European Journal of Social Psychology, *Eur. J. Soc. Psychol.* (2013)
Published online in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/ejsp.1941

Fast track report

Counter-prejudicial training reduces activation of biased associations and enhances response monitoring

JIMMY CALANCHINI[1]*, KAREN GONSALKORALE[2], JEFFREY W. SHERMAN[1] AND KARL CHRISTOPH KLAUER[3]

[1]University of California, Davis, Davis, USA; [2]University of Sydney, Sydney, Australia; [3]University of Freiburg, Freiburg im Breisgau, Germany

Abstract

Although implicitly measured bias was once assumed to be highly stable, subsequent research has shown that it is, in fact, malleable. One technique for altering implicit bias is through counter-prejudicial training. At least two broad mechanisms may drive this effect. First, training people to respond in counter-prejudicial ways may diminish the extent to which biased associations are activated in memory. Second, training may strengthen processes that reduce the influence of biased associations on responses. Participants received either counter-prejudicial, pro-prejudicial, or no training and then completed an implicit measure of bias. Application of the quadruple process model revealed support for both mechanisms: Counter-prejudicial training produced less activation of biased associations as well as enhanced detection of appropriate responses compared with pro-prejudicial or no training. Implications for the development of bias-reduction training are discussed. Copyright © 2013 John Wiley & Sons, Ltd.

Although implicitly measured bias was once assumed to be highly stable (Bargh, 1999; Dovidio & Fazio, 1992; Fazio, Jackson, Dunton, & Williams, 1995), subsequent research has shown that it is, in fact, malleable (Blair, 2002). One technique for altering implicit bias is through counter-prejudicial training (Gawronski, Deutsch, Mbirkou, Seibt, & Strack, 2008; Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000; Kawakami, Dovidio, & van Kamp, 2005, 2007; Plant, Peruche, & Butz, 2005). At least two broad mechanisms may drive this effect. First, training people to respond in counter-prejudicial ways may diminish the extent to which biased associations are activated in memory. Second, training may strengthen control processes that reduce the influence of biased associations on implicit task performance.[1]

Recent data suggest that a combination of reduced activation of biased associations and enhanced control underlies the effects of training. It is well established that people who are internally but not externally motivated to control prejudice (high IMS/low EMS) show less implicit bias than people who either are unmotivated or are externally motivated (high EMS) to control prejudice (Amodio, Devine, & Harmon-Jones, 2008; Amodio, Harmon-Jones, & Devine, 2003; Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002). Gonsalkorale, Sherman, Allen, Klauer, and Amodio (2011) examined differences in implicit attitude processes among these groups and found that people who are internally but not externally motivated have both less activation of biased associations and enhanced control over the influence of those associations compared with other respondents.

One prominent account of the development and influence of attitudes among those with high IMS/low EMS is that these individuals train themselves to respond in unbiased ways (e.g., Monteith, Ashburn-Nardo, Voils, & Czopp, 2002). This training involves learning to identify contexts in which biased responses are likely and learning to replace biased responses with more egalitarian responses. One result is less implicit bias. If externally directed counter-prejudicial training affects implicit bias in a similar fashion to self-training, then external training may also be expected to both reduce the activation of biased associations and enhance the operation of control during the completion of implicit measures of bias. The purpose of the current research was to examine this possibility.

The quadruple process model (quad model; Sherman et al., 2008) is well suited for this kind of process-level investigation. The quad model is a multinomial model

*Correspondence to: Jimmy Calanchini, 134 Young Hall, 1 Shields Ave, University of California, Davis, Davis, CA 95616, USA.
E-mail: jcalanchini@ucdavis.edu

[1]By "implicit," we mean indirect, in contrast to explicit or direct. Throughout this paper, we will use different terms in order to distinguish among underlying evaluative associations ("implicit association"), the tools that have been developed to measure such associations ("implicit measure" and "implicit task"), and behavioral outcomes of implicit measures such as reaction time effects ("implicit attitude" and "implicit bias"). The term "implicit" has become a synonym for "indirect" in common vernacular, and it is as such that we use it here. However, we do not assume that responses on implicit measures are isomorphic with underlying evaluative associations that instigate responses on the measure.

Received 2 November 2012, Accepted 27 February 2013

(Batchelder & Riefer, 1999; Riefer & Batchelder, 1988) that provides a mathematical means to estimate the independent contributions of different processes to implicit task performance. The model proposes that implicit task performance depends jointly on the activation of implicit associations in memory (activation [AC]), the ability to detect correct responses on the task (detection [D]), success at overcoming biased associations when they would produce an incorrect response (overcoming bias [OB]), and the influence of general guessing or response biases that may influence behavior in the absence of other available guides to response (guessing [G]). Whereas the AC parameter reflects the relatively automatic influence of underlying associations, the D and OB parameters reflect the operation of more controlled processes that may work in opposition to the associations and reduce demonstrated implicit bias (e.g., Sherman et al., 2008). The quad model has been extensively validated (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Sherman et al., 2008).

THE PRESENT STUDY

Research has shown that people with high internal and low external motivations to control prejudice show less implicit bias than others. Gonsalkorale et al. (2011) applied the quad model and showed that this effect was related to less activation of biased associations (AC) and greater levels of detection (D) among high IMS/low EMS participants. If externally directed training influences bias via the same mechanisms as those involving IMS and EMS, then we should expect that participants trained to respond in a counter-prejudicial manner should subsequently demonstrate reduced activation of biased associations (AC) and enhanced detection of correct responses (D). To investigate this possibility, participants received counter-prejudicial training, pro-prejudicial training, or no training and then completed an implicit measure of racial bias.

Method

Two hundred thirty-six undergraduates (144 women) participated for partial course credit. First, participants in the training conditions completed a task in which they responded to a series of pictures of Black and White men paired with positive and negative words. The training procedures were similar to those used by Kawakami et al. (2000). Participants in the counter-prejudicial condition ($N = 67$) were trained to affirm Black–positive and White–negative picture pairings by pressing the YES key (the + key) and to disaffirm Black–negative and White–positive picture pairings by pressing the NO key (the ~ key). Participants in the pro-prejudicial condition ($N = 70$) were trained to affirm Black–negative and White–positive picture pairings and to disaffirm Black–positive and White–negative picture pairings. The stimuli consisted of four blocks of 60 trials in which pictures of 10 Black and 10 White men with neutral expressions (taken from Minear & Park, 2004) were randomly paired with 20 positive and 20 negative words. Each 60-trial block consisted of 15 trials of Black/positive, Black/negative, White/positive, and White/negative pairings. Each picture/word pair remained onscreen until the participant made a correct response. Participants in the no training condition ($N = 99$) completed no task.

Next, each participant completed an evaluative Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998), in which they paired pleasant and unpleasant words with pictures of Black and White men. Participants first completed two 20-trial practice blocks, in which they discriminated pleasant from unpleasant words, and White from Black faces. The third and fourth blocks were critical blocks consisting of 20 and 40 trials, respectively. Participants were instructed to press one key whenever they saw a picture of a White person or a pleasant word, and another key whenever they saw picture of a Black person or an unpleasant word. The keys used to categorize Black and Whites faces were switched in subsequent blocks. The fifth block was a 20-trial practice block in which participants discriminated Black from White faces using the new key assignments. The sixth and seventh blocks were critical blocks consisting of 20 and 40 trials, respectively. Participants were instructed to press one key whenever they saw a picture of a White person or an unpleasant word, and another key when they saw a picture of a Black person or a pleasant word. Participants who respond more quickly when "Black" shares a key with "unpleasant" (commonly referred to as a "compatible" trial) than when it shares a key with "pleasant" (commonly referred to as an "incompatible" trial) are thought to have an implicit preference for Whites relative to Blacks (Greenwald et al., 1998). Target category and attribute labels remained on the top left and top right of the screen throughout the task, while stimulus pictures and words appeared at the center of the screen. A red "X" appeared whenever participants made an error, and they were required to correct the error before moving onto the next trial. Latencies were recorded to the correct response. Participants were instructed to make their classifications as quickly and accurately as possible.

Pleasant words were always categorized using the "I" key and unpleasant words using the "E" key. Both the pictures and words used in the IAT were different from the ones used in the training task. The words used as stimuli in both the training task and the IAT are reported in the appendix.[2]

Results

IAT Bias

Implicit Association Test scores were calculated according to the algorithm described by Greenwald, Nosek, and Banaji (2003). A one-way ANOVA revealed a significant effect of training, $F(2, 233) = 12.99$, $p < .001$, $\eta^2_{\text{partial}} = 0.10$. Participants who received counter-prejudicial training showed less IAT bias ($M = 0.36$, 95% $CI = [0.28, 0.45]$) than participants who received pro-prejudicial training ($M = 0.57$, 95% $CI = [0.49, 0.66]$), $t(135) = 3.68$, $p < .001$, Cohen's $d = 0.63$, and participants who received no training ($M = 0.64$, 95% $CI = [0.57, 0.71]$), t

[2]Some participants completed the IAT with a 675-millisecond response deadline, whereas others completed it with no deadline. This variable did not interact with training condition and was not included in subsequent analyses.

(164) = 4.93, $p < .001$, Cohen's $d = 0.77$. There was no difference in IAT bias between participants who received pro-prejudicial and no training, $t(167) = 1.20$, $p = .23$, Cohen's $d = 0.19$.

Modeling

We applied the quad model in order to explore the processes responsible for the differences between training conditions. The structure of the quad model is depicted as a processing tree in Figure 1. In the tree, each path represents a likelihood. Processing parameters with lines leading to them are conditional on all preceding parameters. For instance, OB is conditional on both AC and D. The conditional relationships described by the model form a system of equations that predicts the numbers of correct and incorrect responses in different conditions (e.g., compatible and incompatible trials). For example, there are three ways in which an incorrect response can be returned on an incompatible trial, in which Black and pleasant share a response key. The first is the likelihood that biased associations are activated (AC), detection succeeds (D), and OB fails (1 − OB), which can be represented by the equation AC × D × (1 − OB). The second is the likelihood that the biased associations are activated (AC) and detection fails (1 − D), which can be represented by the equation AC × (1 − D). The third is the likelihood that biased associations are not activated (1 − AC), detection fails (1 − D), and a bias toward guessing "unpleasant" (1 − G) produces an incorrect response, which can be represented by the equation (1 − AC) × (1 − D) × (1 − G). As such, the overall likelihood of producing an incorrect response on an incompatible trial is the sum of these three conditional probabilities: [AC × D × (1 − OB)] + [AC × (1 − D)] + [(1 − AC) × (1 − D) × (1 − G)]. The respective equations for each item category (e.g., Black faces, White faces, pleasant words, and unpleasant words in both compatible and incompatible blocks) are then used to predict the observed proportions of errors in a given data set. The model's predictions are compared with the actual data to determine the model's ability to account for the data. A χ^2 estimate is computed for the difference between the predicted and observed errors. To best approximate the model to the data, the parameter values are changed through maximum likelihood estimation until they produce a minimum possible value of the χ^2. The final parameter values that result from this process are interpreted as relative levels of the processes.

For each group, we calculated parameter estimates of AC, D, OB, and G. The G parameter was coded so that higher scores represented a bias toward guessing with the positive (pleasant) key. Two separate AC parameters were estimated: one measuring the extent to which associations between Black and unpleasant (BAC) were activated in performing the task and another measuring the extent to which associations between White and pleasant (WAC) were activated. The overall error rate for the IAT was 11.5% and χ^2 for model fit was 26.71, $df = 1$, $p < .001$. Tests of model fit are dependent on sample size, such that minute deviations from the model can jeopardize model fit when power is high (Cohen, 1988). However, the effect size of the difference between the actual data and the model's predicted data was $w = 0.03$, indicating good fit when controlling for power.

As Table 1 shows, BAC estimates were lower for participants who received counter-prejudicial training ($M = 0.09$) than for participants who received pro-prejudicial training ($M = 0.14$), $\Delta\chi^2 = 6.89$, $df = 1$, $p < .01$, $w = 0.02$, or no training ($M = 0.13$), $\Delta\chi^2 = 4.33$, $df = 1$, $p = .04$, $w = 0.01$. Similarly, WAC estimates were lower for participants who received counter-prejudicial training ($M = 0.09$) than for participants who received pro-prejudicial training ($M = 0.15$), $\Delta\chi^2 = 7.63$, $df = 1$, $p < .01$, $w = 0.02$, or no training ($M = 0.14$), $\Delta\chi^2 = 6.61$, $df = 1$, $p = .01$ $w = 0.02$. Additionally, D estimates were higher for participants who received counter-prejudicial training ($M = 0.85$) than for participants who received pro-prejudicial training ($M = 0.82$), $\Delta\chi^2 = 4.14$, $df = 1$, $p = .04$, $w = 0.02$, or no training ($M = 0.81$), $\Delta\chi^2 = 8.33$, $df = 1$, $p < .01$, $w = 0.02$. However, OB estimates were not different for participants who received counter-prejudicial training and participants who received either pro-prejudicial training, $\Delta\chi^2 = 0.64$, $df = 1$, $p = .42$, $w = 0.007$ or no training, $\Delta\chi^2 = 1.59$, $df = 1$, $p = .21$ $w = 0.009$. Similarly, G estimates were not different for participants who received counter-prejudicial training and participants who received either pro-prejudicial training, $\Delta\chi^2 = 0.21$, $df = 1$, $p = .65$, $w = 0.004$, or no training, $\Delta\chi^2 = 0.01$, $df = 1$, $p = .92$ $w = 0.004$. None of the parameters differed between the pro-prejudicial training and no training

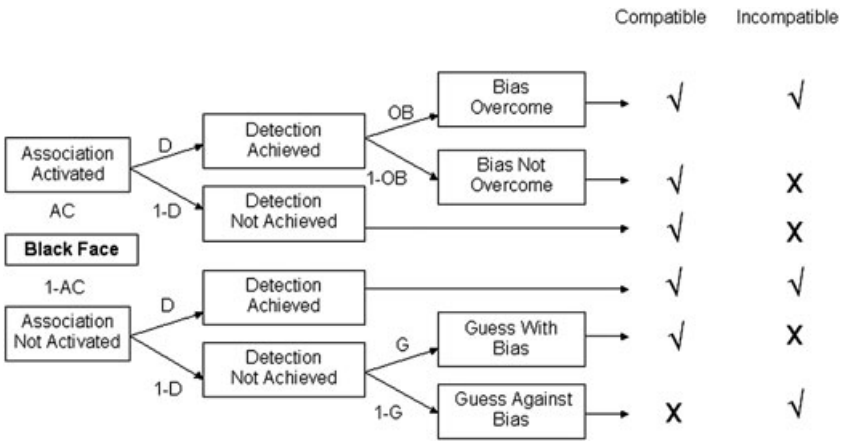

Figure 1. The quadruple process model (quad model). Each path represents a likelihood. Parameters with lines leading to them are conditional upon all preceding parameters. The table on the right side of the figure depicts correct (√) and incorrect (X) responses as a function of process pattern

Table 1. IAT parameter estimates

	No training		Pro-prejudicial training		Counter-prejudicial training	
BAC	0.13	[0.10–0.15]	0.14	[0.11–0.17]	0.09	[0.06–0.11]
WAC	0.14	[0.12–0.16]	0.15	[0.12–0.17]	0.09	[0.07–0.12]
D	0.81	[0.80–0.83]	0.82	[0.80–0.84]	0.85	[0.83–0.86]
OB	0.94	[0.81–1.06]	0.89	[0.75–1.04]	0.79	[0.58–0.99]
G	0.53	[0.49–0.56]	0.55	[0.50–0.59]	0.53	[0.48–0.57]

Note: BAC = Black–unpleasant associations; WAC = White–pleasant associations; D = detection; OB = overcoming bias; G = guessing. [95% confidence intervals].

conditions, all $\Delta\chi^2 < 0.57$, $ps > .46$, $ws < 0.005$. Thus, counter-prejudicial training decreased the activation of both anti-Black and pro-White associations, and increased the likelihood of detecting correct responses, compared with either pro-prejudicial or no training.[3]

DISCUSSION

Previous research has reported that counter-prejudicial training leads to decreased bias. Two competing theoretical accounts have been proposed to explain this effect, and the purpose of the present research was to directly test these explanations. Specifically, we tested whether training decreases the extent to which biased associations are activated, increases control, or both. To test these accounts, we used the quad model to separately estimate the contributions of multiple processes to implicit bias. Results showed that participants who received counter-prejudicial training showed both less activation of biased associations and enhanced detection of appropriate responses compared with participants who received either pro-prejudicial training or no training, providing evidence for both accounts.

These findings expand our understanding of the process of prejudice reduction. Previous research showed that high IMS/low EMS people demonstrate less implicit bias, and that this effect is related to both reduced activation of biased associations and increased detection of correct responses among these respondents (Gonsalkorale et al., 2011). However, it is an open question whether high IMS/low EMS individuals have less biased associations because of extensive practice at monitoring and replacing biased responses, if they simply never had biased associations in the first place, or if having less biased associations facilitates the development of effective response monitoring. The results of the present study indicate that extensive practice at monitoring and replacing biased responses can both enhance detection and reduce the activation of biased associations.

Interestingly, extensive practice at affirming biased responses neither increased the activation of biased associations nor affected detection. That biased associations were unaffected by pro-prejudicial training suggests a ceiling effect for bias, or perhaps it is simply harder to strengthen associations than it is to reduce them. Although further research is needed to explore these possibilities, it is encouraging that bias in this case was easier to reduce than enhance. Moreover, the fact that detection was unaffected by pro-prejudicial training highlights the importance of the content of the training: Detection is not enhanced through practice alone, but rather specifically through practice attending to counter-prejudicial pairings.

It is possible that participants in this experiment may have been aware of the purpose of the training, as they may have been in previous demonstrations of training effects. The goal of this research was to examine the processes through which training-induced reductions in bias are achieved, intentionally or not. We make no strong claim regarding the (un)intentional nature of the reactions to the training and do not view the question of intentionality as central to the research question.

Finally, these findings also expand prior research on bias-reduction training. Specifically, the effect of counter-prejudicial training on associations and control has implications for the development of bias-reduction interventions. Individuals may display implicit bias either because they have biased associations that are activated or because they are unable to exert control in responding to the task. Presumably, interventions to reduce bias would need to address participants' specific processing deficit (i.e., associations vs. control). However, counter-prejudicial training similar to what was used in this study would appear to be effective at reducing bias both for people who have strongly biased associations and people who have poor control.

[3]Readers may be interested in the correlations between the parameter estimates and IAT performance. In the counter-prejudicial training condition, IAT bias correlates with BAC, $r = .31$, $p = .012$; and WAC, $r = .38$, $p = .002$; but not D, $r = -.01$, $p = .92$; OB, $r = .21$, $p = .09$; or G, $r = .01$, $p = .98$. In the pro-prejudicial training condition, IAT bias correlates with BAC, $r = .55$, $p < .001$; and WAC, $r = .33$, $p = .006$; but not D, $r = .09$, $p = .48$; OB, $r = .17$, $p = .17$; or G, $r = .23$, $p = .06$. In the no training condition, IAT bias correlates with BAC, $r = .35$, $p < .001$; and WAC, $r = .25$, $p = .02$; but not D, $r = .12$, $p = .22$; OB, $r = .07$, $p = .52$; or G, $r = .09$, $p = .40$. The meaning of such correlations is unclear, however, because both quad model parameters and IAT scores (using the improved scoring algorithm: Greenwald et al., 2003) are derived from performance accuracy data. Deriving separate estimates (i.e., quad parameters and IAT scores) from the same data can result in correlated random errors, which may bias the variance/covariance structure that exists between them and thereby complicate any interpretation of correlations between the estimates (Klauer, 2010). Because subject-level correlations necessarily depend on individual-level parameter estimates, they are especially vulnerable to the biasing effects of correlated random errors. That is, because individual-level parameter estimates are derived from a small number of observations, they are especially vulnerable to this problem (Cohen, Sanborn, & Shiffrin, 2008). In contrast, aggregation across conditions reduces the impact of random errors, which is why we report only aggregate-level analyses in the body of this paper.Note that these correlations are based on both a different means of estimating IAT bias and a different set of equations for estimating the D parameter in the quad model than those reported by Conrey et al. (2005). Specifically, the original quad model equations reported in Conrey et al. (2005) for estimating D were based, in part, on the single-category IAT practice items, which was problematic for many reasons. Since then, every paper published by our lab using the quad model has estimated D only from the dual-category IAT test trials (Sherman et al., 2008).

ACKNOWLEDGEMENT

This research was supported by a grant from the National Science Foundation (BCS 0820855) to the third author.

REFERENCES

Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2008). Individual differences in the regulation of intergroup bias: The role of conflict monitoring and neural signals for control. *Journal of Personality and Social Psychology*, *94*(1), 60–74.

Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality and Social Psychology*, *84*(4), 738–753.

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S Chaiken, & Y Trope (1999) (Eds.), *Dual-process theories in social psychology* (pp. 361–382). New York, NY, US: Guilford Press.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review*, *6*(1), 57–86.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, *6*(3), 242–261.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, A. L., Sanborn, A. N., & Shiffrin, R. M. (2008). Model evaluation using grouped or individual data. *Psychonomic Bulletin and Review*, *15*, 692–712.

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: The quad Model of implicit task performance. *Journal of Personality and Social Psychology*, *89*(4), 469–487.

Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology*, *82*(5), 835–848.

Dovidio, J. F., & Fazio, R. H. (1992). New technologies for the direct and indirect assessment of attitudes. In *Questions about questions: Inquiries into the cognitive bases of surveys* (pp. 204–237). New York, NY, US: Russell Sage Foundation.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobstrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*(6), 1013–1027.

Gawronski, B., Deutsch, R., Mbirkou, 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*(2), 370–377.

Gonsalkorale, K., Sherman, J. W., Allen, T. J., Klauer, K. C., & Amodio, D. M. (2011). Accounting for successful control of implicit racial bias: The roles of association activation, response monitoring, and overcoming bias. *Personality and Social Psychology Bulletin*, *37*(11), 1534–1545.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216.

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*(5), 871–888.

Kawakami, K., Dovidio, J. F., & van Kamp, S. (2005). Kicking the habit: Effects of nonstereotypic association training and correction processes on hiring decisions. *Journal of Experimental Social Psychology*, *41*(1), 68–75.

Kawakami, K., Dovidio, J. F., & van Kamp, S. (2007). The impact of counterstereotypic training and related correction processes on the application of stereotypes. *Group Processes & Intergroup Relations*, *10*(2), 139–156.

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika*, *75*(1), 70–98.

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 630–633.

Monteith, M. J., Ashburn-Nardo, L., Voils, C. I., & Czopp, A. M. (2002). Putting the brakes on prejudice: On the development and operation of cues for control. *Journal of Personality and Social Psychology*, *83*(5), 1029–1050.

Plant, E. A., Peruche, B. M., & Butz, D. A. (2005). Eliminating automatic racial bias: Making race non-diagnostic for responses to criminal suspects. *Journal of Experimental Social Psychology*, *41*(2), 141–156.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, *95*(3), 318–339.

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review*, *115*(2), 314–335.

APPENDIX

Words used as stimuli in the training task:

Negative words:	Positive words:
Accident	Brilliant
Awful	Celebrate
Cancer	Cheer
Crash	Diamond
Destroy	Excitement
Disaster	Fabulous
Filth	Freedom
Grief	Gift
Gross	Glee
Hatred	Health
Noxious	Heaven
Painful	Glad
Poison	Lucky
Pollute	Paradise
Rotten	Rainbow
Stink	Splendid
Tragedy	Sunrise
Vomit	Superb
War	Triumph
Yucky	Vacation

Words used as stimuli in the IAT:

Negative words:	Positive words:
Agony	Happy
Death	Laughter
Evil	Love
Hatred	Peace
Sickness	Pleasure