incentives, have been proposed. The theory presented on these variables has not specified whether these are likely to be mediators or moderators so they will be presented independently. Each will be further explored in the following paragraphs.

Identity salience. A confusing aspect of ST is that individuals have multiple identities, and any one of them might be salient to the individual during the task. Therefore, different identities may be stereotyped in different, and possibly contradicting, ways (Carr & Steele, 2010). Research has shown that, by altering which identity is made salient, triggers of ST can cause stereotype susceptibility or stereotype lift (Ambady et al., 2001; Shih et al., 1999). For example, female Asian children could do better or worse on a math test compared to controls when they colored in pictures of stereotypically Asian or stereotypically feminine activities

(Ambady et al., 2001). The same results have been found among female Asian undergrads (Shih et al., 1999), as well as with student athletes who were primed with either a student or an athlete identity before taking a math exam (Yopyk & Prentice, 2005). Questions still arise as to which identity is more likely to be salient in any given condition, and multiple seemingly logical yet possibly contradictory ideas have been presented to answer this question. Arguments have been made that whichever identity is most salient or threatening to the task at hand (Yopyk & Prentice, 2005), and whichever identity is most distinctive in the environment are the ones more likely to be salient (Mehra, Kilduff, & Brass, 1998). Clearly, more research must be done in order to understand the impact of multiple identities on the ST-stereotype susceptibility relationship.

Incentives. Another aspect of ST that is currently confusing is the impact of financial and other real world incentives on ST. Because most studies have been conducted in a laboratory (which only has limited incentives for individuals to perform well compared to the real world) the impact of incentives needs further work (Steele & Davies, 2003). Like most other intervening variables, the impact of financial incentives has had mixed results on the ST-stereotype susceptibility relationship, either by exacerbating the group difference regardless of ST activation, or by overshadowing the effect of stereotype susceptibility (Fryer, Levitt, & List, 2008; Sackett, 2003). Studies tested in the real world which had real life incentives – such as decisions of college credit, promotion, or hiring – have also found conflicting results. As mentioned earlier, no results of statistical and practical significance were found by Stricker and Ward (2004) on the Calculus AP exam when ST was triggered or not. Donaher and Crandall (2008), however, used different criteria to evaluate these same results and found that the impact was of significance. When ST was tested in fire and police stations, some results showed that

there was stereotype susceptibility, while others showed no effect (Chung-Herrera et al., 2005; Gillespie et al., 2010).

RESEARCH IN THE ORGANIZATIONAL SCIENCES

The bulk of research on ST has been conducted in academic test situations. While originally developed to explain the gaps in academic settings, research has shown that ST applies to many other settings (Steele, 2003). Previously mentioned studies showed that ST can affect behavioral as well as cognitive outcomes (Steele et al., 2002; Stone et al., 1999), which implies that ST is probably applicable to many settings, and therefore not limited to the classroom.

Although little research has been done on how ST can affect individuals in organizations, this is an area which should be investigated by organizational scientists. For instance, there have been many studies which have found ST to be applicable to situations which would be of concern to scholars in the organizational sciences. Mainly, these studies have focused on gender and racial stereotypes prevalent in organizations. As mentioned earlier, previous studies which have looked at organizations have focused primarily on personnel selection tests (Sackett, 2003; Steele & Davies, 2003). When ST literature was applied to preparation sessions for firefighter selection, some recommendations from the literature worked to reduce effects of stereotype susceptibility in this real-world study while others did not prove effective (Gillespie et al., 2010). Also, archival data showed that, when applicants were asked for demographic information before taking a cognitive exam for employment selection, the stereotype susceptibility effect appeared but was not present when this information was collected *after* the exam (Kirnan, Alfieri, Bragger, & Harris, 2009). These studies show that the effects of ST do impact individuals in organizations, and one could see how this would in turn impact the organizations themselves. For example, when individuals are able to perform at their best, the organization benefits from

increased productivity (Brough & O'Driscoll, 2010) and companies who employ individuals who are under less stress, such as that caused by ST, have been found to have stronger financial and humanistic futures (Cooper & Cartwright, 1994). For these reasons, ST is an important theory which should be applied to situations in organizations and to phenomena of interest to organizational science scholars.

Obstacles

There are two main obstacles which have hindered the real world expansion of ST research. Many articles have mentioned these in some form or another (Cullen, Hardison, & Sackett, 2004; Steele, 2010; Steele, Spencer, & Aronson, 2002). These arguments are summed in total and credit should be given to the above listed authors as well as others in the field. The first obstacle to testing the ST effect in organizations is pragmatism. As Steele and Davies (2003) note, in order to test ST effect, the threat must be removed from the situation. To do this, many studies have told participants that tests were non-diagnostic, showed no gender/race differences, or were “puzzles” or other non-stereotyped activities. Such manipulations may not be plausible in organizational settings (where, even if participants were told that a selection test was non-diagnostic, the chance that this would be believed is small). Further, there exist a multitude of implementation issues, including convincing managers to test a possibly stressful situation, finding an accurate measure to test the dependent variable used to measure stereotype susceptibility, and keeping participants in one condition from talking to those in another condition. These are all issues of concern, just to name a few. An additional obstacle to be considered in any real-life ST scenario is the ethical implications of such a study. If ST really does cause individuals in the stereotype group to underperform, then it is unethical to impose ST in a situation that determines an individual's future, even if those situations would be of value to

test. While these obstacles are of concern, they should act as a springboard for further discussion and research on the subject, perhaps with innovative methods and naturally occurring field experiments. Further inquiry into the prevalence and long-term ramifications of ST and stereotype susceptibility on individuals in organizations is needed.

Women in Organizations

In spite of these obstacles, it is important to consider the impact that ST is having in organizations, for example by investigating the effect that ST has on the gender pay gap and the glass ceiling. ST has been found to affect women in many organizational contexts. When ST is activated, women show decreased performance in management decisions (Bergeron et al., 2006), leadership (Burnette et al., 2010; Hoyt et al., 2010), negotiations (Kray et al., 2001), and financial decision making (Carr & Steele, 2010), as well as lower intentions to become entrepreneurs (Gupta & Bhawe, 2007). As mentioned earlier, simply being in a room full of men can trigger ST for women. This means that in the upper echelons, where women are over represented by men 28:1 (Catalyst, 2012), women are constantly experiencing ST. It simply has not been acknowledged or studied in organizational contexts.

Women in business have been plagued by negative stereotypes since they entered the workforce (Kanter, 1993), and much research has been done on this issue (Gatrell, C. Cooper, & Kossek, 2010). The stereotypical qualities women are seen to possess are not perceived as being consistent with good managerial practice as those which men are seen to possess. This idea is so prevalent that it has even received its own catch phrase “think manager-think male” (Schein, 1975). This pervasive mentality has led to women being left open to judgment based on negative stereotypes - whether they are just starting out, on a managerial track, or have managed to work their way into a top position in their company (Eagly & Karau, 2002). Outside of the article cited

earlier by Bergeron et al. (2006) (which did find that women underperformed in comparison to men in masculine-stereotyped managerial roles but not in feminine-stereotyped managerial roles), and the other articles mentioned which have looked at ST and managerial-related activities like leadership and negotiations (Burnette et al., 2010; Hoyt et al., 2010; Kray et al., 2001), little work has been done in real-world organizational contexts . These studies indicate that ST causes stereotype susceptibility in women in organizations, and this is important to investigate because of the implications which may arise due to it. When women unknowingly act in accordance with stereotypes because of ST, they are not only hurting their own chances for advancement, but are perpetuating the stereotype as well. This is detrimental to working women as a whole, the organizations that employ them, and society at large.

Long-Term Implications

Other than consequences for individuals in organizations, recent work has also urged researchers to consider the long-term impacts on individuals and, in turn, on society. For instance, Steele (2010) suggested that the implications are not just temporal, but may be lifelong and extremely detrimental. ST has been blamed for limiting, in part, the number of women advancing in science, technology, engineering, and math fields, causing them to be less likely to enroll in these majors and subsequently less likely to seek related career-advancing opportunities (Danaher & Crandall, 2008; M. C. Murphy, Steele, & Gross, 2007). One could assume that similar detriments are occurring to women in management fields because of the rampant stereotypes previously mentioned (Schein, 1975). Unfortunately, these stereotypes are still in effect not only in the United States but around the world (Schein, 2001). Furthermore, there is some evidence which suggests that, besides limiting opportunity, ST also causes personal health issues. ST has been shown to lead to increased blood pressure (Blascovich et al., 2001), and it is

assumed that if this increase is sustained – for example when someone works in an environment which constantly imposes ST – this could lead to hypertension (Blascovich et al., 2001; James, 1994; Steele, 2010). These reasons simply underscore the importance of investigating ST when researching organizational groups and individuals who are stereotyped.

FUTURE RESEARCH

Real World Settings

Future research should focus on the effects of ST in real world settings. While there is much to be learned from lab studies with student participants, there is also great harm in limiting studies to such methodologies. Generally, the use of student populations in lab studies brings to question the generalizability of experiments, as those working often have very different defining characteristics than students, and organizational contexts and intricacies usually cannot be accurately portrayed in lab experiments (Sears, 1986). Furthermore, there has been a push toward understanding the impact that different contexts have on studies (Johns, 2006). This brings about two issues for ST research. First, it is difficult to study the context of organizations when studies are conducted with students in a laboratory. Therefore, studies must be done in real world setting in order to understand the organizational context. Second, there may be contrasting effects of context (Johns, 2006) which will not be clarified until multiple studies have been conducted in different organizational contexts. It is important not only to start conducting studies in the field, but to conduct multiple studies across organizational contexts in order to understand the impact that context has on ST. Because of this, many calls have been made for field studies to be conducted (Carr & Steele, 2010; Sackett, 2003; Sackett et al., 2001; Steele & Davies, 2003).

Nullified ST

Beyond embarking on field studies, it is also important to consider how to effectively nullify ST. As mentioned earlier, the issue from the seminal study of ST was to find a non-ST condition (Steele, 2010). This has led to serious obstacles in field studies, since most real world situations inherently contain ST, and it is generally neither plausible nor ethical to convince participants that an exam or situation which can influence their life track is not diagnostic (Cullen et al., 2004). Not being able to remove ST from situations has led to confounding results in the past, where both control and experimental groups showed effects of stereotype susceptibility (Steele & Davies, 2003). Future research must find a manner (perhaps in training simulations) in which to achieve a nullified threat condition so that real world experiments can be compared to results found in the lab. Similarly, the influence of intervening variables like mediators and moderators needs to be studied in populations similar to those affected by stereotypes in organizations every day. Studies organized following these parameters would yield results more generalizable to working populations.

Mitigating Factors

Once a body of work in field settings has been compiled, research should focus on applying ST mitigating factors found in lab experiments to samples in organizations. This would help those stereotyped to achieve their full potential in organizations, and not be hindered by stereotype susceptibility. Thus far, there has only been one study that attempted to do this. As described earlier, this was when information from ST theory was applied to pre-test sessions for applicants who wanted to become firefighters. The multiple experiments presented in this article (Gillespie et al., 2010) showed mixed results. There have been calls to understand which critical mechanisms affect ST, and how to address these in real life situations (Ben-Zeev et al., 2005). Studies have found that organizations should try to create an identity-safe environment (Davies

et al., 2005) and teach individuals about ST (Johns, 2006) and thought suppression enhancement techniques such as misattribution or reframing (Ben-Zeev et al., 2005; Logel et al., 2009).

Another idea is to make the stereotype so blatant that stereotype reactance is triggered (Hoyt et al., 2010; Kray et al., 2001). Findings supporting this method, however, are currently extremely controversial, and contradicting findings have been found (Nguyen & Ryan, 2008). Additionally, this course of action could easily open organizations to discrimination lawsuits, so additional research would be suggested before advising such an approach. Stereotyped individuals can also help themselves by creating self-integrity statements, which emphasize attributes not affected by the negative stereotype (Cohen et al., 2006), and by learning more about ST in general, as both of these methods have been shown to remove stereotype susceptibility effects (Johns, 2006). Any or all of these measures could be used to preliminarily address the issue of how ST impacts individuals in organizations.

CONCLUSION

As pointed out by Steele and Aronson (1995) in their original work, the findings of studies based on this theory are actually immensely positive. Since it is known that stereotype susceptibility is due to contextual factors, there is something that can be done to change the environment. If the difference in outcomes were due to an inherent lack of ability in one group, then this could not be changed simply by altering the description of the task. However, since this has proven not to be the case, the possibility of leveling the playing field by eliminating situations which result in ST remains. This is a positive implication in that it allows researchers to look for a solution to this issue, and will hopefully grant society the opportunity to change stereotype susceptibility for future generations.

One of the major issues with ST is that it becomes a self-fulfilling prophecy. When a group is stereotyped, individuals in the group are subjected to ST, and because of ST they are likely to succumb to stereotype susceptibility and act in accordance with the stereotype. This results in the stereotype being reinforced for that individual and any observers who may have witnessed their response. This issue is of utmost concern since, when left to itself, it causes confounding difficulties. An external presence must put a stop to the cycle in order to halt the continuous negative downward trend. There is one advantage here, however, in that it seems that the cycle could be stopped at any point. Picking the easiest intervention point therefore advantages those who are concerned with this issue. While changing society's stereotypes about individuals may be difficult – especially for one individual in a finite amount of time – changing the context or knowledge of an individual so that they can effectively handle ST and not suffer from stereotype susceptibility is a definitely possibility. In fact, methods of doing so are already being proposed, and future research into the applications of these methods has been called for. By continuing this vein of research, academics can address the issue and attempt to stop the cycle, potentially helping to alleviate the pay gap and glass ceiling for stereotyped groups in organizations, as well as providing many other benefits.

In summary, this paper addresses the development of the research stream on ST. Based on the implications of ST and stereotype susceptibility research, a contextual shift is proposed that ST theory be explored in the organizational science fields, specifically with regard to gender and minority stereotypes which hinder progression in organizations. Finally, it is suggested that future research address this context shift, and appropriately find ways to nullify ST in organizations and to help improve organizational outcomes and the work life of those within them. These suggestions would embrace the positive aspect of ST by working on changing when

ST is triggered and when the negative effects of stereotype susceptibility are felt, as well as putting an end to the vicious cycle of self-reinforcing stereotypes.

References

- Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. 2001. Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science*, 12(5): 385–390.
- Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. 1999. When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, 35(1): 29–46.
- Bargh, J. A., Chen, M., & Burrows, L. 1996. Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71(2): 230–244.
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. 1998. Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74(5): 1252–1265.
- Ben-Zeev, T., Fein, S., & Inzlicht, M. 2005. Arousal and stereotype threat. *Journal of Experimental Social Psychology*, 41(2): 174–181.
- Bergeron, D. M., Block, C. J., & Echtenkamp, A. 2006. Disabling the able: Stereotype threat and women's work performance. *Human Performance*, 19(2): 133–158.
- Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. M. 2001. African Americans and high blood pressure: The role of stereotype threat. *Psychological Science*, 12(3): 225–229.
- Brough, P., & O'Driscoll, M. P. 2010. Organizational interventions for balancing work and home demands: An overview. *Work & Stress*, 24(3): 280–297.
- Burnette, J. L., Pollack, J. M., & Hoyt, C. L. 2010. Individual differences in implicit theories of leadership ability and self-efficacy: Predicting responses to stereotype threat. *Journal of Leadership Studies*, 3(4): 46–56.

- Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. 2005. Why do women underperform under stereotype threat? *Psychological Science*, 16(7): 572–578.
- Carr, P. B., & Steele, C. M. 2009. Stereotype threat and inflexible perseverance in problem solving. *Journal of Experimental Social Psychology*, 45(4): 853–859.
- Carr, P. B., & Steele, C. M. 2010. Stereotype threat affects financial decision making. *Psychological Science*, 21(10): 1411–1416.
- Catalyst. 2012. Women CEOs of the Fortune 1000 : Catalyst.
<http://www.catalyst.org/publication/271/women-ceos-of-the-fortune-1000>, March 18, 2012.
- Chung-Herrera, B. G., Ehrhart, M. G., Ehrhart, K., Hattrup, K., & Solamon, J. 2005. A new vision of stereotype threat: Testing its effects in a field setting. *Academy of Management Annual Meeting Proceedings*: I1–I6. Presented at the Academy of Management Annual Meeting Proceedings, Academy of Management.
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. 2006. Reducing the racial achievement gap: A social-psychological intervention. *Science*, 313(5791): 1307–1310.
- Cooper, C. L., & Cartwright, S. 1994. Healthy Mind; Healthy Organization— A Proactive Approach to Occupational Stress. *Human Relations*, 47(4): 455–471.
- Croizet, J.-C., & Claire, T. 1998. Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, 24(6): 588–594.
- Croizet, J.-C., Després, G., Gauzins, M.-E., Huguet, P., Leyens, J.-P., & Méot, A. 2004. Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin*, 30(6): 721–731.

- Cullen, M. J., Hardison, C. M., & Sackett, P. R. 2004. Using SAT-grade and ability-job performance relationships to test predictions derived from stereotype threat theory. *Journal of Applied Psychology*, 89(2): 220–230.
- Danaher, K., & Crandall, C. S. 2008. Stereotype threat in applied settings re-examined. *Journal of Applied Social Psychology*, 38(6): 1639–1655.
- Davies, P. G., Spencer, S. J., & Steele, C. M. 2005. Clearing the air: Identity safety moderates the effects of stereotype threat on women's leadership aspirations. *Journal of Personality and Social Psychology*, 88(2): 276–287.
- Davis, C., & Simmons, C. 2009. Stereotype threat: A review, critique, and implications. *Handbook of African American Psychology*, 211–222. Thousand Oaks.
- Eagly, A. H., & Karau, S. J. 2002. Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3): 573–598.
- Farr, J. L. 2003. Introduction to the special issue: Stereotype threat effects in employment settings. *Human Performance*, 16(3): 179–180.
- Fryer, R. G., Levitt, S. D., & List, J. A. 2008. Exploring the impact of financial incentives on stereotype threat: Evidence from a pilot study. *The American Economic Review*, 98(2): 370–375.
- Gatrell, C., Cooper, C., & Kossek, E. (Eds.). 2010. *Women and management*. <http://eprints.lancs.ac.uk/48294/>, March 20, 2012, Cheltenham, UK: Edward Elgar Pub.
- Gillespie, J. Z., Converse, P. D., & Kriska, S. D. 2010. Applying recommendations from the literature on stereotype threat: Two field studies. *Journal of Business and Psychology*, 25(3): 493–504.

- Gupta, V. K., & Bhawe, N. M. 2007. The influence of proactive personality and stereotype threat on women's entrepreneurial intentions. *Journal of Leadership & Organizational Studies*, 13(4): 73–85.
- Hoyt, C. L., Johnson, S. K., Murphy, S. E., & Skinnell, K. H. 2010. The impact of blatant stereotype activation and group sex-composition on female leaders. *The Leadership Quarterly*, 21(5): 716–732.
- Inzlicht, M., & Ben-Zeev, T. 2000. A threatening intellectual environment: why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, 11(5): 365–371.
- Inzlicht, M., McKay, L., & Aronson, J. 2006. Stigma as ego depletion. *Psychological Science*, 17(3): 262–269.
- James, S. A. 1994. John Henryism and the health of African-Americans. *Culture, Medicine and Psychiatry*, 18(2): 163–182.
- Johns, G. 2006. The essential impact of context on organizational behavior. *The Academy of Management Review ARCHIVE*, 31(2): 386–408.
- Kanter, R. M. 1993. *Men and women of the corporation* (2nd ed.).
<http://books.google.com/books?id=B-5GztW4IZQC>, New York: Basic Books.
- Keyishian, A. 2012, July 19. The Pregnant CEO: Should You Hate Marissa Mayer? - Forbes.
Forbes. <http://www.forbes.com/sites/learnvest/2012/07/19/the-pregnant-ceo-should-you-hate-marissa-mayer/>, July 31, 2012.
- Kirnan, J. P., Alfieri, J. A., Bragger, J. D., & Harris, R. S. 2009. An investigation of stereotype threat in employment tests. *Journal of Applied Social Psychology*, 39(2): 359–388.

- Koenig, A. M., & Eagly, A. H. 2005. Stereotype threat in men on a test of social sensitivity. *Sex Roles*, 52(7): 489–496.
- Kray, L. J., Thompson, L., & Galinsky, A. 2001. Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology*, 80(6): 942–958.
- Logel, C., Iserman, E. C., Davies, P. G., Quinn, D. M., & Spencer, S. J. 2009. The perils of double consciousness: The role of thought suppression in stereotype threat. *Journal of Experimental Social Psychology*, 45(2): 299–312.
- Martin, B. 1961. The assessment of anxiety by physiological behavioral measures. *Psychological Bulletin*, 58(3): 234–255.
- Martin, R. 2012, July 21. Focus on Marissa Mayer’s brain, not her pregnancy. *CNN*.
<http://www.cnn.com/2012/07/20/opinion/martin-mayer-pregnant/index.html>, July 31, 2012.
- Mayer, D. M., & Hanges, P. J. 2003. Understanding the stereotype threat effect with “culture-free” tests: An examination of its mediators and measurement. *Human Performance*, 16(3): 207–230.
- Mehra, A., Kilduff, M., & Brass, D. J. 1998. At the margins: A distinctiveness approach to the social networks of underrepresented groups. *The Academy of Management Journal*, 41(4): 441–452.
- Mendes, W. B., Blascovich, J., Lickel, B., & Hunter, S. 2002. Challenge and threat during social interactions with white and black men. *Personality and Social Psychology Bulletin*, 28(7): 939–952.

- Murphy, M. C., Steele, C. M., & Gross, J. J. 2007. Signaling threat. *Psychological Science*, 18(10): 879–885.
- Nadler, J. T., & Clark, M. H. 2011. Stereotype threat: A meta-analysis comparing African Americans to Hispanic Americans¹. *Journal of Applied Social Psychology*, 41(4): 872–890.
- Nguyen, H.-H. D., & Ryan, A. M. 2008. Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology*, 93(6): 1314–1334.
- Ployhart, R. E., Ziegert, J. C., & McFarland, L. A. 2003. Understanding racial differences on cognitive ability tests in selection contexts: An integration of stereotype threat and applicant reactions research. *Human Performance*, 16(3): 231–259.
- Sackett, P. R. 2003. Stereotype threat in applied selection settings: A commentary. *Human Performance*, 16(3): 295–309.
- Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. 2001. High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative-action world. *American Psychologist*, 56(4): 302–318.
- Schein, V. E. 1975. Relationships between sex role stereotypes and requisite management characteristics among female managers. *Journal of Applied Psychology*, 60(3): 340–344.
- Schein, V. E. 2001. A Global Look at Psychological Barriers to Women's Progress in Management. *Journal of Social Issues*, 57(4): 675–688.
- Sears, D. O. 1986. College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, 51(3): 515–530.

- Shapiro, J. R., & Neuberg, S. L. 2007. From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review*, 11(2): 107–130.
- Shih, M., Pittinsky, T. L., & Ambady, N. 1999. Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, 10(1): 80–83.
- Smith, J. L. 2004. Understanding the process of stereotype threat: A review of mediational variables and new performance goal directions. *Educational Psychology Review*, 16(3): 177–206.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. 1999. Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35(1): 4–28.
- Steele, C. M. 1997. A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6): 613–629.
- Steele, C. M. 1998. Stereotyping and its threat are real. *American Psychologist*, 53(6): 680–681.
- Steele, C. M. 2003. Through the back door to theory. *Psychological Inquiry*, 14(3-4): 314–317.
- Steele, C. M. 2010. *Whistling Vivaldi: and other clues to how stereotypes affect us*. W. W. Norton & Company.
- Steele, C. M., & Aronson, J. 1995. Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5): 797–811.
- Steele, C. M., & Aronson, J. 2004. Stereotype threat does not live by Steele and Aronson (1995) alone. *American Psychologist*, 59(1): 47–48.
- Steele, C. M., & Davies, P. G. 2003. Stereotype threat and employment testing: A commentary. *Human Performance*, 16(3): 311–326.

- Steele, C. M., Spencer, S. J., & Aronson, J. 2002. Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology*: 379–440. <http://www.sciencedirect.com/science/article/pii/S0065260102800090>, March 6, 2012, Academic Press.
- Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. 1999. Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology*, 77(6): 1213–1227.
- Stricker, L. J., & Ward, W. C. 2004. Stereotype threat, inquiring about test takers' ethnicity and gender, and standardized test performance. *Journal of Applied Social Psychology*, 34(4): 665–693.
- Walton, G. M., & Cohen, G. L. 2003. Stereotype lift. *Journal of Experimental Social Psychology*, 39(5): 456–467.
- Whaley, A. L. 1998. Issues of validity in empirical tests of stereotype threat theory. *American Psychologist*, 53(6): 679–680.
- Wheeler, S. C., & Petty, R. E. 2001. The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin*, 127(6): 797–826.
- Yopyk, D. J. A., & Prentice, D. A. 2005. Am I an athlete or a student? Identity salience and stereotype threat in student–athletes. *Basic and Applied Social Psychology*, 27(4): 329–336.