



# Do Findings from Laboratory Experiments on Preferential Selection Generalize to Cognitively-Oriented Tasks? A Test of Two Perspectives

Edgar E. Kausel<sup>1</sup> · Jerel E. Slaughter<sup>2</sup> · Joel M. Evans<sup>3</sup> · Jordan H. Stein<sup>4</sup>

Published online: 13 September 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

## Abstract

This investigation examined whether previous findings in preferential selection using laboratory simulations, which have used leadership tasks and perceived performance, generalize to cognitively oriented tasks and actual performance. We tested competing perspectives derived from two theoretical accounts of stereotype threat theory: regulatory focus and executive control interference. Non-stigmatized (Whites and Asians) and stigmatized (Hispanics and Blacks; total  $n = 513$ ) individuals first took a cognitive ability test to be selected for a subsequent task and a chance to win a cash prize. They were then randomly assigned to an explanation concerning selection for a proofreading task based on merit, gender, or race. Results tended to support the regulatory focus view. The main study showed there were no significant differences in performance quantity or quality among participants who were selected based on merit or gender. Among those selected on race, stigmatized participants had lower performance quantity but higher performance quality (i.e., they were slower but more accurate) than non-stigmatized participants. A follow-up study ( $n = 252$ ) found that stigmatized people selected based on race had more prevention concerns than non-stigmatized people. We discuss previous findings in preferential selection research utilizing experiments and conclude that the regulatory focus perspective can account for these results. Our research also shows that by using different outcomes, it is possible to qualify the allegedly harmful effects of preferential selection. This study is the first to experimentally examine the effects of preferential selection on actual task performance in cognitively oriented tasks.

**Keywords** Preferential selection · Stereotype threat · Laboratory experiments · Regulatory focus · Affirmative action · Minorities · Task performance · Cognitive tasks · Generalizability

---

✉ Edgar E. Kausel  
ekausel@uc.cl

Jerel E. Slaughter  
jslaught@eller.arizona.edu

Joel M. Evans  
jevans@alfaisal.edu

Jordan H. Stein  
jstein@matherlifeways.com

- <sup>1</sup> School of Management, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Macul, Santiago, Chile
- <sup>2</sup> Department of Management & Organizations, University of Arizona, P. O. Box 210108, Tucson, AZ 85721-0108, USA
- <sup>3</sup> Prince Mohammad Bin Salman College of Business and Entrepreneurship, King Abdullah Economic City, Saudi Arabia
- <sup>4</sup> Institute on Aging, Mather LifeWays, 1603 Orrington Avenue, Suite 1800, Evanston, IL 60201, USA

On June 23, 2016, closing a case referred to as *Fisher v. University of Texas II*, the Supreme Court of the USA reaffirmed that a university could weigh an applicant's race in admissions decisions (Totenberg 2016). However, given the Supreme Court's 4-to-3 ruling, preferential selection remains a contentious issue. Authors in the popular press have debated the legality and fairness of affirmative action policies (AAPs), as well as the efficacy of implementing such policies, over the last several years (Caplan 1995; Craig 1995; Levy 2016; Wade 1995; Will 2001). A particular aspect within AAPs has to do with preferential selection. A body of empirical work by Heilman and colleagues based on laboratory simulations tends to suggest that preferential selection may inadvertently thwart diversity efforts by hampering the performance of individuals it is intended to help (e.g., Heilman 2012; Heilman and Alcott 2001; Heilman and Blader 2001; Heilman et al. 1991, 1992; Turner and Pratkanis 1993). This research suggests that when women are led to believe they are preferentially selected based on gender; they report lower self-evaluations of performance.

These studies have been influential. Doverspike et al. (2000) considered Heilman and colleagues' simulation method one of the four most common approaches to studying the effects of AAPs. In a meta-analysis examining the effects of AAPs on self-evaluations of performance (Leslie et al. 2014), of the 21 studies included, 12 had used these laboratory simulations. Stewart and Shapiro (2000), after reviewing the effect of AAPs on performance, concluded: "Heilman and her colleagues have produced a critical mass of empirical research regarding the effects of affirmative action policies and procedures" (p. 230). Crosby et al. (2003) also reviewed this research and concluded that the "unintended negative effects of affirmative action are disturbing" (p. 102).

Despite these approving critiques, lab studies of affirmative action have also received criticism (Doverspike et al. 2000; Turner and Pratkanis 1993). One strong source of criticism is due to concerns about the external validity of these studies; in other words, authors have questioned the degree to which these laboratory studies mirror what occurs in actual organizations (Crosby et al. 2006). A second criticism, and the focus of this paper, concerns the fact that these laboratory simulations have previously relied almost exclusively on communication and leadership tasks. For example, Heilman et al. (1987) used a task in which participants adopted the role of a leader who then instructed a follower in drawing geometric figures. They were then asked to rate their own competence on the task. Heilman et al. found that women's (but not men's) self-evaluations were lower when women were selected based on gender relative to those selected based on merit.

The primary goal of the present study, therefore, is to examine whether previous laboratory findings based on leadership and communication tasks extend to more common cognitively oriented tasks. A second, related goal has to do with the fact that nearly all work simulating preferential selection has used only self-evaluation as the dependent variable.<sup>1</sup> Given that most of the tasks used in this work have been communication tasks, there has been no actual participant performance that could be scored quantitatively. Thus, our second goal is to test hypotheses about *actual task performance* following preferential selection.

An emphasis on cognitively oriented tasks is particularly relevant in light of two competing predictions regarding the effects of preferential selection on cognitively oriented task performance. These competing predictions are found in relation to *stereotype threat* theory (Roberson and Kulik 2007; Spencer et al. 2016; Steele and Aronson 1995), which suggests that when an individual enters a situation in which a stereotype is known to exist, concerns about confirming that stereotype arise. This extra pressure leads these individuals to underperform relative to those who do not experience the

threat (Walton and Cohen 2003, Walton and Spencer 2009). Thus, for example, selecting women or minorities based on gender or race—as opposed to merit—may cause them to experience this threat, because beneficiaries of affirmative action may infer they were only selected based on their minority status, and may then harbor doubts about their competence to perform a task or job effectively.

There are two main explanations for why people under stereotype threat decrease their performance in most tasks. The *executive control interference* view argues that stereotype threat undermines performance by increasing emotional load, which limits executive resources required by the task at hand (Inzlicht et al. 2011; Rydell et al. 2009; Schmader and Johns 2003). The *regulatory focus* perspective suggests that stereotype threat tends to affect performance by inducing regulatory foci, meaning the individual adopts either a promotion-focused achievement orientation or a prevention-focused failure orientation (Barber 2017; Barber et al. 2015; Grimm et al. 2009). Specifically, advocates of this view argue that negative stereotypes trigger a prevention focus (Barber and Mather 2013a). This prompts in turn a conservative strategy: People under prevention focus tend to avoid errors. In contrast, positive stereotypes tend to activate a promotion focus, which prompts an approach strategy, motivating people to attempt successes.

As noted, the fact that only communication and leadership tasks have been used in these studies is problematic, because both perspectives of stereotype threat can explain these results. While the decrease in performance among females who were preferentially selected can be explained by an interference in cognitive functioning, it can also be explained by a prevention focus. This is because hesitant behavior and contractive nonverbal displays—cues of a prevention focus (Fennis and Stel 2011)—are negatively related to assessments of communication skills, especially when using self-assessments (Carney et al. 2010; Cuddy et al. 2015). Thus, these two alternative perspectives on stereotype threat in experimental preferential selection research remain relatively untested from a comparative standpoint. The final goal of this research, therefore, is to test the two views of stereotype threat theory described above: the executive control interference view and the regulatory focus view (Barber and Mather 2013a, b).

The remainder of this paper is structured as follows. In the following sections, we review the literature on laboratory simulations of preferential selection. We then describe the two perspectives on stereotype threat theory that have been used to explain previous results, as well as the hypotheses derived from these views. Next, we describe the present experimental study that was conducted to test the hypotheses and conclude with a discussion of the implications and limitations of our study.

<sup>1</sup> There are two exceptions, which we discuss in the final section of the present paper (Brown et al. 2000; Turner and Pratkanis 1993).

## Literature Review, Theory, and Hypotheses

### Laboratory Simulations of Preferential Selection

As mentioned above, for the better part of 30 years, Heilman and her colleagues have been investigating the negative effects of preferential selection on how women are perceived by others (e.g., Heilman 2012; Heilman et al. 1992; Heilman et al. 1993; Heilman et al. 1996), how women perceive themselves (Heilman and Alcott 2001; Heilman et al. 1990; Heilman et al. 1987), and women's decisions and behaviors after experiencing preferential selection (Heilman et al. 1991, 2015).

Before describing some examples of the results observed in this area, it is important to describe the general methodology used in these studies, which is somewhat similar to the methodology used in the present investigation. Typically, individuals sign up for an experiment and arrive at the laboratory. They are met by the experimenter and told that they and the other participant (usually a confederate) will take a pretest that will assess their one-way communication skills and then participate in a communication task with the other participant (the confederate). In the one-way communication task, the participant and the confederate sit back-to-back, and the leader instructs the follower in the drawing of three complex geometric figures. In reality, the participant is always chosen as the leader and the confederate is always selected as the follower. The most common manipulation is the method of assignment of the participant to the leadership role. In the merit condition, participants are told that they are assigned the leader role because they scored higher on a communication pretest. In the preferential selection condition, participants are told that communication pretest scores are typically used to select a leader, but because there is a shortage of female (male) participants, on this particular day, they have been assigned the leader role specifically because the participant in question is female (male). Then, after completing the performance task, all participants complete questionnaires including items that assess the dependent variables in question.

What is particularly compelling about the results of these studies is that only female participants appear to experience negative self-perceptions associated with preferential selection on the basis of gender. For example, Heilman et al. (1987) found that females who were preferentially selected rated themselves lower on leadership ability, performance competence, and desire to remain a leader, as compared to females who were selected on the basis of pretest scores. However, the self-ratings of males on these variables did not differ across the preferential selection and merit conditions. As another example, Heilman et al. (1991) examined whether preferential selection causes individuals to lower their own perception of competence to the point that they select less challenging over more challenging tasks, when given the

opportunity to make their own task choices. Results showed that this effect occurred only for women and not for men.

Heilman and colleagues have concluded that these effects are due to multiple processes operating simultaneously. First, women have been shown to be less confident about their general performance capability than men (Barber and Odean 2001; Lenney 1977; Maccoby and Jacklin 1974). Second, this difference in self-efficacy is exacerbated when the task under consideration is a stereotypically male task (Betz and Hackett 1981). In keeping with the idea that leadership roles in organizations are stereotypically male (Eagly and Carli 2004), Heilman and her colleagues referred to the assigned role in these studies as the role of “leader.” Third, for individuals who naturally harbor doubts about their competence (i.e., women in stereotypically male roles), the absence of information confirming their ability to perform the tasks—as is the case when women are preferentially selected—only serves to further increase self-doubt. Men, who do not harbor doubts about their abilities to succeed in these roles, are not differentially affected by merit-based and preferential selection procedures (Heilman et al. 1991).

### Race-Based Preferential Selection and Cognitive Tasks

The present investigation differs from studies conducted by Heilman and colleagues in a number of ways. The first difference is that we elected to use a cognitive ability test as the pretest ostensibly used to select participants. Related to this, the second task was a proofreading task in which individuals were asked to compare a manuscript with several typos and grammatical mistakes to a master document; they were also asked to make the correct proofreading marks, which were also provided to the participant.

The second difference is that we used ethnic minorities (Hispanics and Blacks) instead of females, and the preferential selection was made on the basis of race rather than gender. Although there is one study using a laboratory simulation method in which participants are selected based on race (Stewart and Shapiro 2000; see below), Heilman et al. (1987) suggested that the stigmatization effects of gender-based preferential selection should extend to other forms of affirmative action policies: “Whenever individuals harbor doubts about their competence, regardless of whether those doubts are warranted, preferential selection on the basis of nonwork-related criteria is likely to have deleterious consequences for their self-perceptions and self-evaluation” (p. 68).

For this reason, the use of cognitively loaded tasks, and the inclusion of Hispanics and Blacks as participants, is consistent with this view. This is based on the notion that the type of task matching the stereotype of a particular group can provide the strongest conditions for stigmatization effects to occur. Research shows that Blacks and Hispanic score lower on standardized cognitive ability tests than do Whites, Non-

Hispanics,<sup>2</sup> and Asians (Berry et al. 2014; Roth et al. 2001; Rushton and Jensen 2005). This has led to the proliferation of stereotypes that Blacks and Hispanics are lower on cognitive ability and lower on academic achievement than Whites and Asians (e.g., Brown and Lee 2005; Devine 1989; Devos and Torres 2007; Weyant 2005).

We were able to locate only one experiment regarding the effects of race-based preferential selection on performance. Stewart and Shapiro (2000) conducted a replication and extension of Heilman et al.'s (1987) study, in which participants engaged in a one-way communication (leadership) task. In Stewart and Shapiro's study, however, they were preferentially selected based on race or merit (they also were selected on gender but we will not discuss these findings here). The authors found that African Americans who were selected on the basis of race, and who were given negative feedback about their performance on the leadership task, actually provided the highest self-evaluations of their abilities. The authors explained this result in terms of ego protection in order to maintain self-esteem.

Another possible explanation of these results is that there is no clear stereotype suggesting that African Americans have weaker leadership skills than Whites. As noted by Heilman et al. (1987, 1991), in order for selection based on demography to have negative consequences, there must exist a stereotype of poor performance in a domain. When selection is based on demography rather than merit, there is no information to counteract self-doubt, leading to lower self-evaluations. A leadership and communication paradigm, as used by Heilman to find stigmatization effects for women preferentially selected on the basis of gender, may be unlikely to produce the same effects for minority group members, as no stereotype is known to exist for members of minority groups and leadership ability.

In sum, three conclusions can be made from this literature review on the effect of gender-based and race-based preferential selection on performance. First, most research has focused on communication or leadership tasks. Second, in theory, race-based selection should have similar effects as gender-based selection, but the single study that examined this issue found the opposite effect, possibly due to the task. Third, little research has examined actual performance as opposed to perceived performance, because of the nature of the task (i.e., tasks in which there are non-countable outcomes).

## Two Perspectives on Stereotype Threat

Stereotype threat occurs when members of an identity group expect that others may see them according to a negative stereotype about their group, and the concern that their behavior

may confirm that stereotype (Roberson and Kulik 2007; Spencer et al. 2016). Those concerns cause people to underperform in a way consistent with the stereotype (Steele et al. 2002). In a meta-analysis, Walton and Spencer (2009) suggested that stereotype threat can explain an important amount of variance in the race gap in academics and the gender gap in mathematics. Stereotype threat, therefore, can have important real-world consequences.

As noted above, in our experiment, we used a cognitive ability pretest and a cognitively oriented task. In a cognitive ability testing situation, stereotype threat can be induced by telling individuals they are taking a test that is diagnostic of their abilities (Steele and Aronson 1995), by priming the specific stereotype (Schmader and Johns 2003), or by simply asking them to indicate their race prior to the beginning of the test (Shih et al. 1999). Among Hispanic and Black participants who were to be preferentially selected on the basis of race, we expected stereotype threat processes to occur because we told participants they were taking a test of intellectual abilities that was known to be a valid predictor of job performance, and that high scores would make them eligible for another cognitive task. While Hispanic and Black participants selected on merit would be able to ease some of the self-doubt about their abilities after being given feedback about their high scores, this doubt should remain for those selected on race (e.g., Heilman et al. 1987, 1992).

**The Executive Control Interference View** The specific way in which stereotype threat can affect outcomes is a matter of some debate (Barber and Mather 2013b; Grimm et al. 2009). One perspective is the *executive control interference* view (e.g., Beilock 2008; Schmader and Johns 2003; Rydell et al. 2009; Forbes and Schmader 2010). This view posits that stereotype threat affects crucial executive functions because people make an effort to deal with the imbalance between their positive self-views and negative self-relevant stereotypes. This effort implies utilizing their cognitive resources, which takes up resources from executive functions needed to perform optimally (Schmader et al. 2008). For example, Rydell et al. (2014) suggested that activating gender-based math stereotypes can negatively impact the executive function of *updating*, which is the ability to use attentional control to maintain and update relevant information while performing tasks. Rydell et al. (2014) suggested that this executive function was the main mediator in explaining why stereotype threat reduced math performance among women. Other studies have suggested that stereotype threat affects older adults' ability to use controlled memory processes (e.g., Mazerolle et al. 2012) and executive attention processes (Richeson and Shelton 2003). Schmader and Johns (2003) found that women and Hispanics who were exposed to stereotype threat had reduced working memory capacity.

<sup>2</sup> For brevity, henceforth, we use the term "Whites" to refer to individuals who are White, non-Hispanic.

Thus, from the executive control interference perspective, stereotype threat should decrease performance quality. Because Hispanic and Black individuals should be more likely to spend their cognitive resources dealing with the concern of confirming a stereotype, they will have fewer resources available to perform the task (e.g., attention processes, working memory, and updating). As a result, these individuals are expected to make more mistakes, thereby reducing their performance quality. In terms of performance quantity, however, stereotype threat should induce the opposite effect and increase performance quantity. Researchers from this perspective have argued that individuals subjected to stereotype threat engage in more effort, because they wish to disconfirm the stereotype that is salient in the situation (O'Brien and Crandall 2003; Oswald and Harvey 2000; Schmader et al. 2008). Indeed, McFall et al. (2007) have argued that individuals subjected to stereotype threat could be “trying very hard during task performance to disprove the negative stereotypes directed at their group” (p. 562). As such, this effort should produce enhanced speed or performance quantity (see O'Brien and Crandall 2003), but at the expense of more errors due to memory interference (Rydell et al. 2009).

**The Regulatory Focus View** According to regulatory focus theory, there are important differences in the processes through which people pursue their goals (Higgins 1997, 1998; Li et al. 2011). Higgins (1997) argued that two distinct self-regulatory systems could account for these differences. In the first system, people have a promotion focus in which they put emphasis on attainment aspirations, advancement and accomplishments. The second system deals with a prevention focus, in which people put emphasis on protection, safety, and responsibility (Brockner and Higgins 2001). Crowe and Higgins (1997) proposed that because a promotion focus is concerned with a strategic inclination to make progress, people are concerned with goal attainment and with ensuring “hits.” Because a prevention focus is concerned with a strategic inclination to be prudent and precautionary, people are concerned with avoiding mistakes. Consistent with this view, Wallace and Chen (2006), using 50 work groups, found that while prevention focus was related to safety performance, promotion focus was related to productivity. Similarly, in four experiments, Förster et al. (2003) found that people with a promotion focus had higher performance quantity but lower performance quality than those with a prevention focus.

A number of researchers have proposed that stereotype threat is related to regulatory focus (e.g., Barber and Mather 2013a; Wong and Gallo 2016). Grimm et al. (2009) found that priming people with a negative stereotype induced a prevention focus, whereas priming individuals with a positive stereotype induced a promotion focus (see also Seibt and Förster 2004). The rationale behind this idea is that a negative stereotype activates a negative reference point, which in turn triggers

the adoption of a minimal goal. Not meeting a minimum goal becomes the negative event, while meeting it is the non-negative event. This induces a prevention-focused state of alertness. In contrast, a positive stereotype sets a positive reference point, which triggers the adoption of a maximum goal. Achieving it becomes the positive event, while not achieving it becomes the non-positive event. Seibt and Förster (2004) found that positive stereotypes led to better performance in a creativity task (where quantity and divergent thinking are important), whereas negative stereotypes led to better performance in an analytical task (where lack of errors is important). Similarly, Barber and Mather (2013a, b) found that stereotype threat reduced older adults' memory errors (intrusion rates during free-recall tests, as well as decrease of false memories).

The regulatory focus view, therefore, makes different predictions than the executive control interference view. In our particular context, Hispanics and Blacks who are selected on the basis of race should experience stereotype threat. Consistent with previous findings, this should create a prevention focus, which in turn would decrease performance quantity. However, given that these individuals would be more vigilant and make fewer errors, they would increase their performance quality. In short, this view predicts that Hispanics and Blacks selected based on race would be slower but more accurate.

Interestingly, in the context of our experiment, this view would also suggest that Whites (non-Hispanics) and Asians who are selected on the basis of their race would increase their performance quantity and decrease their performance quality. Because a positive stereotype is made salient (i.e., Whites and Asians perform well on cognitive ability tests), selecting Whites and Asians on the basis of their race would induce in them a promotion focus. In turn, this would make Whites and Asians faster (more quantity) but careless (less quality).

In sum, both perspectives predict an interactive effect between race and selection method on performance, but the form of this predicted effect is different. In both perspectives, stigmatized (Hispanics and Blacks) and non-stigmatized (Whites and Asians) individuals selected based on merit have similar performance levels. However, among those preferentially selected based on race, the predictions are in opposite directions. The executive control interference view predicts that Hispanics and Blacks will have better performance quantity but lower performance quality than Whites and Asians. In contrast, the regulatory focus view predicts that Hispanics and Blacks will have better performance quality but lower performance quantity than Whites and Asians.

In addition, in order to test the notion of match between task and stigmatized groups, we also included a condition in which participants were preferentially selected based on their gender. Because the main task involved in this experiment (the proofreading task) was cognitive in nature, and more verbally oriented than quantitatively oriented, women who were selected based on gender were unlikely to feel stereotype threat.

Thus, in contrast to previous results (e.g., Heilman et al. 1987), we expected to find no interaction effect between gender and selection method (gender-based vs. merit-based) in either performance quantity or quality.

## Methods

### Participants

Participants were 513 undergraduate business and psychology students at a large university in the southwestern USA. They were recruited by an internet-based experiment system and were given course credit for their participation. The mean age of participants was 21.24 years ( $SD = 3.19$ ), and 306 (57.2%) of them were women. The racial composition of the sample included 289 (67.9%) Whites, 66 (12.9%) Asians, 136 (26.5%) Hispanics, and 22 (4.3%) Blacks. The mean self-reported GPA was 3.32 ( $SD = 0.42$ ), with most (81.0%) in at least their third year at the university. Of the 513 participants, 461 reported receiving their high school degree in the USA or a country whose official language was English and 52 in a different country.

Prior to conducting the analyses for the study, we eliminated from the sample all participants who were not Whites, Asians,<sup>3</sup> Hispanics, or Blacks, as our hypotheses were only relevant to individuals from these specific racial groups. Thus, we eliminated 13 participants (6 Middle-Easterners and 7 from “other races”) from the overall sample. The 513 participants reported above exclude these individuals.

### Design

The study was a  $2 \times 2 \times 3$  factorial between-subjects design, involving three independent variables: gender (male or female), race (Whites and Asians vs. Hispanics and Blacks), and selection method (merit-based, gender-based, or race-based). Similar numbers of participants (165–175) were assigned to each selection method condition.

### Procedure

Upon arriving at the laboratory, participants were provided an introductory overview of the procedures. The experimenter first announced that participants were to complete a cognitive ability test, emphasizing that the test is known to be a valid predictor of job performance. Once they had completed this test, participants were told that it would be scored by one of the experimenters. In addition, participants were informed that, depending on their scores, they might be eligible to

participate in a proofreading task. Subjects were further informed that the top three performers on the proofreading task would each earn a \$100 cash prize, and they were reminded that the opportunity to participate in the proofreading task was contingent on their cognitive ability test score.

After this introduction, participants were asked to complete a demographic questionnaire. Once the demographic questionnaire was completed, each participant was seated in a private room, where they completed the cognitive ability test. Then, an experimenter collected the test and asked participants to wait for a couple minutes while the test was scored.

### Experimental Manipulation

The experimental manipulation was similar to that used by Heilman and colleagues (Heilman et al. 1987, 1991, 1996). After completing the test, participants waited for approximately 3 min in their individual rooms, while one of the experimenters ostensibly scored their tests. In all conditions, the experimenter returned and said: “We’ve been selecting participants on the basis of skill and ability, and that’s why we used the intellectual ability test you completed a couple of minutes ago.”

What followed varied based on condition. In the merit condition, subjects were told: “Because you scored better than average on the test, you have been selected to complete the task.” In the preferential selection condition, however, the experimenter indicated that the study required a certain percentage of individuals from a specific race in order to ensure that the sample accurately represented the demographic profile of the University’s student body. Participants in these conditions were told they were being selected based on their specific race, depending on the respective condition. For example, Hispanic participants in the race-based preferential selection condition were told,

But today we’re going to have to do things a little differently, because we need at least 15% of the participants completing the proofreading task to be Hispanic, given that Hispanic individuals comprise 15% of the overall student population at (university name). So regardless of how you did on the test, because you are Hispanic, you have been selected to complete the task.

Similarly, individuals in the gender-based preferential selection condition were told the following:

But today we’re going to have to do things a little differently, because we need at least 53% (47%) of the participants completing the proofreading task to be female (male), given that females (males) comprise 53% (47%) of the overall student population at (university name). So regardless of how you did on the test, because you are a female (male), you have been selected to complete the task.

<sup>3</sup> Because Asian Americans receive higher scores on cognitive ability tests (Roth et al. 2001), we included them in the non-stigmatized group, along with Whites.

## Proofreading Task

Participants were then given 12 min to complete the performance-based task, which consisted of proofreading a 678-word business-related article. Participants received the instructions and a list of the notation marks required to carry out the task, as well as a sample of how a corrected article should look. In addition, two versions of the article were handed to subjects. “Master” was the original version of the article, without errors. “Proof” was almost identical to Master, but it contained 27 errors. Subjects were asked to find the errors in Proof and to make the appropriate corrective marks on it: One mark on the text and another mark on the margin (as suggested in the *Proofreader’s Marks* of the APA manual, Fifth edition; American Psychological Association 2001, pp. 337–338). The instructions also informed participants that their final score would be a composite of quantity (identify a large number of errors) and quality (use correct symbols and avoid false recognitions). In this way, we mentioned both aspects of performance, and participants could center their attention on either of these, depending on their regulatory focus.

## Dependent Variables

Our two dependent variables were participants’ quantity and quality of performance. Quantity was measured as the number of attempts participants made to identify errors. In other words, this was the number of marks made (Fong and Tosi (2007) called this measure “effort”). Performance quantity ranged from 0 to 35 ( $M = 13.36$ ,  $SD = 5.17$ ).

Quality was measured by a ratio. The numerator was the composite of number of correct recognition of errors, plus the number of times participants used the appropriate marks, minus the number of correction of errors that did not exist (for a similar measure, see Fong and Tosi 2007). The denominator was the number of marks made. For each mark participants made, performance quality could range from  $-1$  (an attempt of correcting an error that did not exist) to 3 (a correct recognition of error, an appropriate mark made on the text, and an appropriate mark made in the margin). The quality score of those participants who did not make marks was set to 0 (15 participants).<sup>4</sup> The minimum quality score earned was 0 and the maximum was 3 ( $M = 1.88$ ,  $SD = 0.77$ ).

<sup>4</sup> The number of participants who did not make any marks was not trivial. That said, when we were pilot-testing this study, we conducted some post-session informal interviews. One participant who had no marks explained that in the 12 min he had to complete the task, he was unable to figure out which marks to use. Thus, the task may have been perceived as difficult by some individuals, which may explain why a few of them did not record any marks. Based on this assessment, we decided to conduct the analyses with and without those who had recorded no marks. The results were robust to this test; they did not change by including or excluding these participants.

## Results

### Manipulation Check

Following the proofreading task, participants were provided an open-ended form asking them to explain why they believed they were selected for the proofreading task. It was expected that participants would infer their selection based on either their performance on the cognitive ability test (in the merit condition) or on their race (in the race condition). The vast majority of the participants (446 of 513, or 87%) indicated the correct method of assignment. Of the 67 participants who did not answer as expected, 21 were in the merit condition, 29 were in the gender condition, and 8 were in the race condition. Examples of people who indicated the wrong reason for being selected include “To see how I would perform in a performance test” (merit condition) and “So they could understand what type of student are (sic) at (university), in other words how good students are at proofreading” (race condition). Although these are not what we defined as correct, there were no cases of individuals in the preferential selection condition indicating that they were selected on merit, nor were there any cases of individuals in the merit condition indicating that they were selected preferentially. Thus, we retained all cases for data analysis.<sup>5</sup>

### Test of Hypotheses

Table 1 presents the means, standard deviations, and correlations of the major study variables. Before testing the hypotheses, for each outcome, we first ran three-way ANOVAs with selection method (merit vs. race), race (Whites and Asians vs. Hispanics and Blacks), and gender as independent variables, to test potential three-way interactions. Using this same analysis, we then tested hypotheses with two-way interactions (selection method and race as the independent variables). After this, when appropriate, we used the Fisher LSD method (Hays 1994; Kirk 1995) and conducted follow-up pairwise contrasts to test simple effects. In addition, we also examined two-way interactions between selection method (merit vs. race) and gender as the independent variables, to test the matching idea (i.e., whether differences between males and females differed depending on the selection method).

**Performance quantity** Recall that, in terms of performance quantity, the executive control interference view predicts that Hispanics and Blacks selected based on race would have higher performance (due to more effort) than Whites selected based on race. The regulatory focus view predicts that

<sup>5</sup> Results reported below were robust to including or excluding people who answered the manipulation checks incorrectly.

**Table 1** Means, standard deviations, and intercorrelations of main study variables

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7
1. Gender	0.57	0.50	–						
2. Race	0.36	0.48	.02	–					
3. Age	21.22	3.18	–.04	.08	–				
4. GPA	3.33	0.42	.05	–.19	–.07	–			
5. Dummy race	0.34	0.47	–.01	–.03	.05	.00	–		
6. Dummy gender	0.34	0.47	–.03	.09	–.03	–.02	–.51		
7. Performance quality	1.88	0.77	.02	.05	–.06	.14	–.02	–.01	
8. Performance quantity	13.36	5.17	.09	–.04	–.10	.06	–.05	.04	–.08

*N* = 513. For gender, 0 = male, 1 = female. For race, 0 = Whites and Asians, 1 = Hispanics and Blacks. Dummy race = first dummy variable, contrasting race and merit as selection methods; second dummy variable, contrasting gender and merit as selection methods. Correlations  $|\text{.09}|$  and higher significant at  $p < .05$

Hispanics and Blacks selected based on race would have lower performance quantity than Whites selected based on race.

We first conducted an omnibus three-way ANOVA. As seen in Table 2, we found that the three-way selection method  $\times$  race  $\times$  gender was not significant,  $F(2, 501) = 1.49$ , *ns*. In the same table, it can be seen that the selection method  $\times$  race interaction effect was significant,  $F(2, 501) = 4.03$ ,  $p < .05$ ,  $\eta^2 = .02$  (see Table 3 and Fig. 1). After this, we conducted pairwise contrasts to determine simple main effects. We found no significant differences between stigmatized (Hispanic and Black) individuals ( $M = 13.51$ ;  $SE = .71$ ) and non-stigmatized (White and Asian) individuals ( $M = 13.47$ ;  $SE = .49$ ) who were selected based on merit,  $F(1, 507) = .00$ , *ns*. However, among those selected based on race, we did find that stigmatized individuals had poorer performance ( $M = 11.44$ ;  $SE = .66$ ) than non-stigmatized individuals ( $M = 13.61$ ;  $SE = .48$ ),  $F(1, 507) = 8.55$ ,  $p < .01$ ,  $d = -.49$ . These results supported the regulatory focus view; the executive control interference view received no support.

Regarding differences between men and women depending on the selection method (merit vs. gender-based), we failed to find a significant interaction. In particular, we found that the gender  $\times$  selection method term was not statistically significant,  $F(2, 501) = 2.09$ , *ns*.

**Table 2** Analysis of variance table for performance quantity, main study

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	$\eta^2$
Selection method (SM)	121.26	2	60.63	2.33	0.10	.01
Race (R)	11.90	1	11.90	0.46	0.50	.00
Gender (G)	53.09	1	53.09	2.04	0.15	.00
SM $\times$ R	209.58	2	104.79	4.03	0.02	.02
SM $\times$ G	108.90	2	54.45	2.09	0.13	.01
R $\times$ G	40.18	1	40.18	1.54	0.22	.00
SM $\times$ R $\times$ G	77.32	2	38.66	1.49	0.23	.01
Error	13,039.56	501	26.03			

**Performance quality** Recall that, in terms of performance quality, the executive control interference view predicts that Hispanic and Black participants selected based on race would have lower performance than Whites and Asians selected based on race. The regulatory focus view predicts that Hispanics and Blacks selected based on race would have higher performance quality than Whites and Asians selected based on race.

As with performance quantity, we first conducted a three-way omnibus ANOVA. Table 4 and Fig. 2 shows that the three-way selection method  $\times$  race  $\times$  gender was not significant,  $F(2, 501) = .39$ , *ns*. However, we did observe a significant race  $\times$  selection method interaction,  $F(2, 501) = 5.69$ ,  $p < .01$ ,  $\eta^2 = .02$ . Follow-up analyses suggested that, within participants selected based on merit, there were no significant differences between stigmatized (Hispanic and Black) individuals ( $M = 1.82$ ;  $SE = .11$ ) and their non-stigmatized (White and Asian) counterparts ( $M = 1.93$ ;  $SE = .07$ ),  $F(1, 501) = .62$ , *ns*. In contrast, within participants selected based on race, Hispanic and Black participants demonstrated higher performance ( $M = 2.16$ ;  $SE = .10$ ) than Whites and Asians ( $M = 1.73$ ;  $SE = .07$ ),  $F(1, 501) = 12.10$ ,  $p < .001$ ,  $d = .57$  (see Table 5). Results concerning performance quality, therefore, supported the regulatory focus view and did not support the executive control interference view of stereotype threat.<sup>6</sup>

When testing whether differences between men and women depending on the selection method (merit vs. gender-based), as with performance quantity, we found no such effect. The gender  $\times$  selection method interaction was not statistically significant,  $F(2, 501) = 1.69$ , *ns*.<sup>7</sup>

<sup>6</sup> We also conducted these analyses, for performance quantity and quality, using GPA, highest degree (English speaking country or not), and cognitive ability as covariates. Results remained unchanged in terms of statistical significance. All these results are available from the first author upon request.

<sup>7</sup> Similar results were found when we operationalized performance quality differently. We used the number of correction of errors that did not exist. In other words, this is the number of marks that should not have been made. There was a significant race  $\times$  selection method interaction,  $F(2, 501) = 4.68$ ,  $p < .05$ ,  $\eta^2 = .02$ . Within participants selected based on race, minorities showed less errors ( $M = .73$ ;  $SE = 1.85$ ) than non-minorities ( $M = 2.94$ ;  $SE = 4.71$ ),  $F(1, 501) = 10.96$ ,  $p < .001$ .



**Table 3** Performance quantity means (standard errors) across conditions, main study

Selection method		Race		Total
		Whites and Asians	Hispanics and Blacks	
Merit	Males	13.48 (0.72)	14.47 (1.24)	13.98 (0.72)
	Females	13.47 (0.65)	13.06 (0.85)	13.26 (0.54)
	Total	13.47 (0.49)	13.51 (0.71)	13.62 (0.45)
Gender	Males	12.55 (0.77)	13.03 (0.88)	12.78 (0.58)
	Females	13.76 (0.69)	15.10 (0.81)	14.33 (0.53)
	Total	13.16 (0.52)	14.07 (0.60)	13.61 (0.39)
Race	Males	12.29 (0.73)	11.63 (0.98)	11.96 (0.61)
	Females	14.93 (0.64)	11.24 (0.89)	13.09 (0.53)
	Total	13.61 (0.48)	11.44 (0.66)	12.52 (0.41)

### Follow-Up Study

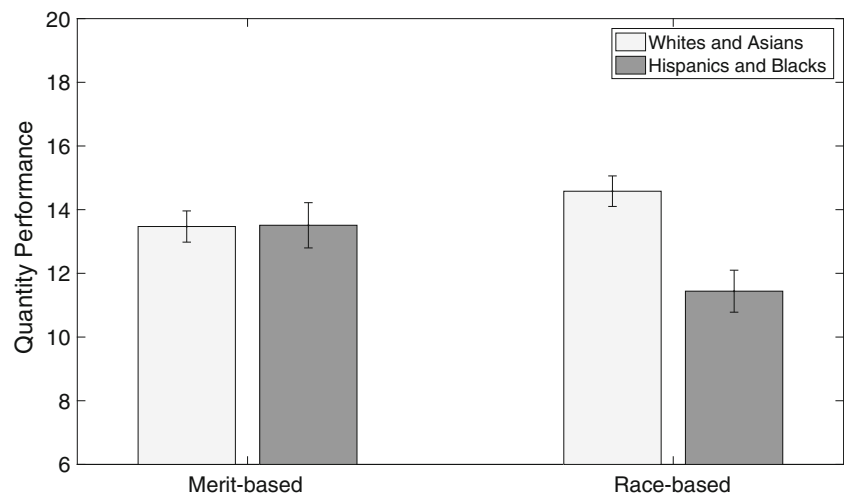
**Purpose** In the main study, we did not measure the hypothesized mechanism (i.e., regulatory focus) because by doing so we could have created demand effects or similar artifacts. The aim of the follow-up study was, therefore, to test whether the same manipulation used in the main experiment would trigger different regulatory focus levels among stigmatized vs. non-stigmatized individuals, as hypothesized. More specifically, we examined whether the effect of selection method (merit-based vs. race-based) interacted with the participants’ race (Whites and Asians vs. Blacks and Hispanics) on the participants’ regulatory foci (Table 5).

**Participants and Procedure** Two hundred and fifty-two undergraduate business students at a large university were recruited for the study in exchange of course credit. They were on average 21 years old and 132 (52%) were women. Most of them (61%) were White, 9 (4%) were Black, 31 (12%) were Asian, and 41 (16.2%) were Hispanic. Eighteen participants (7%) did not indicate their race or indicated other races. They were excluded from the analyses, which left us with a total of 234 participants.

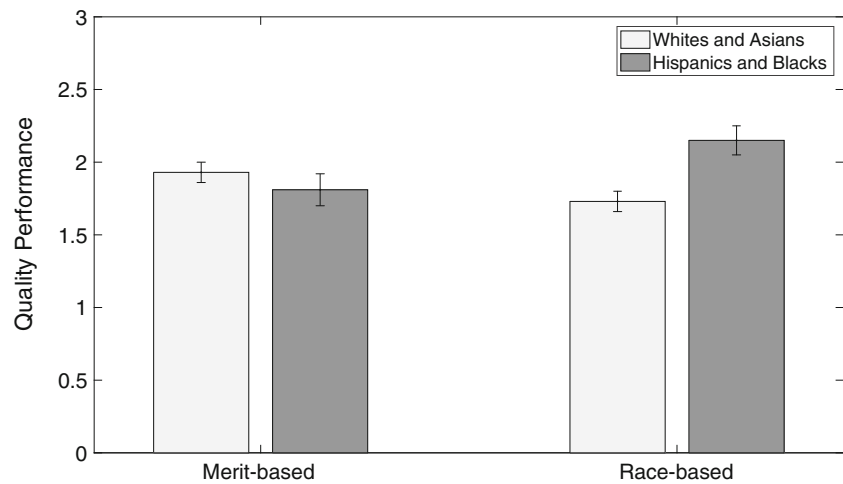
This study was nearly identical to the main study except for the final section. First, upon arriving to the laboratory, participants completed a demographic questionnaire. Second, they were asked to complete a cognitive ability test and were told that depending on their scores they might be eligible to participate in a proofreading task. They were also told that, depending on their performance in the proofreading task, they could earn a \$100 cash prize. Third, they completed the cognitive ability test. Fourth, participants were told either that (a) they were selected because of their performance in the cognitive ability test or (b) they were selected for the proofreading task because of their race, regardless of how they did on the test.

The fifth and final section of the follow-up study differed from the main study. Recall that in the main study, we asked participants to complete a 12-min proofreading task. Here, participants (a) completed three regulatory focus measures and (b) were given 4 min to complete a proofreading task. The proofreading task was not scored because it was not the focus of this study; it was included only to be consistent with the instructions given to participants at the outset. Participants took in total around 40 min to complete the whole session.

**Fig. 1** Interactive effects of selection method × race on performance quantity



**Fig. 2** Interactive effects of selection method  $\times$  race on performance quality



**Dependent Variables** We included three measures of regulatory focus: one implicit and two explicit measures. First, we used an implicit measure proposed by Johnson (2006; see also Friedman and Förster 2001). This measure seems appropriate, given that researchers argue that regulatory focus often operates outside of people’s awareness and control (e.g., Johnson and Steinman 2009). Word fragments were created by removing letters from an existing word. The fragments were constructed in such a way that participants could form promotion-oriented, prevention-oriented, or neutral words. Participants were presented with five target word fragments, in addition to some filler word fragments. For example, one of the word fragments was “a \_\_\_d.” In this example, participants had three spaces to fill. Those who wrote “award” received one point on the promotion scale score and zero otherwise; those who wrote “avoid” received one point on the prevention scale score and zero otherwise. Other example target words were “\_\_\_tive” (“positive” vs. “negative”) and “\_ain” (“gain” vs. “pain”). Participants’ promotion and prevention foci scores were created by averaging the number of promotion and prevention-oriented words that they generated, effectively creating two proportion scores.

In addition to the implicit measure, we also used two explicit measures of regulatory focus. The first explicit measure was

adapted from Wallace and Chen (2006). Participants were asked to rate 4 statements on a 5-point scale (1 = *strongly disagree*; 5 = *strongly agree*) regarding their strategy in the upcoming (proofreading) task. The statements were “In the following task, I plan to accomplish a lot” (promotion), “I plan to get a lot done in a short amount of time” (promotion), “I plan to avoid mistakes” (prevention), and “I plan to follow meticulously the task rules and instructions” (prevention). These questions were averaged for each measure. Internal consistencies were .72 for promotion focus and .73 for prevention focus.

The second explicit measure we used was created by Li et al. (2011). Participants were asked to indicate whether different items described their general mindset (“what you have been thinking about in the past 5 minutes”). There were six items in total and the four critical items were “my dreams” and “my ambitions” (promotion focus); “my worries” and “my duties” (prevention focus). Responses consisted of yes/no answers. Scores for promotion and prevention foci were obtained by averaging the number of times participants indicated “yes” (i.e., “yes” was scored as “1”; “no” was scored as “0”) for each item in the corresponding category. Thus, scores for both promotion and prevention foci for this measure were proportions that ranged from 0 to 1.

**Table 4** Analysis of variance table for performance quality, main study

Source	SS	df	MS	F	p	$\eta^2$
Selection method (SM)	0.57	2	0.28	0.49	0.61	.01
Race (R)	0.81	1	0.81	1.40	0.24	.00
Gender (G)	0.10	1	0.10	0.16	0.69	.00
SM $\times$ R	6.57	2	3.29	5.69	0.00	.02
SM $\times$ G	1.96	2	0.98	1.69	0.19	.01
R $\times$ G	0.01	1	0.01	0.02	0.88	.00
SM $\times$ R $\times$ G	0.45	2	0.23	0.39	0.68	.00
Error	289.35	501	0.58			

**Results and Discussion** Our analytical strategy was to first conduct, for each criterion (i.e., promotion focus and prevention focus), a 2 (selection method: merit-based vs. race-based)  $\times$  2 (race: Whites and Asians vs. Hispanics and Blacks) multivariate analysis of variance (MANOVA) using the implicit and the two explicit measures as dependent variables. We then conducted 2  $\times$  2 ANOVAs using each dependent variable. When appropriate, we conducted follow-up pairwise contrasts to test simple effects.

We first focused on promotion focus. The 2  $\times$  2 MANOVA revealed no selection method  $\times$  race interaction effect on promotion focus, Wilk’s  $\Lambda = 1.00$ ,  $F(3, 228) = .50$ , *ns*. We also conducted follow-up 2  $\times$  2 ANOVAs and found no interaction

**Table 5** Performance quality means (standard errors) across conditions, main study

Selection condition		Race		Total
		Whites and Asians	Hispanics and Blacks	
Merit	Males	1.81 (0.11)	1.81 (0.18)	1.81 (0.11)
	Females	2.05 (0.10)	1.84 (0.13)	1.95 (0.08)
	Total	1.93 (0.07)	1.82 (0.11)	1.88 (0.06)
Gender	Males	1.85 (0.12)	1.76 (0.13)	1.81 (0.09)
	Females	1.92 (0.10)	1.89 (0.12)	1.91 (0.08)
	Total	1.89 (0.08)	1.83 (0.09)	1.86 (0.06)
Race	Males	1.82 (0.11)	2.21 (0.15)	2.02 (0.09)
	Females	1.63 (0.09)	2.09 (0.09)	1.86 (0.08)
	Total	1.73 (0.07)	2.15 (0.10)	1.94 (0.06)

effect, both when using the implicit measure ( $F [1, 230] = .02, ns$ ) and the two explicit measures (for each:  $F [1, 230] = .54, ns$ ;  $F [1, 230] = .29, ns$ ) of promotion focus as dependent variable.

We then turned our attention to prevention focus. The  $2 \times 2$  MANOVA revealed a significant selection method  $\times$  race interaction effect on prevention focus, Wilk’s  $\Lambda = .94, F (2, 229) = 4.90, p < .01, \eta^2 = .06$  (see Table 6). Follow-up ANOVAs also suggested a significant interaction effect when using the implicit measure of prevention focus as the dependent variable,  $F (1, 230) = 7.20, p < .01, \eta^2 = .04$ . When using the first explicit measure (Wallace and Chen 2006) of prevention focus, the selection method  $\times$  race interaction effect was marginally significant,  $F [1, 230] = 3.34, p = .069, \eta^2 = .01$ . When using the second explicit measure (Li et al. 2011), the selection method  $\times$  race interaction effect was not significant,  $F (1, 230) = .60, ns$ .

Next, keeping our attention on prevention focus, we conducted pairwise contrasts to test simple effects. In terms of the multivariate test, within participants selected based on race, there were significant differences between stigmatized and non-stigmatized groups,  $F (3, 228) = 4.57, p < .01, \eta^2 = .06$ . The univariate tests suggested that, when using the implicit measure as the outcome, stigmatized people ( $M = .27; SE = .03$ ) had greater prevention focus than non-stigmatized individuals ( $M = .18; SE = .02$ ),  $F (1, 230) = 4.57, p < .05, d = .58$ . When using the first explicit measure, stigmatized individuals reported having marginally significantly more prevention focus ( $M = 4.45; SE = .12$ ) than non-stigmatized individuals ( $M = 4.21; SE = .06$ ),  $F (1, 230) = 2.75, p = .081, d = .45$ . When using the second explicit measure, a similar

significant difference also emerged ( $M_{\text{stigmatized}} = .78; SE = .09; M_{\text{non-stigmatized}} = .58; SE = .04; F [1, 230] = 4.08, p < .05, d = .50$ ). See Table 7 for means and standard deviations across conditions and racial groups.

In sum, results from the follow-up study provided support for our expectations regarding prevention focus. We found no effect of race-based preferential selection on promotion focus; we did find an effect on prevention focus, particularly when using the implicit measure. More specifically, these results suggest that when Black and Hispanic individuals were selected on the basis of their race, they had increased prevention concerns, relative to White and Asian individuals in the same condition.

## Discussion

In the present study, we examined the potential effects of specific affirmative action policies—preferential selection—on performance quality and quantity in laboratory simulations. While previous research using these simulations has generally focused on leadership and communication tasks, perceived performance, and gender-based preferential selection, we examined cognitively oriented tests and tasks, actual performance, and race-based preferential selection. We also tested hypotheses derived from two perspectives of why stereotype threat occurs: the executive control interference view and the regulatory focus view (Barber and Mather 2013b). According to the executive control interference view (Rydell et al. 2009; Schmader et al. 2008), individuals under stereotype threat increase their effort to counter the stereotype, but due to fewer available resources they are not able to perform well. The regulatory focus view proposes that stereotype threat generates a prevention motivational state because negative stereotypes activate a negative reference point (Barber and Mather 2013a; Barber 2017).

Results tended to lend support to the regulatory focus view; the executive control interference view received no support. Among those who were selected on the basis of merit, there were no differences between stigmatized (Hispanic and

**Table 6** Multivariate analysis of variance table for prevention focus, follow-up study

Source	Wilk’s $\Lambda$	df	F	p	$\eta^2$
Selection method (SM)	.99	3	0.89	0.44	.01
Race (R)	.97	3	2.57	0.06	.03
SM $\times$ R	.94	3	4.90	0.00	.06

**Table 7** Implicit and explicit prevention focus across conditions and measures, follow-up study

Measure	Selection condition	Race		Total
		Whites and Asians	Hispanics and Blacks	
Implicit	Merit	.22 (0.15)	.15 (0.12)	.20 (0.15)
	Race	.18 (0.15)	.27 (0.13)	.20 (0.16)
	Total	.20 (0.14)	.20 (0.15)	.20 (0.15)
Explicit 1	Merit	4.40 (0.55)	4.32 (0.63)	4.38 (0.54)
	Race	4.21 (0.51)	4.45 (0.56)	4.25 (0.53)
	Total	4.31 (0.54)	4.37 (0.60)	4.32 (0.56)
Explicit 2	Merit	0.64 (0.31)	0.73 (0.34)	0.66 (0.32)
	Race	0.58 (0.35)	0.78 (0.30)	0.62 (0.33)
	Total	0.61 (0.35)	0.75 (0.32)	0.64 (0.34)

Implicit: Johnson (2006) measure; explicit 1: Wallace and Chen (2006) measure; explicit 2: Li et al. (2011) measure

Black) and non-stigmatized (White and Asian) participants. However, among those participants selected based on race, Hispanics and Blacks demonstrated lower performance quantity but higher performance quality (i.e., they were slower but more accurate) than Whites and Asians. This can be explained by a prevention focus that was likely triggered by the negative stereotype among Hispanic and Black individuals. A follow-up study suggested that minorities selected based on race had more prevention concerns than non-minorities, while this was not the case when they were selected based on merit. These results, therefore, qualify the generalizability of Heilman and colleagues' results, which pointed at the detrimental effects of preferential selection. We conclude that their findings, using communication and leadership tasks, are not entirely consistent with results when using cognitively related tasks.

These findings are in line with several papers showing the importance of regulatory focus in understanding stereotype threat outcomes. For example, Seibt and Fröster (2004) found that whereas positive in-group stereotypes led to more creative performance, negative stereotypes led to better analytical performance (studies 4 and 5). Using two tasks (a word-selection task and a task in which participants had to connect numbered dots), these authors also found that participants presented with a negative stereotype demonstrated slower but more accurate performance than a control group, whereas participants presented with a positive stereotype demonstrated faster but less accurate performance (studies 2 and 3). Additionally, Chalabaev et al. (2014) studied girls who were told to avoid errors in a soccer task (dribbling a ball through a slalom course). They found that those who faced a negative stereotype performed better (made fewer mistakes) than those who did not face a negative stereotype (see also Chalabaev et al. 2012). Finally, a number of researchers have found that older adults have improved working memory performance (avoiding false recognitions and false memories and less intrusion rates during free-recall tests) under stereotype threat (Barber 2017; Barber and Mather 2013a, b; Barber et al. 2015; Wong and Gallo 2016).

Our results have implications at two levels. From a theory standpoint, the regulatory focus perspective can account for our findings as well as previous findings in the preferential selection literature. Heilman and her colleagues (e.g., Heilman et al. 1987) used leadership and communication tasks, in which a promotion focus is important to perform well (Carney et al. 2010; Cuddy et al. 2015). Brown et al. (2000, study 1) conducted a study in which the task consisted of solving questions similar to those included in the Analytical Reasoning section of the GRE. They found that women who were preferentially selected demonstrated lower performance than those in a control condition. However, it is likely that a promotion focus would positively affect an outcome such as number of problems solved (Grimm et al. 2009). Indeed, Brown et al. (2000) concluded that the “pattern of performance differences appears to be due to a decrement in the number of problems answered rather than to a decrement in problem-solving accuracy” (p. 740; see also Steele and Aronson 1995). Turner and Pratkanis (1993; see also Nacoste 1989) found that women who were preferentially selected performed worse than those who were selected based on merit. However, the outcome was performance in a brainstorming task, in which the specific criterion was the number of uses given to objects such as a soda can. Again, a promotion focus tends to favor these outcomes (Fröster et al. 2003). Thus, overall, prior results are consistent with the regulatory focus perspective.

An interesting and related implication has to do with the speed-accuracy tradeoff (Heitz 2014).<sup>8</sup> Recent research in cognitive ability tests suggests that managing such trade-offs are important for standardized test performance (Ackerman and Ellingsen 2016). Our results suggest that, under stereotype threat (i.e., minorities selected based on race), minorities tend to be slower but more accurate. This suggests that in testing settings minorities may put too much emphasis on

<sup>8</sup> We thank an anonymous reviewer for suggesting this idea.

avoiding mistakes (accuracy or performance quality) and not enough on answering many items (speed or performance quantity). Relatedly, time constraints in cognitive ability tests may be an important issue to examine as well. Indeed, there is some evidence that time constraints may have a negative impact on women and minorities, regardless of actual ability. For example, whereas gender differences exist in some spatial tasks such as mentally rotating 3D objects (Voyer et al. 1995), time limit tends to increase these differences (Voyer 2011). Similar gender differences under time constraint have been found in other domains (De Paola and Gioia 2016). However, we are not aware of such negative impact across different ethnicities. Nor are we aware of racial differences in the speed-accuracy tradeoff. This is an interesting area for future research.

From a practical point of view, our research may help identify conditions in which negative stereotypes unequivocally have negative effects on performance from others in which they do not. For example, if assessment centers use in-basket tasks and the criterion is quantity (e.g., number of emails answered), minorities who feel threatened by their ethnicity (e.g., because they were told that the in-basket task is diagnostic of their abilities; Steele and Aronson 1995) may demonstrate poor performance. However, if the outcome is quality (e.g., fewer mistakes in writing emails), stereotype threat is less likely to be an issue. Our research shows, therefore, that the use of different outcomes qualifies the harmful effects of negative stereotypes, by showing that individual reactions to negative stereotypes demonstrate a mix of behavioral tendencies that results in better or worse performance, depending on how performance is measured. Relatedly, one could argue that affirmative action plans (AAPs) may not be as harmful to minorities' performance as previously thought. Thus, our research suggests that we should qualify some of Heilman and colleagues' findings regarding task performance.

### Limitations and Conclusion

A limitation of the present investigation—and this line of research as a whole—is that the type of affirmative action policies used in this line of research (i.e., preferential selection) is blatant and, in most cases, illegal (Evans 2003). The law generally prohibits stronger forms of preference that include selection of unequal or unqualified candidates. Relatedly, there are issues associated with the external validity of this experimental method (Taylor 1994; Crosby et al. 2006). It could very well be that the race-based selection setting created here had a short-term effect on task performance that may vanish if participants perform more trials. In addition, this paradigm may be a long way from what actually happens in organizations; it would be very rare that a person is told the reasons they were hired by a manager. As the editorial team suggested, it is time to study the effects of preferential using other

methodologies in order to better understand how applicable these findings are to real life. A second limitation is that we did not include a control condition. In general, studies from this literature do not include a control condition (although there are exceptions; Brown et al. 2000). However, had we included a control condition, we may have been able to further determine whether each of the manipulations (merit-based vs. race-based) had opposing effects; or whether just the preferential selection manipulation had a strong effect on performance.

This investigation examined whether results from laboratory experiments on preferential selection found by Heilman and colleagues (e.g., Heilman et al., 1986) generalized to cognitively oriented tasks. In particular, we tested hypotheses of the effects of race-based preferential selection on quality and quantity of individual performance. Despite the above limitations, this study adds to scholarly understanding of the dynamics of stereotype threat, and that by using different outcomes, it is possible to qualify and specify the potentially harmful effects of stereotype threat. Given the increasingly diverse workforce in the USA, future research should test the potential effects on job performance in actual organizational settings.

### References

- Ackerman, P. L., & Ellingsen, V. J. (2016). Speed and accuracy indicators of test performance under different instructional conditions: Intelligence correlates. *Intelligence*, 56(Supplement C), 1–9. <https://doi.org/10.1016/j.intell.2016.02.004>.
- American Psychological Association. (2001). *The publication manual of the American Psychological Association* (5th ed.). Washington, D.C.: American Psychological Association.
- Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, 116, 261–292.
- Barber, S. J. (2017). An examination of age-based stereotype threat about cognitive decline: Implications for stereotype-threat research and theory development. *Perspectives on Psychological Science*, 12(1), 62–90. <https://doi.org/10.1177/1745691616656345>.
- Barber, S. J., & Mather, M. (2013a). Stereotype threat can reduce older adults' memory errors. *The Quarterly Journal of Experimental Psychology*, 66(10), 1888–1895. <https://doi.org/10.1080/17470218.2013.840656>.
- Barber, S. J., & Mather, M. (2013b). Stereotype threat can both enhance and impair older adults' memory. *Psychological Science*, 24(12), 2522–2529. <https://doi.org/10.1177/0956797613497023>.
- Barber, S. J., Mather, M., & Gatz, M. (2015). How stereotype threat affects healthy older adults' performance on clinical assessments of cognitive decline: The key role of regulatory fit. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, gbv009. <https://doi.org/10.1093/geronb/gbv009>.
- Beilock, S. L. (2008). Math performance in stressful situations. *Current Directions in Psychological Science*, 17(5), 339–343. <https://doi.org/10.1111/j.1467-8721.2008.00602.x>.
- Berry, C. M., Cullen, M. J., & Meyer, J. M. (2014). Racial/ethnic subgroup differences in cognitive ability test range restriction:

- Implications for differential validity. *Journal of Applied Psychology*, 99(1), 21–37. <https://doi.org/10.1037/a0034376>.
- Betz, N. E., & Hackett, G. (1981). The relationship of career-related self-efficacy expectations to perceived career options in college women and men. *Journal of Counseling Psychology*, 28(5), 399–410. <https://doi.org/10.1037/0022-0167.28.5.399>.
- Brockner, J., & Higgins, E. T. (2001). Regulatory focus theory: Implications for the study of emotions at work. *Organizational Behavior and Human Decision Processes*, 86, 35–66.
- Brown, R. P., Charnsangavej, T., Keough, K. A., Newman, M. L., & Rentfrow, P. J. (2000). Putting the “affirm” into affirmative action: Preferential selection and academic performance. *Journal of Personality and Social Psychology*, 79, 736–747.
- Brown, R. P., & Lee, M. N. (2005). Stigma consciousness and the race gap in college academic achievement. *Self and Identity*, 4(2), 149–157. <https://doi.org/10.1080/1357650044000227>.
- Caplan, L. (1995). Why affirmative action is divisive, difficult—and necessary. *The Washington Post*, p. C03.
- Carney, D. R., Cuddy, A. J. C., & Yap, A. J. (2010). Power posing brief nonverbal displays affect neuroendocrine levels and risk tolerance. *Psychological Science*, 21(10), 1363–1368. <https://doi.org/10.1177/0956797610383437>.
- Chalabaev, A., Dematte, E., Sarrazin, P., & Fontayne, P. (2014). Creating regulatory fit under stereotype threat: Effects on performance and self-determination among junior high school students. *Revue Internationale de Psychologie Sociale*, 27, 119–132.
- Chalabaev, A., Major, B., Sarrazin, P., & Cury, F. (2012). When avoiding failure improves performance: stereotype threat and the impact of performance goals. *Motivation and Emotion*, 36(2), 130–142.
- Craig, J. G. (1995). *Fairly inconsistent: The meaning of “fairness” is usually in the eye of the begrudged* (p. D3). Pittsburgh: Pittsburgh Post Gazette.
- Crosby, F. J., Iyer, A., Clayton, S., & Downing, R. A. (2003). Affirmative action: Psychological data and the policy debates. *American Psychologist*, 58(2), 93–115. <https://doi.org/10.1037/0003-066X.58.2.93>.
- Crosby, F. J., Iyer, A., & Sincharoen, S. (2006). Understanding affirmative action. *Annual Review of Psychology*, 57, 585–611. <https://doi.org/10.1146/annurev.psych.57.102904.190029>.
- Crowe, E., & Higgins, E. T. (1997). Regulatory focus and strategic inclinations: Promotion and prevention in decision-making. *Organizational Behavior and Human Decision Processes*, 69(2), 117–132. <https://doi.org/10.1006/obhd.1996.2675>.
- Cuddy, A. J. C., Wilmoth, C. A., Yap, A. J., & Carney, D. R. (2015). Preparatory power posing affects nonverbal presence and job interview performance. *Journal of Applied Psychology*, 100(4), 1286–1295. <https://doi.org/10.1037/a0038543>.
- De Paola, M., & Gioia, F. (2016). Who performs better under time pressure? Results from a field experiment. *Journal of Economic Psychology*, 53(Supplement C), 37–53. <https://doi.org/10.1016/j.joep.2015.12.002>.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56(1), 5–18. <https://doi.org/10.1037/0022-3514.56.1.5>.
- Doverspike, D., Taylor, M. A., & Arthur Jr., W. (2000). *Affirmative action: A psychological perspective*. Hauppauge, NY, US: Nova Science Publishers.
- Devos, T., & Torres, J. A. C. (2007). Implicit identification with academic achievement among Latino college students: The role of ethnic identity and significant others. *Basic and Applied Social Psychology*, 29, 293–310.
- Eagly, A. H., & Carli, L. L. (2004). Women and men as leaders. In J. Antonakis, A. T. Cianciolo, & R. J. Sternberg (Eds.), *The nature of leadership* (pp. 279–301). Thousand Oaks: Sage Publications, Inc.
- Fennis, B., & Stel, M. (2011). The pantomime of persuasion: Fit between sales-agent nonverbal communication and influence strategies. *Advances in Consumer Research*, 39, 441–442.
- Fong, E. A., & Tosi Jr., H. L. (2007). Effort, performance, and conscientiousness: An agency theory perspective. *Journal of Management*, 33, 161–179.
- Forbes, C. E., & Schmader, T. (2010). Retraining attitudes and stereotypes to affect motivation and cognitive capacity under stereotype threat. *Journal of Personality and Social Psychology*, 99, 740–754.
- Förster, J., Higgins, E. T., & Bianco, A. T. (2003). Speed/accuracy decisions in task performance: Built-in trade-off or separate strategic concerns? *Organizational Behavior and Human Decision Processes*, 90(1), 148–164. [https://doi.org/10.1016/S0749-5978\(02\)00509-5](https://doi.org/10.1016/S0749-5978(02)00509-5).
- Friedman, R. S., & Förster, J. (2001). The effects of promotion and prevention cues on creativity. *Journal of Personality and Social Psychology*, 81, 1001–1013.
- Grimm, L. R., Markman, A. B., Todd, W., & Baldwin, G. C. (2009). Stereotype threat reinterpreted as a regulatory mismatch. *Journal of Personality and Social Psychology*, 96(2), 288–304. <https://doi.org/10.1037/a0013463>.
- Hays, W. (1994). *Statistics* (5th ed.). Fort Worth: Wadsworth Publishing.
- Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, 32, 113–135. <https://doi.org/10.1016/j.riob.2012.11.003>.
- Heilman, M. E., & Alcott, V. B. (2001). What I think you think of me: Women’s reactions to being viewed as beneficiaries of affirmative action. *Journal of Applied Psychology*, 86, 574–582.
- Heilman, M. E., & Blader, S. L. (2001). Assuming preferential selection when the admissions policy is unknown: The effect of gender rarity. *Journal of Applied Psychology*, 86, 188–193.
- Heilman, M. E., Block, C. J., & Lucas, J. A. (1992). Presumed incompetent? Stigmatization and affirmative action efforts. *Journal of Applied Psychology*, 77, 536–544.
- Heilman, M. E., Kaplow, S. R., Amato, M. A. G., & Stathatos, P. (1993). When similarity is a liability: Effects of sex-based preferential selection on reactions to like-sex and different-sex others. *Journal of Applied Psychology*, 78, 917–927.
- Heilman, M. E., Lucas, J. A., & Kaplow, S. R. (1990). Self-derogating consequences of sex-based preferential selection: The moderating role of initial self-confidence. *Organizational Behavior and Human Decision Processes*, 46, 202–216.
- Heilman, M. E., Manzi, F., & Braun, S. (2015). Presumed incompetent: perceived lack of fit and gender bias in recruitment and selection. In A. M. Broadbridge & S. L. Fielden (Eds.), *Handbook of gendered careers in management: Getting in, getting on, getting out* (Vol. 90, pp. 90–104). Cheltenham, UK: Elgar.
- Heilman, M. E., McCullough, W. F., & Gilbert, D. (1996a). The other side of affirmative action: Reactions of non-beneficiaries to sex-based preferential selection. *Journal of Applied Psychology*, 81, 347–357.
- Heilman, M. E., Rivero, J. C., & Brett, J. F. (1991). Skirting the competence issue: Effects of sex-based preferential selection on task choices of men and women. *Journal of Applied Psychology*, 76, 99–105.
- Heilman, M. E., Simon, M. C., & Repper, D. P. (1987). Intentionally favored, unintentionally harmed? The impact of sex-based preferential selection on self-perceptions and self-evaluations. *Journal of Applied Psychology*, 72, 62–68.
- Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience*, 8. <https://doi.org/10.3389/fnins.2014.00150>.
- Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, 52(12), 1280–1300. <https://doi.org/10.1037/0003-066X.52.12.1280>.
- Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 1–46). Cambridge: Academic Press.

- Inzlicht, M., Tullett, A. M., Legault, L., & Kang, S. K. (2011). Lingering effects: stereotype threat hurts more than you think. *Social Issues and Policy Review*, 5(1), 227–256. <https://doi.org/10.1111/j.1751-2409.2011.01031.x>.
- Johnson, R. E. (2006). Uncovering the motivational processes underlying justice: The implicit cognitive, affective, and conative effects of experiencing (un)fairness. Doctoral dissertation, University of Akron, Akron, OH.
- Johnson, R. E., & Steinman, L. (2009). The use of implicit measures for organisational research: An empirical example. *Canadian Journal of Behavioural Science*, 41, 202–212.
- Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences*. Pacific Grove: Brooks/Cole.
- Lenny, E. (1977). Women's self-confidence in achievement settings. *Psychological Bulletin*, 84, 1–13.
- Leslie, L. M., Mayer, D. M., & Kravitz, D. A. (2014). The stigma of affirmative action: A stereotyping-based theory and meta-analytic test of the consequences for performance. *Academy of Management Journal*, 57(4), 964–989. <https://doi.org/10.5465/amj.2011.0940>.
- Levy, H. O. (2016). Supreme court has spoken on affirmative action. Now, colleges should boost income diversity. Retrieved June 25, 2016, from <http://www.foxnews.com/opinion/2016/06/24/supreme-court-has-spoken-on-affirmative-action-now-colleges-should-boost-income-diversity.html>
- Li, A., Evans, J., Christian, M. S., Gilliland, S. W., Kausel, E. E., & Stein, J. H. (2011). The effects of managerial regulatory fit priming on reactions to explanations. *Organizational Behavior and Human Decision Processes*, 115(2), 268–282. <https://doi.org/10.1016/j.obhdp.2011.01.003>.
- Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences (Vol. xiii)*. Palo Alto: Stanford University Press.
- McFall, S., Jamieson, J. P., & Harkins, S. (2009). Testing the generalizability of the mere effort account of the evaluation–performance relationship. *Journal of Personality and Social Psychology*, 96, 135–154.
- Mazerolle, M., Régner, I., Morisset, P., Rigalleau, F., & Huguet, P. (2012). Stereotype threat strengthens automatic recall and undermines controlled processes in older adults. *Psychological Science*, 956797612437607, 723–727. <https://doi.org/10.1177/0956797612437607>.
- Nacoste, R. W. (1989). Affirmative action and self-evaluation. In F. A. Blanchard & F. J. Crosby (Eds.), *Affirmative action in perspective* (pp. 103–109). New York: Springer-Verlag.
- O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin*, 29(6), 782–789.
- Oswald, D. L., & Harvey, R. D. (2000). Hostile environments, stereotype threat, and math performance among undergraduate women. *Current Psychology*, 19(4), 338–356. <https://doi.org/10.1007/s12144-000-1025-5>.
- Richeson, J. A., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science*, 14(3), 287–290. <https://doi.org/10.1111/1467-9280.03437>.
- Roberson, L., & Kulik, C. T. (2007). Stereotype threat at work. *The Academy of Management Perspectives*, 21(2), 24–40. <https://doi.org/10.5465/AMP.2007.25356510>.
- Roth, P. L., Be Vier, C. A., Bobko, P., Switzer III, F. S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, 54, 297–330.
- Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, 11, 235–294.
- Rydell, R. J., McConnell, A. R., & Beilock, S. L. (2009). Multiple social identities and stereotype threat: Imbalance, accessibility, and working memory. *Journal of Personality and Social Psychology*, 96(5), 949–966. <https://doi.org/10.1037/a0014846>.
- Rydell, R. J., Van Loo, K. J., & Boucher, K. L. (2014). Which executive functions are impaired by stereotype threat and account for different threat-related outcomes? *Personality and Social Psychology Bulletin*, 40, 377–390.
- Schmader, T., & Johns, M. (2003). Convergent evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, 85, 440–452.
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, 115, 336–356.
- Seibt, B., & Förster, J. (2004). Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci. *Journal of Personality and Social Psychology*, 87, 38–56.
- Shih, M., Pittinsky, T., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, 10, 80–83.
- Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*, 67(1), 415–437. <https://doi.org/10.1146/annurev-psych-073115-103235>.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual performance of African-Americans. *Journal of Personality and Social Psychology*, 69, 797–811.
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–440). New York: Academic Press.
- Stewart, M. M., & Shapiro, D. L. (2000). Selection based on merit versus demography: Implications across race and gender lines. *Journal of Applied Psychology*, 85, 219–231.
- Taylor, M. C. (1994). Impact of affirmative action on beneficiary groups: Evidence from the 1990 General Social Survey. *Basic and Applied Social Psychology*, 15(1–2), 143–178. [https://doi.org/10.1207/s15324834baspl501&2\\_7](https://doi.org/10.1207/s15324834baspl501&2_7).
- Totenberg, N. (2016, June 24). Supreme court upholds affirmative action in college admissions. *NPR.org*. Retrieved June 25, 2016, from <http://www.npr.org/2016/06/23/483275212/supreme-court-upholds-affirmative-action-in-college-admissions>
- Turner, M. E., & Pratkanis, A. R. (1993). Effects of preferential and meritorious selection on performance: An examination of intuitive and self-handicapping perspectives. *Personality and Social Psychology Bulletin*, 19(1), 47–58. <https://doi.org/10.1177/0146167293191006>.
- Voyer, D. (2011). Time limits and gender differences on paper-and-pencil tests of mental rotation: A meta-analysis. *Psychonomic Bulletin & Review*, 18(2), 267–277. <https://doi.org/10.3758/s13423-010-0042-0>.
- Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, 117(2), 250–270.
- Wade, R. (1995). Affirmative action: A small patch on a festering problem. *Minneapolis Star Tribune*, p. 4.
- Wallace, C., & Chen, G. (2006). A multilevel integration of personality, climate, self-regulation, and performance. *Personnel Psychology*, 59(3), 529–557. <https://doi.org/10.1111/j.1744-6570.2006.00046.x>.
- Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, 39, 456–467.
- Walton, G. M., & Spencer, S. J. (2009). Latent ability grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science*, 20(9), 1132–1139. <https://doi.org/10.1111/j.1467-9280.2009.02417.x>.
- Weyant, J. M. (2005). Implicit stereotyping of Hispanics: Development and validity of a Hispanic version of the implicit association test. *Hispanic Journal of Behavioral Sciences*, 27(3), 355–363. <https://doi.org/10.1177/0739986305276747>.
- Will, G. F. (2001). Affirmative action out of control. *The Washington Post*, p. A19.
- Wong, J. T., & Gallo, D. A. (2016). Stereotype threat reduces false recognition when older adults are forewarned. *Memory*, 24(5), 650–658. <https://doi.org/10.1080/09658211.2015.1036885>.