

Kicking the habit: Effects of nonstereotypic association training and correction processes on hiring decisions[☆]

Kerry Kawakami,^{a,*} John F. Dovidio,^b and Simone van Kamp^c

^a Department of Psychology, York University, 4700 Keele Street, Toronto, Canada M3J 1P3

^b Colgate University, USA

^c University of Nijmegen, The Netherlands

Received 10 July 2003; revised 11 May 2004

Available online 15 June 2004

Abstract

The primary aim of the present research was to examine the effect of training in associating nonstereotypic traits with men and women on hiring decisions. While previous findings demonstrate that training can reduce the uncontrolled activation of stereotypes, the present results show that training by itself may not reduce the more controlled application of stereotypes. Across both no training and training conditions, participants chose male over female candidates for a leadership function. However, extensive nonstereotypic training did reduce sex discrimination when the training phase was disassociated from the choice of candidate task or when the participant's cognitive capacity was limited. These findings provide further evidence for the effectiveness of training in decreasing intergroup biases and for the potential influence of mental correction in moderating the effects of nonstereotypic training.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Stereotyping; Prejudice; Automatic processes; Discrimination; Implicit stereotypes; Gender bias; Sexism

Although in the last 50 years there has been a substantial increase in the representation of women in the workplace, disparities in the status of woman relative to men persist (Kite et al., 2001). Stereotypes have long been identified as a fundamental factor contributing to this type of discrimination (Allport, 1954). Considerable evidence has revealed a significant relationship between stereotypes, which are characteristics associated with a particular social group, and discrimination, which is biased treatment of others based on their group membership (Dovidio, Brigham, Johnson, & Gaertner, 1996). For example, Rudman and Glick (1999) have shown that because power and leadership are qualities

associated with men and submissiveness and gentility are characteristics associated with women, men may be chosen more often than women for supervisory roles in employment contexts. Thus, one way of decreasing sex discrimination may be related to reducing one of the important contributors to bias, stereotyping (Deaux & LaFrance, 1998).

There are two basic processes involved in stereotyping (Bargh, 1999; Devine, 1989; Kawakami, Dion, & Dovidio, 1998). Whereas, stereotype *activation* involves the increased cognitive accessibility of characteristics associated with a particular group, stereotype *application* represents the use of stereotypes when responding to a group member. A primary goal of the present research was to examine sex discrimination in an experimental setting and to explore the effectiveness of a strategy, which was originally designed to reduce the activation of stereotypes, on this form of bias.

Although stereotype activation has been described in the past as a habit, a bad habit that is hard to “kick” (Devine, 1989; Devine & Monteith, 1993), one strategy

[☆]The authors would like to express their gratitude to Laurie Rudman, Thomas Mussweiler, and Irene Blair for their helpful comments on an earlier version of the manuscript. The research reported in this paper was supported by a Social Science and Humanities Council of Canada (SSHRC) grant and a Netherlands Science Foundation (NWO) grant to the first author.

* Corresponding author. Fax: +1-416-736-5814.

E-mail address: kawakami@yorku.ca (K. Kawakami).

that has recently been utilized to reduce the automatic activation of stereotypes is extensive training in negating category-trait associations (Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000). In particular, Kawakami and her colleagues found that just as practice is considered to be essential in learning stereotypes, practice may also be the key to unlearning these associations. In one study, participants were presented with photographs of Blacks and Whites and stereotypic and nonstereotypic traits under each photograph on a computer monitor. Participants in a training condition designed to reduce stereotype activation were instructed to negate racial stereotypes by responding “NO” to a photograph of a White person and a trait associated with Whites or a photograph of a Black person and a trait associated with Blacks. They were further instructed to respond “YES” to stereotype inconsistent word–picture pairings. Participants in the control condition were given the opposite instructions.

The results demonstrated that participants in the control condition, both before and after training, exhibited a persistent stereotype activation effect on a person categorization task. Although participants who were extensively trained to negate racial stereotypes initially also demonstrated stereotype activation, this effect was eliminated by the extensive training. Furthermore, Kawakami et al. found that practice effects of this type lasted up to 24 h following the training.

The aim of the present research was to examine whether the training strategy that had been shown to reduce stereotype activation would also be effective in ameliorating the preference of men over women for managerial positions. Although training appears to be a promising strategy in reducing *stereotype activation*, it is unclear what the impact of training would be on *behavioral responses* to group members, that is, to discrimination. One important factor in considering the effectiveness of such a strategy is the controllability of responses related to measurements of different types of bias (Bargh, 1999; Devine, 1989; Gilbert & Hixon, 1991; Stangor & Lange, 1994). While measures of stereotype activation are often automatic and therefore not amenable to strategic control, the application of stereotypes in job selection situations is a more controlled process. So even though the training may be effective in reducing the activation of stereotypes and other types of responses that are automatic and not controllable by nature (Bargh, 1999), when people can control their evaluations such as when they are deliberating on job candidates, training may be less effective and more influenced by momentary motivations.

Recent research related to the relationship between explicit and implicit processes reveals that reducing automatic activations may not necessarily reduce a more controlled form of bias (Blair, 2001; Dovidio, Kawakami, & Beach, 2000; Kawakami et al., 1998). Implicit

measures of stereotyping are generally related to spontaneous responses and are not correlated with deliberative responses because these responses may be differentially shaped by a number of factors (Dovidio, Kawakami, & Gaertner, 2002; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995). Thus, when investigating strategies related to the reduction of biases, it is critical to examine their impact separately for explicit and implicit processes.

Because the evaluation and choice of male and female job candidates is typically a deliberative process, it is possible that participants may attempt to manipulate their initial responses in systematic ways. Specifically, recent theorizing on mental correction (Wegener & Petty, 1997; Wilson & Brekke, 1994) indicates that when people think they are being unduly influenced and are able to correct for these influences, they may modify their assessments in a direction opposite to the perceived bias. Because the training technique to reduce bias used by Kawakami et al. (2000) is direct, obvious, and somewhat heavy-handed, people exposed to this training may resist these attempts at influence and strategically modify their responses to correct for the potential effect of this strategy.

In the present research, in particular, analogous to the procedures of Kawakami et al. (2000), some participants received extensive training (480 trials) aimed at reducing gender stereotyping (the Nonstereotypic Association Training condition). Then, in an ostensibly different study about employment decisions, participants were asked to choose among male and female candidates for a leadership position. In a control condition (the No Training condition), participants proceeded directly to the employment decision phase. Because participants are likely not to want to be influenced by the training, they may try to correct for the perceived effects of the training so that their responses reflect their “real” opinion. Theories of mental correction (Wegener & Petty, 1997; Wilson & Brekke, 1994) posit that when people are *motivated* and *able*, they may modify their assessments in a direction opposite to the perceived bias (Mussweiler & Neumann, 2000). If participants think that the training will bias their decisions in favor of female job candidates, they may try to minimize this influence and be less favorable towards a female than a male candidate. Thus, in correcting for the anticipated impact of the training, participants in the Nonstereotypic Association Training condition may exhibit a level of sex discrimination which is similar to participants in the No Training condition.

For correction to occur, however, people must first recognize that they have been influenced (e.g., by recognizing a direct relationship between their responses and the source of influence) and then have the mental

resources necessary to control their response (Martin, Seta, & Crelia, 1990; Schwarz & Bless, 1992; Wegener & Petty, 1997). If people are relatively unaware of an unwanted influence on their evaluations and/or are unable to adjust their responses because they do not have the necessary cognitive capacity (Wilson & Brekke, 1994), they will not modify their response and the uncorrected impact of the strategy will be apparent.

To examine whether correction processes could moderate the effects of training on subsequent hiring decisions, we initially questioned participants in a pilot study about their awareness of the potential impact of training and their anticipated reactions. We described the training task and the subsequent job application task and probed participants' naive theories related to these tasks. The findings supported our assumption that participants were aware that the training could systematically influence their responses and that they believed that the experimenter expected the training to reduce sex biases. On the basis of correction theory, it is therefore possible that participants would modify their responses to adjust for this influence under controlled processing conditions.

In the present study, we examined the possible influence of correction processes related to nonstereotypic association training in two ways—by disassociating the two tasks and by adding a cognitive load manipulation. Although we expected participants to correct for the effect of the perceived training when they received training and then directly performed the candidate selection task, we expected correction processes to be reduced when the training was disassociated from the job application task by including an intermediate filler task (the Nonstereotypic Association Plus Filler Task condition). We hypothesized that if the relationship between the nonstereotypic association training and the candidate selection task was less obvious, participants would be less likely to correct for the training. Without the operation of correction processes, we expected that nonstereotypic association training would be effective in reducing sex discrimination against women in the candidate selection task.

Besides awareness, mental resources are also necessary for correction to occur. Martin et al. (1990), for example, emphasized the importance of cognitive capacity to correction processes in a study that blatantly primed positive and negative concepts. In their study, half of the participants were simply asked to form an impression of a target person and the other participants were given an additional task (i.e., listening to a tape recording and tracking recited digits). Their results demonstrated that when not under cognitive load, participants corrected for the prime and formed less favorable impressions following the positive than the negative prime. In contrast, participants under cognitive load assimilated their evaluations to the prime and

formed more favorable impression following the positive than the negative prime.

In the present context, we further examined the possibility that participants may correct for the perceived influence of nonstereotypic association training by including a cognitive load manipulation (the Nonstereotypic Association Training Plus Probe Reaction Task condition). Specifically, while reviewing the applicants during the job application phase, participants were given a secondary task to impede their ability to correct (Gilbert, Pelham, & Krull, 1988). If participants do correct for the effects of training in general, participants under cognitive load would be less able to correct and the “true” effect of training on discrimination would become apparent. Thus, participants in this condition, like those in the Nonstereotypic Association Plus Filler Task condition, were expected to demonstrate reduced bias favoring male candidates in the job application task as a result of nonstereotypic association training.

In summary, the goal of the present research was to investigate whether nonstereotypic association training would be effective in reducing sex discrimination and to explore the potential moderating role of correction processes that might limit its effectiveness. Specifically, we hypothesized that because of correction processes, participants in the Nonstereotypic Association Training condition would show levels of discrimination equivalent to that of the No Training control group when hiring decision directly followed nonstereotypic training. In both of these conditions, participants were expected to prefer male over female candidates. However, when correction processes were seen as less relevant because the tasks were separated in time by an unrelated intermediate task (the Nonstereotypic Association Training Plus Filler Task condition) or when participants were unable to correct because they were impeded by cognitive load (the Nonstereotypic Association Training Plus Probe Reaction Task condition), the training was expected to be effective in ameliorating sex discrimination. Bias against female candidates was expected to be reduced in the latter two conditions and participants were expected to select men and women to the same extent.

Method

Participants and design

Seventy (33 female, 19 male and 18 who did not indicate their sex) undergraduate students in the Netherlands participated in the experiment, receiving approximately \$6. The main independent variable, Training, included four levels: No Training, Nonstereotypic Association Training, Nonstereotypic Association Training Plus Filler Task, and Nonstereotypic Association Training Plus Probe Reaction Task.

Procedure

Upon arrival, participants were assigned to one of four conditions and informed that the session would involve either one study (No Training condition) or two separate studies (the Training conditions). The primary aim of the “first study,” described as a decision making task, was to train participants to associate nonstereotypic words with specific social categories. The goal of the “second study,” described as a personnel selection task, was to examine the effect of the training on sex discrimination.

In the *training phase*, participants were seated behind a Macintosh Performa computer and a button box and were informed that they would be presented with a series of photographs of men and women. Furthermore, they were informed that underneath each photograph there would be two traits, one on the left and one on the right. They were instructed to select the trait that was not culturally associated with the gender category in the photograph and to press the appropriate button on the button box as quickly and as accurately as possible. For half of trials the correct answer was on the right side of the button box, for the other half of trials the correct answer was on the left side of the button box. For example, when participants were presented with a photograph of a woman with the traits “sensitive” and “strong,” the trait “strong” is the correct answer because strong is not culturally associated with women. Alternatively, if participants were presented with a photograph of a man with the traits “weak” and “messy,” the trait “weak” is the correct answer because that trait is not associated with men. Although the two traits that were presented simultaneously differed in their gender association, they were both either positive or negative. The photograph and the traits remained on the screen until the participant responded, after which a blank screen appeared for 1500 ms before the presentation of the next trial.

In total, participants received 480 trials divided into six blocks. In each block, 20 black and white photographs of men and women scanned from college year-

books were presented with one of 20 stereotypes of men and one of 20 stereotypes women under the photograph (see Table 1). Each block, therefore, consisted of 80 trials in which each male and female stereotype (40 in total) was presented with a photograph of a man or a woman. The trials were presented in a random order. After each block, participants were given a break and were asked to press the mouse when they were ready to continue the experiment.

To examine the effects of the training on sex discrimination, all participants were presented with the *job application task*. In this phase, participants were told that a neighboring university was carrying out a national study and had asked their psychology department to help investigate job-hiring procedures. Specifically, participants were presented with an actual advertisement for a chairperson of a District Doctor’s Association selected from a national newspaper. They were asked to read the advertisement along with four ostensibly real resumes and cover letters for the job.

The advertisement described a leadership job in which the successful candidate would supervise doctors in his or her district. Specifically, the candidate was expected to advise policy on financial, organizational, and juridical affairs, as well as supervise and set into motion action plans for health care in the district. The candidate was required to have strong management and negotiating skills, to be able to network well with people in the field, and to supervise the office staff. The function required an academic level of thinking and working, management qualities and experience, and inspirational leadership skills.

The resumes and cover letters of four candidates described their education and experience. Materials were developed such that all applicants were suitable for the advertised job. Two of the applicants, however, were given male names (i.e., Sander van Cleef and Anton Fleuren) and two of the applicants were given female names (i.e., Marian Koenders and Helma Driessen). The specific application and the sex of the applicant were counterbalanced. Participants were presented with one of four different orders of the applications. After reading

Table 1
Stereotypic stimulus words

Female Traits in Training		Male Traits in Training	
Modest	Gentle	Strong	Brave
Kind	Tender	Adventurous	Courageous
Tidy	Caring	Athletic	Powerful
Attentive	Understanding	Practical	Active
Sensitive	Impulsive	Realistic	Handy
Naïve	Jealous	Macho	Closed
Talkative	Vain	Aggressive	Stubborn
Anxious	Weak	Coarse	Abrupt
Vulnerable	Fickle	Untidy	Loud
Careful	Complaining	Headstrong	Messy

all the applications, participants were asked to choose the best candidate for the job.

Participants in the No Training condition performed only this task. Specifically, their procedure included reading the job advertisement and the applications and selecting the best applicant. Participants in the Nonstereotypic Association Training condition first completed the training task and then immediately afterwards were presented with the job applicant task. Participants in the Nonstereotypic Association Training Plus Filler Task condition, alternatively, performed an intermediate task *before* proceeding to the job application task. Specifically, after the training phase, to disassociate the training from the candidate selection task, these participants were presented with 15 arithmetic sums [e.g., $(2 \times 13 - 11) \times 16 - 15 = ___$] and instructed to calculate as many answers in their head as possible in 10 min and to write their responses on a piece of a paper. Although participants in the Nonstereotypic Association Training Plus Probe Reaction Task condition were given the candidate selection task directly after the nonstereotypic training as in the first two conditions, these participants were also given an additional probe reaction task while completing the latter phase. This additional activity was designed to utilize some of the participants' cognitive capacity (Macrae, Bodenhausen, Milne, & Wheeler, 1996; Pendry & Macrae, 1994), and thus, deplete resources for conscious control during the candidate selection phase. Specifically, while evaluating the applicants for the advertised position, these participants were told that a simulated light bulb located in the center of a computer screen would become illuminated at random intervals. Their task was simply to switch off the light bulb as fast as possible each time by pressing the appropriate button on a button-box. The presentation of the probe stimulus, which continued throughout the job application task, was completely randomized for each participant, and the intervals between each occurrence varied between 13 and 31 s.

Results

In this section, we first examined changes in speed of responding to nonstereotypic traits within the training phase and then consider differences in hiring decisions as a function of the four training conditions.

Nonstereotypic association training

The data from three participants were unavailable because of computer malfunction. For the rest of the participants, response latencies related to incorrect answers and outlier latencies that were more than three standard deviations from the mean (13.72%) were excluded. Analyses were performed on logarithmically

transformed response times, but the untransformed means (in milliseconds) are reported in the text for illustrative purposes.

To examine the effect of the amount of training on speed of responding, a 2 (Male vs. Female Photograph) \times 6 (Blocks) \times 3 (Training: Training vs. Training Plus Filler Task vs. Training Plus Probe Reaction Task) analysis of variance was performed on the transformed response latencies with the Photograph and Block variables as repeated measures. The Block variable was analyzed for linear trend.

Consistent with the findings of Kawakami et al. (2000) which indicated that participants are adept at learning new nonstereotypical associations, a significant linear effect for Block was obtained, $F(1, 51) = 110.88$, $p < .001$. The systematic decrease in response latencies from blocks 1 through 6 ($M_s = 2178, 2110, 1922, 1827, 1728, 1650$) indicates that participants were able to respond with increasing ease to new nonstereotypic associations with more practice. Moreover, as expected because the nonstereotypic training procedure was identical across the three training conditions (Nonstereotypic Association Training, Nonstereotypic Association Training Plus Filler Task, and Nonstereotypic Association Training Plus Probe Reaction Task), neither the Training \times Block interaction, $F(10, 245) = .73$, $p = .69$, or the Training \times linear effect, $F(2, 51) = .09$, $p = .77$, was significant.

Choice of candidate

Our initial analysis of the candidate selected for the job demonstrated that the choice of men and women differed systematically across the four conditions (No Training vs. Training vs. Training Plus Filler Task vs. Training Plus Probe Reaction Task), $\chi^2(3, N = 70) = 10.93$, $p = .01$ (see Fig. 1).¹

As expected based on correction theorizing, no difference was found between the participants' choice of male and female candidates in the No Training condition (67% men and 33% women) and the Nonstereotypic Association Training condition (78% men and 22% women), $\chi^2(1, N = 36) = .55$, $p = .46$. Overall, sex discrimination was observed across these two conditions: Men were chosen significantly more frequently than women, 72% vs. 28%, $\chi^2(1, N = 36) = 7.11$, $p < .01$.

Comparison of these two conditions with the Nonstereotypic Association Training Plus Filler Task condition provides further evidence for the operation of correction processes in the present context. As expected, the participants' choice for male and female candidates differed between the Nonstereotypic Association Train-

¹ Because the main effect and interaction effects for sex of participant were not significant ($p > .34$), sex was not included in the reported analyses.

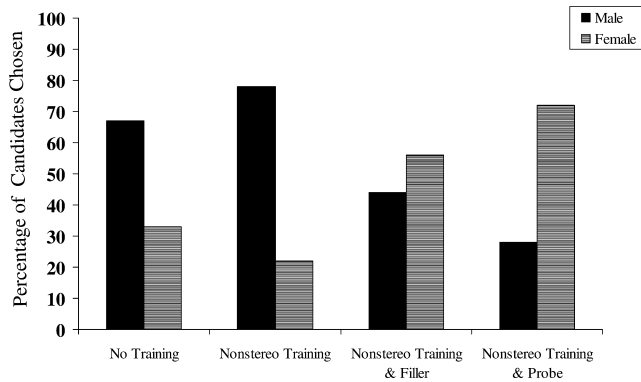


Fig. 1. The effects of nonstereotypic training and correction processes on choice of male and female candidates.

ing Plus Filler Task condition and the combination of the No Training and the Nonstereotypic Association Training conditions, $\chi^2(1, N = 52) = 3.87$, $p < .05$. Whereas, participants in the latter conditions demonstrated discrimination against women (see previous analysis), participants who received training and completed a filler task before judging the job applications did not differ in their choice of men and women (44% men vs. 56% women), $\chi^2(1, N = 16) = 0.25$, $p < .62$.

Comparison of responses in the Nonstereotypic Association Training Plus Probe Reaction Task condition also provides evidence for correction processes by differing from those in the combination of the No Training and the Nonstereotypic Association Training conditions, $\chi^2(1, N = 54) = 9.69$, $p < .003$. In contrast to the latter two conditions, which showed a marked preference for men (see previous analysis), the Nonstereotypic Association Training Plus Probe Reaction Task condition indicated that women were selected somewhat more often than men (72% women vs. 28% men), $\chi^2(1, N = 18) = 3.56$, $p < .06$.

As further anticipated, the two conditions that were designed to mitigate the potential effects of correction processes, the Nonstereotypic Association Training Plus Filler Task and the Nonstereotypic Association Training Plus Probe Reaction Task conditions, did not differ from one another in the pattern of selection of men and women for the managerial position, $\chi^2(1, N = 34) = .95$, $p < .33$.

Discussion

The results related to the training data demonstrated that with practice participants showed marked improvement in learning new, nonstereotypic associations. Nevertheless, the selection of male or female candidates for a managerial position revealed no difference between Nonstereotypic Association Training and No Training conditions, the same conditions in which Kawakami

et al. (2000) did show a difference in stereotype activation. In the present study, although extensive debriefing confirmed our assumption that participants were indeed attempting to choose the best candidate overall, in these conditions there was an overall pattern of discrimination against women relative to men in recommended hiring for a managerial position (Glick, 1991; Rudman & Glick, 1999).

One reason for this bias favoring men over women likely involves the nature of the advertised position—a managerial position in which the person would supervise doctors, need strong negotiation skills, and provide advice on financial, organizational, and juridical affairs. These responsibilities are more closely associated with the agentic qualities in stereotypes of men than with the communal characteristics involved in stereotypes of women (Glick, Zion, & Nelson, 1988; Rudman, 1998). Women's stereotypical "lack of fit" with supervisory positions may thus lead participants to see them as less desirable than male candidates (Heilman, 1983). Nevertheless, given that we counterbalanced the sex of the candidate and the specific credentials, our results revealed that *given the same documented qualifications*, women were recommended for the position less often than men. Thus, our research provides further evidence for the prevalence of sex discrimination (Rudman & Glick, 1999).

We have proposed that the similar levels of bias in both the Training and No Training conditions implicates the influence of correction processes that limit the effectiveness of training. The results of the Nonstereotypic Association Training Plus Filler Task and the Nonstereotypic Association Training Plus Probe Reaction Task conditions support this interpretation. In particular, when correction processes are curtailed by either disassociating the candidate selection task from the training or by reducing participants' cognitive capacity during the candidate selection phase, nonstereotypic association training is successful in reducing sex discrimination.

These findings suggest that just as in earlier theorizing related to reactance theory (Brehm & Brehm, 1981), if people think they are being unduly influenced they may counter this influence. Reactance and correction processes differ in fundamental ways, however. Whereas reactance theory posits that people, often without full deliberation, will react *against* threats that limit their personal sense of behavioral freedom, correction theory emphasizes more specific, deliberative, and strategic adjustments. In particular, correction theory proposes that people do not simply resist attempts at control, but they assess the direction and the extent of potential influence and then adjust and calibrate their responses to compensate for this impact (Mussweiler & Neumann, 2000; Wegener & Petty, 1997; Wilson & Brekke, 1994). Thus, whereas reactance to the loss of one's personal

control is somewhat automatic and diffuse, correction is a more cognitively demanding process that requires not only mental capacity but personal theories related to specific details concerning the perceived influence and ways to reduce the impact from this source. Our findings that both an intermediate task, which presumably limited awareness of the direct impact of training on candidate selection decisions, and a probe reaction task, which limited people's ability to perform controlled cognitive activities, moderated the effects of training in job candidate selection and implicate the role of the types of volitional assessment and controlled processes outlined in correction theory.

Importantly, in contrast to research that typically demonstrates that cognitive load *increases* stereotyping (Bodenhausen, 1990; Macrae, Hewstone, & Griffiths, 1993), we predicted and found that preference for male candidates *decreased* with cognitive load in the Training Plus Probe Reaction Task condition. Notably, participants in this condition actually tended to favor women over men in hiring (72% vs. 28%). Although unanticipated, these findings are consistent with recent theorizing on the effects of cognitive load on thought suppression (Wegner, 1994; Wegner & Erber, 1992; Wegener & Petty, 1997). In particular, this theorizing suggests that attempts to reduce perceived influences under conditions of high cognitive load may actually exacerbate the impact of this influence. An inability to modify one's responses produced by a concurrent task may therefore not only render corrective attempts to choose less women ineffective but can because of constant monitoring for the unwanted influence of the training accentuate its effects which would result in participants choosing more women when under cognitive load after completing the nonstereotypic association training.

In general, the current work indicates that researchers interested in interventions to combat bias should consider the effects of strategies aimed at changing stereotypes on a variety of levels. Although the automatic activation of stereotypes and attitudes has been shown to predict a range of responses (Dovidio et al., 1997; Fazio et al., 1995), stereotype activation will not always map directly on to manifested behavior (see Dovidio et al., 2002; Kawakami, Dovidio, & Spears, 2004; Kawakami, Young, & Dovidio, 2002). Because decision making when people have the time and opportunity to control their responses may be strongly shaped by personal values and temporary motivations, strategies aimed at changing the automatic activation of stereotypes will not necessarily result in reduced discrimination. Although a major focus in recent research has been on the reduction of stereotype activation (see Blair, 2002; for a review), it is unclear what the effects of these strategies will be on other manifestations of bias and on the use of stereotypes when encountering actual group members.

Two factors may be particularly important when examining the impact of strategies designed to reduce discrimination. First, the "heavy-handedness" of the procedure needs to be carefully considered. As Stephan and Stephan (2001) identify in their comprehensive review of anti-bias education programs, interventions vary substantially in their directness of purpose and method. When the intervention is very obvious, its effectiveness under explicit processing conditions may be compromised. That is, people in these training programs, whether they participate reluctantly or willingly, may subsequently attempt to correct their responses to compensate for such overt attempts to influence their behavior. As our findings illustrate, these correction processes may mitigate at least some of the benefits of the training. The present study, thus underscores the importance of examining the individual's naïve theories related to the effects of the specific procedure and to confront possible influences related to these theories.

Second, it is important to consider the processing conditions related to judgments of target group members. When the conditions in which a response is assessed limit conscious control (e.g., under high cognitive load or when the behaviors are difficult to monitor or manipulate; see Dovidio et al., 1997, 2002; Kawakami et al., 1998; Kawakami & Dovidio, 2001), the relationship between automatic activation and behavior is likely to be relatively strong and direct. We found in the present study that training was effective in reducing sex discrimination when participants experienced high cognitive demand as in the Kawakami et al. (2000) stereotype activation research. However, when the conditions allow conscious control of responses and deliberation, as was the case for the candidate selection phase without an additional probe reaction task, the impact of a strategy aimed at reducing discrimination or the activation of stereotypes may diverge.

Although demonstrating that strategies that have been proven to be successful at reducing stereotype activation may not be as effective in reducing discrimination has theoretical and practical significance, it does not mean that the strategy is any less valuable—its usefulness may simply be limited when impressions are more deliberative. Importantly, however, decision made under both more implicit and explicit conditions can have widespread and important consequences for minority group members and women. Future research, therefore, needs to continue to examine ways to kick the stereotyping habit under *both* automatic and controlled processing conditions by considering the nature of the intervention and of the response that it is targeted to change, as well as the potential role of motivations and other conscious influences that are likely to shape responses in a given context.

References

- Allport, G. W. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.
- Bargh, J. A. (1999). The cognitive monster: The case against controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 361–382). New York: Guilford Press.
- Blair, I. (2001). Implicit stereotypes and prejudice. In G. Moskowitz (Ed.), *Cognitive social psychology: On the tenure and future of social cognition* (pp. 359–374). Mahwah, NJ: Erlbaum.
- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, 6, 242–261.
- Bodenhausen, G. V. (1990). Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science*, 1, 319–322.
- Brehm, S., & Brehm, J. (1981). *Psychological reactance: A theory of freedom and control*. San Diego, CA: Academic Press.
- Deaux, K., & LaFrance, M. (1998). Gender. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (Vol. 1, 4th ed., pp. 788–827). New York: McGraw-Hill.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5–18.
- Devine, P. G., & Monteith, M. J. (1993). The role of discrepancy-associated affect in prejudice reduction. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition and stereotyping: Interactive processes in group perception* (pp. 317–344). San Diego CA: Academic Press.
- Dovidio, J. F., Brigham, J. C., Johnson, B. T., & Gaertner, S. L. (1996). Stereotyping, prejudice, and discrimination: Another look. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 276–319). New York: Guilford.
- Dovidio, J., Kawakami, K., & Beach, K. (2000). Examination of the relationship between implicit and explicit measures of intergroup bias. In R. Brown & S. L. Gaertner (Eds.), *Intergroup processes: Vol. 4. Blackwell handbook of social psychology* (pp. 175–197). Oxford: Blackwell.
- Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, 82, 62–68.
- Dovidio, J., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). The nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, 33, 510–540.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, 69, 1013–1027.
- Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, 60, 509–517.
- Gilbert, D. T., Pelham, B. W., & Krull, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology*, 54, 733–740.
- Glick, P. (1991). Trait-based and sex-based discrimination in occupational prestige, occupational salary, and hiring. *Sex Roles*, 25, 351–378.
- Glick, P., Zion, C., & Nelson, C. (1988). What mediates gender discrimination in hiring decisions? *Journal of Personality and Social Psychology*, 55, 178–186.
- Heilman, M. E. (1983). Sex bias in work settings: The lack of fit model. *Research in Organizational Behavior*, 5, 269–298.
- Kawakami, K., Dion, K. L., & Dovidio, J. F. (1998). Racial prejudice and stereotype activation. *Personality and Social Psychology Bulletin*, 24, 407–416.
- Kawakami, K., & Dovidio, J. F. (2001). Implicit stereotyping: How reliable is it? *Personality and Social Psychology Bulletin*, 27, 212–225.
- Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, 78, 871–888.
- Kawakami, K., Dovidio, J. F., Spears, R. (2004). A change of habit: The consequences of inducing bias for motivation and stereotyping. Under review.
- Kawakami, K., Young, H., & Dovidio, J. F. (2002). Automatic stereotyping: Category, trait, and behavioral activations. *Personality and Social Psychology Bulletin*, 28, 3–15.
- Kite, M. E., Russo, N. F., Brehm, S. S., Fouad, N. A., Hall, C. C. I., Hyde, J. S., & Keita, G. P. (2001). Women psychologists in academe. *American Psychologist*, 56, 1080–1098.
- Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Wheeler, V. (1996). On resisting the temptation for simplification: Counterintentional consequences of stereotype suppression on social memory. *Social Cognition*, 14, 1–20.
- Macrae, C. N., Hewstone, M., & Griffiths, R. (1993). Processing load and memory for stereotype-based information. *European Journal of Social Psychology*, 23, 77–87.
- Martin, L. L., Seta, J. J., & Crelia, R. A. (1990). Assimilation and contrast as a function of people's willingness and ability to expend effort in forming an impression. *Journal of Personality and Social Psychology*, 59, 27–37.
- Mussweiler, T., & Neumann, R. (2000). Sources of mental contamination: Comparing the effects of self-generated versus externally provided primes. *Journal of Experimental Social Psychology*, 36, 194–206.
- Pendry, L. F., & Macrae, C. N. (1994). Stereotypes and mental life: The case of the motivated but thwarted tactician. *Journal of Experimental Social Psychology*, 30, 303–325.
- Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology*, 74, 629–645.
- Rudman, L. A., & Glick, P. (1999). Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler, image of middle management. *Journal of Personality and Social Psychology*, 77, 1004–1010.
- Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgment* (pp. 217–245). Hillsdale, NJ: Erlbaum.
- Stangor, C., & Lange, J. E. (1994). Mental representations of social groups: Advances in understanding stereotypes and stereotyping. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (26, pp. 357–416). New York: Academic Press.
- Stephan, W. G., & Stephan, C. W. (2001). *Improving intergroup relations*. Thousand Oaks, CA: Sage.
- Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, 101, 34–52.
- Wegner, D. M., & Erber, R. (1992). The hyperaccessibility of suppressed thoughts. *Journal of Personality and Social Psychology*, 63, 903–912.
- Wegner, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naïve theories of bias in bias correction. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (29, pp. 141–208). New York: Academic Press.
- Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, 116, 117–142.