



## FlashReport

## Consider the situation: Reducing automatic stereotyping through Situational Attribution Training

Tracie L. Stewart<sup>a,\*</sup>, Ioana M. Latu<sup>a</sup>, Kerry Kawakami<sup>b</sup>, Ashley C. Myers<sup>a</sup>

<sup>a</sup> Department of Psychology, Georgia State University, United States

<sup>b</sup> Department of Psychology, York University, Toronto, Canada

## ARTICLE INFO

## Article history:

Received 23 July 2009

Revised 30 August 2009

Available online 13 September 2009

## Keywords:

Stereotyping

Bias reduction

Intergroup relations

Implicit associations

## ABSTRACT

The present research investigated the effectiveness of a new technique for reducing automatic biases rooted in attribution theory – the Situational Attribution Training Technique. The goal of this strategy extends previous work by targeting the fundamental attributional pillars underlying automatic stereotyping. We aimed to circumvent the well-documented tendency for individuals to be overly reliant on dispositional attributions when perceiving negative stereotype-consistent behaviors performed by outgroup members. By teaching participants to consider situational attributions for such behaviors, we expected a reduction in outgroup stereotyping. Specifically, White participants were trained extensively to choose situational over dispositional explanations for negative stereotype-consistent behaviors performed by Black men. Across two experiments, participants who completed Situational Attribution Training demonstrated reduced automatic racial stereotyping on a person categorization task, relative to control participants who exhibited substantial automatic stereotyping. The implications of these findings for the nature and reduction of intergroup biases are discussed.

© 2009 Elsevier Inc. All rights reserved.

## Introduction

*"We have this insatiable hunger for explanations. [...] This basic need has an important bearing upon group relations. For one thing, we tend to regard causation as something people are responsible for. [...] This quirk, unless it is strenuously disciplined, predisposes us to prejudice."* (Allport, 1954; p. 170).

In 2005, Harvard President Lawrence Summers publicly suggested that innate gender differences were probably the primary reason for women's underrepresentation in math and science domains. His remarks caused a stir in academic and non-academic communities and are at odds with considerable research suggesting that women's underperformance in math and science is linked to situational factors (Krendl, Richeson, Kelley, & Heatherton, 2008; Steele, 1997). Summers' comment reminds us that people tend to underestimate situational constraints on many outgroup member behaviors. A primary concern for intergroup relations is the well-documented tendency for individuals to attribute, in particular, the *negative* behaviors of outgroup members to dispositional factors, especially if the behaviors are stereotype-consistent – a phenomenon coined the *ultimate attribution error*

(UAE; Pettigrew, 1979). In contrast, according to the UAE, positive behaviors performed by outgroup members are generally attributed to situational factors (Allport, 1954; Fiske, 2005).

We contend that the UAE is one of the pillars on which stereotyping stands. When perceivers attribute negative stereotype-consistent behaviors of outgroup members to internal, stable factors ("He could not get a job because he is incompetent"), while underestimating situational constraints ("Jobs are scarce these days"), they are likely to perpetuate outgroup stereotypes. For example, White Americans' stereotype of Blacks as "aggressive" may be perpetuated by attributing the same aggressive shove to dispositional factors for Black actors and to situational factors for White actors, an attributional pattern demonstrated in previous research (Duncan, 1976). Likewise, attributing expressions of anger to dispositional factors for women and to situational factors for men may maintain the "emotional" stereotype of women (Brescoll & Uhlmann, 2008).

Despite the potentially fundamental role of the UAE in processes related to stereotyping, researchers have not systematically targeted dispositional attributions as a means of reducing intergroup bias. In the present research, we introduced a new technique directed at decreasing UAE tendencies. We predicted that extensively training people to make situational, rather than dispositional, explanations for outgroup members' negative stereotype-consistent behaviors would reduce automatic stereotyping. This situational attribution training represents the first bias reduction

\* Corresponding author. Address: Department of Psychology, Georgia State University, Atlanta, GA 30302-5010, United States.

E-mail address: [stewart@gsu.edu](mailto:stewart@gsu.edu) (T.L. Stewart).

technique rooted in attribution theory. Unlike other successful training techniques that target the stereotyping process (e.g., Gawronski, Deutsch, Mbirikou, Seibt, & Strack, 2008; Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000; Stewart & Payne, 2008), our paradigm targets a more fundamental attributional process underlying outgroup stereotypes. As such, this research provides some of the first evidence for the causal relationship between dispositional attributions and stereotyping as well as providing a new and possibly broader method to reduce intergroup biases.

### *Undoing the ultimate attribution error*

Given that attributions may have important consequences for intergroup biases, what happens if we strengthen situational attributions for negative behaviors performed by outgroup members? Findings from “perspective taking” studies indirectly support the hypothesis that situational attributions precipitate lower bias. Compared to controls, participants who took the perspective of a stigmatized person, by imagining what that person felt, showed reduced explicit negative attitudes towards the entire stigmatized group (Batson et al., 1997; Dovidio et al., 2004), as well as reduced stereotype activation (Galinsky & Moskowitz, 2000). Vescio, Sechrist, and Paolucci (2003) found that the efficacy of a perspective taking strategy to reduce explicit prejudice was mediated by increased situational attributions about the target – an attribution type more often used to explain ingroup than outgroup behaviors.

Thus, existing research indirectly suggests that situational attributions for outgroup members’ negative behaviors may impact intergroup biases. However, no studies have directly investigated the causal relationship between situational attributions and automatic stereotype activation. Our research goal was therefore to explore the possibility that training participants to make situational attributions for negative stereotypic behaviors, thereby targeting the UAE, would decrease automatic stereotyping.

### *The present research*

Pretesting confirmed that situational attribution training increased spontaneous activation of situational explanations for stereotype-consistent behaviors.<sup>1</sup> We therefore conducted two experiments to test our prediction that training would also decrease automatic activation of negative stereotypes. Consistent with connectionist models of activation (e.g., Monroe & Read, 2008; van Overwalle & Labiouse, 2004), we predicted that effects of training would extend beyond the traits targeted in training to other unrelated negative stereotypic traits. According to these models, negative stereotypic traits are units connected in a network; once one unit is activated (e.g., a particular negative stereotypic trait), activation can spread to other relevant units (other negative stereotypic traits). Consequently, if training reduces the activation of specific negative stereotypic traits, activation of other negative stereotypic traits would also be expected to decrease. We expected positive stereo-

typic trait activation to be less impacted by training given that the UAE, which our paradigm targets, specifically elicits attributions for outgroup behaviors that maintain negative stereotypes.

The two experiments differed only in comparison group. In Experiment 1, findings for the Situational Attribution Training Condition were compared to a No Training condition. Experiment 2 employed a control condition more comparable to the experimental condition. We expected to obtain the same findings across control conditions. For conciseness, and because of the similarity in the experiments’ methodology and findings, we present their methods and results together.

## **Method**

### *Participants and design*

Seventy-two White undergraduates (50 women) participated in one of two experiments (32 participants in Experiment 1) as a means to fulfill an introductory psychology course requirement. Participants in both experiments were randomly assigned to the Situational Attribution Training Condition or a control condition.

### *Procedure*

#### *Phase 1: training*

Participants assigned to the Situational Attribution Training Condition were told the study investigated how people explain others’ behaviors. The experimenter exemplified the difference between dispositional and situational explanations. Participants were also informed that they were randomly assigned to a condition in which they would make situational attributions for negative behaviors performed by Black men.

After six practice trials with feedback, participants began the training, which was composed of 480 trials divided into six blocks of 80 trials. Each trial began with presentation of a photograph of a Black man, paired with the label “African American” and a sentence describing a Black-stereotypic behavior. Forty behaviors were presented twice per block – four behaviors related to each negative stereotypic trait. The pretested traits were loud, criminal, unintelligent, unreliable, irresponsible, violent, dishonest, dangerous, lazy, and promiscuous. Following a 3000 ms delay, the words “I Choose:” appeared mid-screen, below the behavior description. Two possible explanations of the behavior, one situational and one dispositional, appeared, respectively, on the bottom left- and right-hand side of the screen. The location of the explanations was counterbalanced such that the situational explanation appeared on the right for half of the trials and the left for the remaining half. Participants’ task was to choose the situational explanation of the two by pressing the keyboard key associated with the left- or right-hand side of the screen (see Fig. 1a).

Experiment 1 control participants did not complete any training and proceeded directly to Phase 2 of the study (No Training Control). Experiment 2 control participants were presented with the same photographs and behavioral sentences presented to Situational Attribution Training participants. However, instead of making situational attributions for behaviors, participants counted the number of nouns (240 trials) or verbs in the behavioral sentences (Grammatical Training Control). They made dichotomous decisions using keys associated with the left- or right-hand side of the screen (e.g., choosing “2 or under 2 nouns” or “over 2 nouns;” see Fig. 1b).

#### *Phase 2: the person categorization task*

Next, all participants completed the person categorization task (Banaji & Hardin, 1996) as a measure of automatic stereotype activation. This task was described as a separate experiment conducted by a different researcher. The experimenter explained

<sup>1</sup> Eighteen (White) Situational Attribution Training and Grammar Training Control participants completed a Probe Recognition Task (Ham & Vonk, 2003). On each trial, they briefly (3000 ms) viewed a photo of a Black man paired with one of 20 pretested behavioral sentences indicative of negative Black-stereotypic traits unseen in training. Each display was then replaced by a word probe. Participants’ task was to quickly and accurately indicate whether the probe had appeared in the sentence by pressing one of two keys labeled “yes” and “no”. Sentences randomly appeared six times, followed by a different probe each time (half seen/half not seen). Of interest were words absent in the behavioral sentence but associated (situational probe) or unassociated (control probe) with a situational explanation for that behavior, according to prior pretesting. Following training, participants were slower to correctly reject situational, versus control, probes,  $F(1, 8) = 6.92$ ,  $p < .04$ ,  $\eta^2 = .46$ , suggesting response interference caused by heightened activation of situational inferences. Control participants showed no such difference,  $F(1, 8) = 1.07$ ,  $p = .33$ ,  $\eta^2 = .12$ .

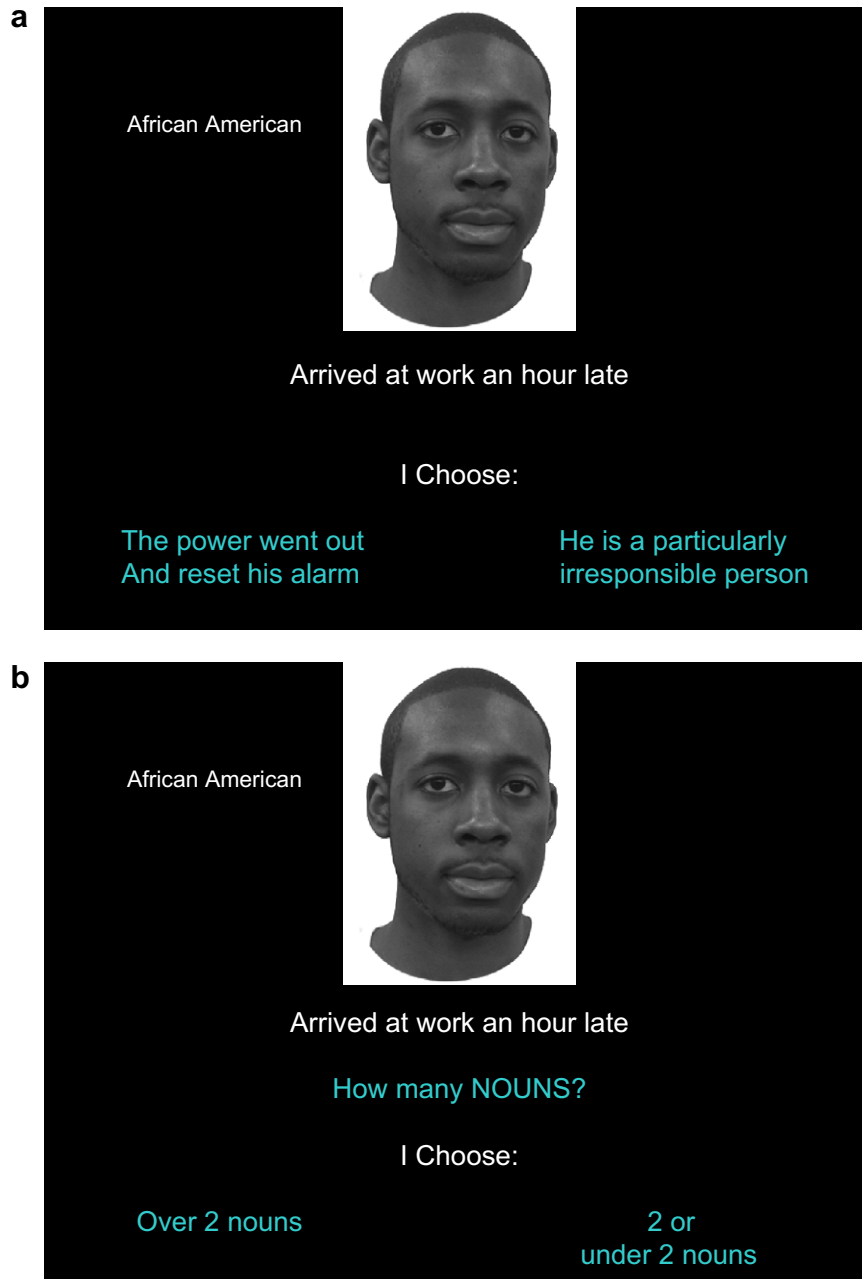


Fig. 1. Example of the screen display on a typical trial in the Situational Attribution Training (a) and Grammatical Training Control (b) conditions.

that the goal was to study how people categorize photographs of others in different groups and that they had been randomly assigned to a condition in which an unrelated distracter word was presented before each photo. For each trial, participants categorized photographs of Black and White men by race after being exposed to a trait prime for 250 ms. Traits included eight positive and eight negative traits unrelated to the traits used in Phase 1 that were determined by pretests to be stereotypic of Blacks (e.g., religious, poor), and 16 positive and 16 negative traits unrelated to Black stereotypes (e.g., elegant, naive). Additionally, eight negative Black-stereotypic traits targeted in Phase 1 were randomly selected for presentation in Phase 2.

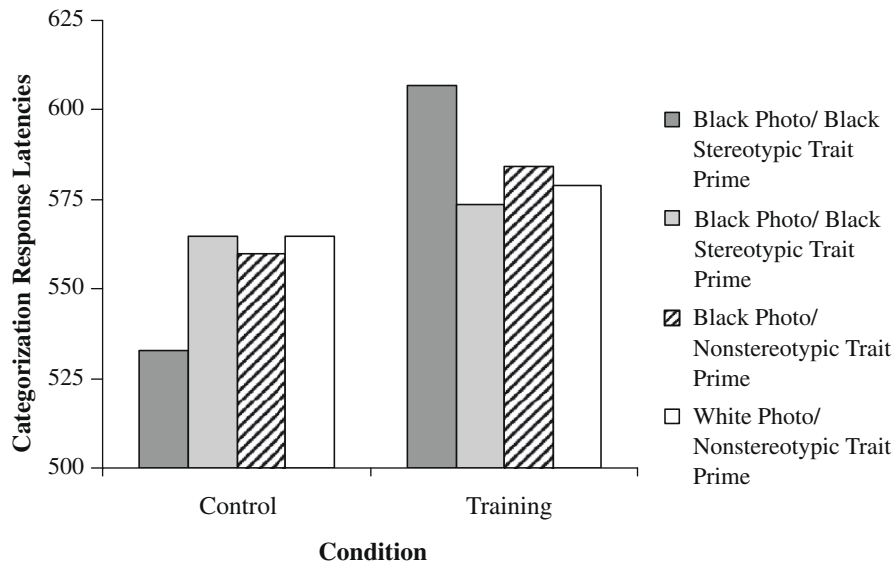
Participants completed two blocks of 56 trials. Within each block, half of the traits of each category (negative and positive Black-stereotypic; negative and positive filler; training) were fol-

lowed by a Black photo and the other half by a White photo. The pairings were counterbalanced such that traits paired with Black photos in one block were paired with White photos in the other block and vice versa. Response latencies for each trial were recorded, with faster responses indicative of greater implicit association between the photo and trait.

## Results

### *Preliminary analyses*

Our dependent measure was response latencies across categorization task trials. To allow comparisons unconfounded by differential prior exposure in the study (Kawakami et al., 2000), trait prime analyses were restricted to traits not used in the training. Response



**Fig. 2.** Control and Situational Attribution Training condition participants' mean response latencies for categorizing photographs of Blacks and Whites by race immediately after negative Black-stereotypic or nonstereotypic trait priming. Note: Response latencies are reported in milliseconds.

latencies were log-transformed to control for outliers; however, we report nontransformed means in the text. In initial analyses, data for the two experiments were merged, and an experiment factor was created. Given nonsignificant effects for the experiment factor, the primary analyses excluded this factor. Data for the control conditions, which did not produce different effects, were combined.<sup>2</sup>

#### Does situational attribution training reduce automatic stereotyping?

To examine which factors impacted the relative categorization speed of White and Black photos, response latencies for categorizing targets as Black or White following trait primes were analyzed in a 2 (Condition: Situational Attribution Training vs. Control)  $\times$  2 (Target Race)  $\times$  2 (Trait Type: Black-stereotypic vs. nonstereotypic)  $\times$  2 (Trait Valence) ANOVA, with repeated measures on the last three factors. Two interactions involving target race reached significance. A Condition  $\times$  Target Race interaction indicated that Black, but not White, photos were categorized particularly quickly in the control conditions,  $F(1, 69) = 8.71, p < .004, \eta^2 = .11$ . This interaction was qualified by the predicted Condition  $\times$  Target Race  $\times$  Trait Type  $\times$  Trait Valence interaction,  $F(1, 69) = 8.85, p < .004, \eta^2 = .11$ .<sup>3</sup>

To explicate this interaction, we next analyzed positive and negative trait primes separately. No significant interactions of target race with condition or trait type were observed for positive trait primes. However, analysis of negative trait primes yielded a significant Condition  $\times$  Target Race interaction,  $F(1, 69) = 7.24, p < .009, \eta^2 = .10$ , qualified by a significant Condition  $\times$  Target Race  $\times$  Trait Type interaction,  $F(1, 69) = 6.21, p < .02, \eta^2 = .08$ . Response latencies for categorization of Black and White photos following stereotypic, but not nonstereotypic, negative trait primes differed as a function of condition. Control participants categorized Black photos significantly faster than White photos following presentation of a negative Black-stereotypic trait prime,  $F(1, 26) = 11.58, p < .002, \eta^2 = .31$ . In contrast, Situational Attribution Training participants exhibited *no difference* in the speed with

which they categorized White and Black photos following negative Black-stereotypic trait primes,  $F(1, 43) = 2.29, p = .14, \eta^2 = .05$  (see Fig. 2).<sup>4</sup> These findings are consistent with our prediction that extensive practice in making situational attributions for negative Black-stereotypic behaviors would reduce implicit associations between negative Black stereotypes and Blacks.

#### Discussion

Stereotyping has been theoretically linked to the tendency to rely on dispositional explanations when explaining negative behaviors of outgroup members. The present research demonstrates that extensive training aimed at undermining this process reduces automatic stereotyping. White participants trained to “consider the situation” when judging negative stereotype-consistent behaviors of Black men showed reduced activation of negative Black stereotypes *not seen or implied* in training. The stereotype reduction benefits of situational attribution training therefore generalized beyond the training received. However, training had no impact on the relative association of Blacks and Whites with positive Black-stereotypic traits or positive or negative nonstereotypic traits. The specific effect of training was to decrease the targeted initial automatic association between Black individuals and negative stereotype-consistent traits.

The present findings are notable in light of considerable evidence that biased attributional processes are automatic, pervasive, and resistant to change (e.g., Geeraert & Yzerbyt, 2007). Even low-prejudiced White individuals have shown a greater tendency to automatically attribute behaviors consistent with Black stereotypes to dispositional factors for Black, versus White, actors (Stewart, Weeks, & Lupfer, 2003). Gilbert and Malone (1995) proposed that automatic dispositional attributions can be consciously corrected to take into account situational constraints on behaviors, but only if sufficient cognitive resources are present. Given that situational attribution training has been shown to successfully in-

<sup>2</sup> Participants more quickly categorized Black than White photos following negative Black-stereotypic trait primes in the No Training,  $F(1, 9) = 8.79, p < .02$ , and Grammatical Training control conditions,  $F(1, 16) = 4.31, p < .05$ . Degree of automatic negative stereotyping did not differ across control conditions,  $F(1, 25) < 1$ .

<sup>3</sup> In preliminary analyses, this finding did not differ across experiments, Experiment  $\times$  Condition  $\times$  Target Race  $\times$  Trait Type  $\times$  Trait Valence,  $F(1, 67) < 1$ .

<sup>4</sup> The same effects were observed in analyses incorporating negative stereotypic traits both seen and not seen in training: Black photos were categorized faster than White photos following these combined trait primes in the control conditions,  $F(1, 26) = 8.93, p < .006, \eta^2 = .26$ , but not the attribution training condition,  $F(1, 43) < 1$ .

crease the *automaticity* of situational inferences of negative behaviors performed by outgroup members, it holds promise to yield bias reduction effects that are particularly long-lasting and that endure in conditions involving low cognitive resources.

Additional research is needed to ascertain the extent to which reduced stereotype activation persists after attributional training and to examine whether factors such as heightened salience of egalitarian norms, increased compunction, or elevated social desirability concerns contributed to the observed effects of the training. Although further research is clearly useful to directly address limitations of the present research, the present studies suggest that targeting the attributional pillars on which stereotypes are built is a potentially important strategy to reduce racial bias and improve intergroup relations.

### Acknowledgments

This research was supported in part by a grant from the Russell Sage Foundation to the first author. We thank Nicole Shelton, Jack Dovidio, Bertram Gawronski, David Washburn, Dominic Parrott, and four anonymous reviewers for helpful comments on previous versions of this manuscript and Claire Lisco, Triman Kohli, Glenna Read, Cristina Parfene, and Patrick Dougall for research assistance.

### References

- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science*, 7, 136–141.
- Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., et al. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology*, 72, 105–118.
- Brescoll, V. L., & Uhlmann, E. L. (2008). Can an angry woman get ahead? Status conferral, gender, and expression of emotion in the workplace. *Psychological Science*, 19, 268–275.
- Dovidio, J. F., ten Vergert, M., Stewart, T. L., Gaertner, S. L., Johnson, J. D., Esses, V. M., et al. (2004). Perspective and prejudice: Antecedents and mediating mechanisms. *Personality and Social Psychology Bulletin*, 30, 1537–1549.
- Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of Blacks. *Journal of Personality and Social Psychology*, 34, 590–598.
- Fiske, S. T. (2005). Social cognition and the normality of prejudice. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice. Fifty years after Allport* (pp. 36–53). Malden, MA: Blackwell Publishing.
- Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and ingroup favoritism. *Journal of Personality and Social Psychology*, 78, 708–724.
- Gawronski, B., Deutsch, R., Mbirikou, S., Seibt, B., & Strack, F. (2008). When “Just Say No” is not enough: Affirmation versus negation training and the reduction of automatic stereotype activation. *Journal of Experimental Social Psychology*, 44, 370–377.
- Geeraert, N., & Yzerbyt, V. Y. (2007). How fatiguing is dispositional suppression: Disentangling the effects of procedural rebound and ego-depletion. *European Journal of Social Psychology*, 37, 216–230.
- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, 117, 21–38.
- Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, 78, 871–888.
- Krendl, A. C., Richeson, J. A., Kelley, W. M., & Heatherton, T. F. (2008). The negative consequences of threat: A functional magnetic resonance imaging investigation of the neural mechanisms underlying women’s underperformance in math. *Psychological Science*, 19, 168–175.
- Monroe, B. M., & Read, S. J. (2008). A general connectionist model of attitude structure and change: The ACS (Attitudes as Constraint Satisfaction) model. *Psychological Review*, 115, 733–759.
- Pettigrew, T. F. (1979). The ultimate attribution error: Extending Allport’s cognitive analysis of prejudice. *Personality and Social Psychology Bulletin*, 5, 461–476.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual ability and performance. *American Psychologist*, 52, 613–629.
- Stewart, B. D., & Payne, B. K. (2008). Bringing automatic stereotyping under control: Implementation intentions as efficient means of thought control. *Personality and Social Psychology Bulletin*, 34, 1332–1345.
- Stewart, T. L., Weeks, M., & Lupfer, M. B. (2003). Spontaneous stereotyping: A matter of prejudice? *Social Cognition*, 21, 263–298.
- Van Overwalle, F., & Baetens, C. (2004). A recurrent connectionist model of person impression formation. *Personality and Social Psychology Review*, 8, 28–61.
- Vescio, T. K., Sechrist, G. B., & Paolucci, M. P. (2003). Perspective taking and prejudice reduction: The mediational role of empathy arousal and situational attributions. *European Journal of Social Psychology*, 33, 455–472.