

Reducing Implicit Bias: How Moral Motivation Helps People Refrain From Making “Automatic” Prejudiced Associations

Félice van Nunspeet
Leiden University

Naomi Ellemers and Belle Derks
Utrecht University

During the past years, evidence has accumulated to suggest that even those who endorse egalitarian views unwittingly make stereotypical associations and show implicit prejudice, which determines how they evaluate and respond to others. Growing knowledge of these effects has led members of the general public and practitioners to conclude that prejudice is to some extent inevitable, as it is seen to rely on “automatic” and “uncontrollable” associations that affect the way we think about others. In this contribution, we review a recent program of research revealing that even implicit bias can be reduced by activating people’s motivation to be moral. We also present evidence from brain potentials showing how neurocognitive processes are adapted to achieve this. We explain how these insights can benefit members of the general public and practitioners who hope to prevent that judgments about others are affected by implicit bias.

Keywords: implicit bias, IAT, morality, event-related potentials, prejudice reduction

With the installation of antidiscrimination legislation and formal presence of equal rights, many of us hope or assume that stereotyping, prejudice, and discrimination¹ are things of the past. Nevertheless, scientists have argued that discrimination has simply “gone underground,” demonstrating that people still make stereotypical inferences when they evaluate others (e.g., Barreto & Ellemers, 2013; Gaertner & Dovidio, 1986). Indeed, during the past years, psychological research has revealed the widespread pres-

ence of so-called implicit associations. These refer to the general tendency to more easily, quickly, and accurately associate individuals with traits or attributes that match (rather than oppose) the stereotypical characteristics and evaluations of their group. Such implicit evaluative and stereotypical associations have even been documented in people who hold egalitarian views and seem to emerge relatively independently of whether or not people indicate endorsing these stereotypes in more explicit statements (e.g., Greenwald & Banaji, 1995; Greenwald, McGhee, & Schwartz, 1998). Nevertheless, the persistence of implicit associations can have far-reaching consequences for actual intergroup behavior. For instance, these associations predict negative interactions with

Félice van Nunspeet, Social and Organizational Psychology Unit, Institute of Psychology, Faculty of Social and Behavioral Sciences, Leiden University; Naomi Ellemers, Faculty of Social Sciences, Utrecht University; Belle Derks, Social and Organizational Psychology, Faculty of Social Sciences, Utrecht University.

This work was supported by the SNS-Reaal KNAW-Merian prize and the NWO-Spinoza prize awarded to Naomi Ellemers by the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organization for Scientific Research (NWO), as well as a NWO-VENI grant (451-08-022) awarded to Belle Derks.

Correspondence concerning this article should be addressed to Félice van Nunspeet, Social & Organizational Psychology Unit, Institute of Psychology, Faculty of Social and Behavioral Sciences, Leiden University, P.O. Box 9555, 2300 RB Leiden, the Netherlands. E-mail: nunspeetvan@fsw.leidenuniv.nl

¹ Stereotyping, prejudice, and discrimination can be defined in different ways. Prejudice refers to evaluative and affective responses toward individuals based on their social group membership, whereas stereotypes imply schematic knowledge and conceptual attributes about a social group. Discrimination is often seen as the actual biased behavior displayed toward other group members. Although we acknowledge these differences—not only in terminology but also, for example, in their differences on the neural level (Amodio, 2014)—we use the terms in the current article as illustrating forms of social bias.

outgroup members (e.g., Amodio & Devine, 2006).

Increasing publicity for and knowledge of such effects has made some members of the general public as well as practitioners in education, the legal system, human resource management, and public policy quite weary, as it seems to suggest there is no way to avoid or reduce biased judgments. However, this conclusion is not warranted (e.g., Monteith, Woodcock, & Gulker, 2013). In the following, we review research demonstrating that people can control implicit bias and do so when they are induced to focus on the moral implications of their behavior. Furthermore, we describe how people reduce their implicitly biased responses by presenting evidence from neuroscientific research methods. Additionally, we explain how these insights can benefit members of the general public and practitioners who aim to avoid implicit biases, as a way to contribute to prejudice reduction.

Implicit Indicators of Prejudice

The traditional examination of prejudice and discrimination relies heavily on what people can and are willing to report about their attitudes. This means that people have to be aware of their own thoughts and attitudes and not reluctant to report them due to social desirability concerns (see also Greenwald et al., 2002). The concerns thus raised about the validity of such explicit indicators of prejudice endorsement prompted researchers to develop more implicit measures of social bias. A well-known example is the Implicit Association Test (IAT; Greenwald et al., 1998): a reaction time measure designed to preclude deliberative responses. The IAT is based on the notion that people will more easily associate subjectively related concepts than subjectively unrelated concepts. When considering an intergroup context, people may, for instance, be more inclined to associate members of a majority group (e.g., Whites) with positive attributes while more easily associating members from a minority group (e.g., Blacks) with negative attributes. In performance on the IAT, such tendencies result in relatively fast responses on prejudice-congruent IAT trials (e.g., where one response key is used to indicate majority group members and positive stimuli) and slower responses on prejudice-

incongruent trials (e.g., where one response key is used for minority group members and positive stimuli). The difference between response latencies on these prejudice-incongruent and prejudice-congruent trials signals ease of evaluative or stereotypical association and is used as an index of implicit bias.

As mentioned above, members of the general public and practitioners may assume that implicit associations that are apparent on implicit measures are triggered automatically and that it is thus impossible to alter biased behavior. However, although such implicit associations seem to be activated unintentionally and outside of one's awareness, this does not mean they cannot be controlled (Bargh, 1994). Indeed, prior research has revealed that implicit associations are malleable (e.g., Blair, 2002; Fiedler & Bluemke, 2005). That is, people can deliberately influence their behavioral responses during a task and thus also their performance on an implicit measure such as the IAT. This is the case, for instance, when they adopt a specific response strategy after they have been informed about how their bias will be measured (Fiedler & Bluemke, 2005). Likewise, IAT responses are modified when people are motivated to enhance their self-image or to emphasize their positive relationship with other individuals (e.g., Blair, 2002). Hence, even associations that seem to be activated automatically can be (re)adjusted or controlled when people are truly motivated to do so.

Moral Motivation and (Implicit) Prejudice Reduction

As indicated above, even implicit biases can be modified, for instance, due to self-concerns and social motives. Additionally, research has revealed that people who are highly internally motivated to respond without prejudice (i.e., who have strong personal egalitarian values) are less prone to implicit bias, especially when their external motivation to respond without prejudice (e.g., concerns about meeting social standards) is low (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002). Moreover, high internal motivation is associated with the intention to actually be free of prejudice (instead of just hiding it), which may thus be a more effective motivation to reduce (implicit) prejudice in the long run (Plant & Devine, 2009). In the current

contribution, we extend these insights by presenting a way in which this internal motivation may be enhanced: reminding people about the moral implications of their behavior.

Morality is often examined in group contexts as a guide of social and norm-based behavior. Specifically, group norms and values—especially when formulated in terms of morality—have an impact on individual and group-based actions. That is, people are more inclined to adhere to ingroup norms when these are presented as “the moral thing to do” rather than “the smart thing to do” (Ellemers, Pagliaro, Barreto, & Leach, 2008) and do so because they think this will yield respect from their fellow group members (Pagliaro, Ellemers, & Barreto, 2011). Moreover, people generally identify more strongly with moral than competent groups and are more proud to be member of groups, teams, and organizations that can contribute to their morality than groups that stand out for their competence (Leach, Ellemers, & Barreto, 2007; van Prooijen & Ellemers, 2015). Evaluating or presenting people’s identity or behavior in terms of moral values can thus enhance their moral intentions and may therefore be used as an instrument to enhance moral behavior toward others and to reduce social bias and prejudice.

However, previous research has shown that when triggering people’s moral motivation to be unbiased toward outgroups, it is crucial to consider how this is done. For example, Legault, Gutsell, and Inzlicht (2011) revealed that when people were primed with the obligation to be unprejudiced, they showed a greater implicit bias on the IAT compared to a no-prime condition. Furthermore, when (White) people are made aware that they may be seen by others as racist, they experience stereotype threat, which in turn causes an increase in implicit bias (but which can be regulated by self-affirmation procedures; Frantz, Cuddy, Burnett, Ray, & Hart, 2004; Rudman, Dohn, & Fairchild, 2007). Such self-concerns about appearing immoral or demonstrating failure to behave according to egalitarian ideals were also ineffective in reducing implicit bias in multiple studies described in Lai et al. (2014). This meta-analysis includes five interventions focused on “appeals to egalitarian values.” In three of these interventions, participants were primed with thoughts and feelings that they themselves or (White) people in gen-

eral cannot live up to egalitarian goals. All three interventions were shown to be ineffective at reducing implicit bias. In another intervention, participants were not asked to reflect upon their own or their group’s behaviors but to watch a video emphasizing that members of different groups all share a sense of humanity. Interestingly, only the final intervention related to egalitarian values was effective at reducing implicit bias. This was when participants were encouraged to adopt a multicultural perspective by having them read and summarize a text advocating multiculturalism and inviting them to come up with two reasons why this is a positive approach to interethnic relations. Importantly, this intervention thus differed from the others in that participants were encouraged to think about their personal stance toward interethnic relations, rather than being reminded of obligations or being confronted with past moral failures.

The effectiveness of such encouragement to consider one’s own reasons for endorsing egalitarian values was also found in other research—independent of the meta-analysis by Lai et al. (2014). Whereas implicit prejudice increased after an obligation prime, Legault et al. (2011) revealed that implicit bias was reduced when participants were encouraged to think about the personal satisfaction and enjoyableness of being unprejudiced. A similar pattern was demonstrated by Does, Derks, and Ellemers (2011), who directly compared the effects of framing social equality as an important moral ideal versus an obligation on people’s (explicit) willingness to strive toward this goal. Their findings indicated that participants were more inclined to strive toward social equality between groups when they were induced to think of this as an important moral ideal, rather than as a moral obligation. Moreover, a follow-up study provided insight into why this is the case: by examining participants’ physiological responses (using electrocardiographic signals), it was found that framing a goal in terms of moral ideals causes a physiological “challenge” response, which helps people to pursue this ideal, whereas the urge to meet an obligation causes a physiological “threat” response that hinders fulfillment of that goal (Does, Derks, Ellemers, & Scheepers, 2012).

Whether an emphasis on moral norms reduces or increases implicit prejudice may thus

depend on how such moral norms are framed and whether they cause a (psychological and/or physiological) threat or challenge response within individuals. Moreover, the meaning of morality may differ between individuals, groups, and situations, which makes it impossible to do what is right according to everyone. However, because people have a need to belong and seek inclusion in groups they identify with (Tajfel & Turner, 1979), they may prioritize behaving according to personal moral values that they share with members of their own (in)group. In a recent empirical research program, it was therefore examined whether presenting the implications of people's behavior in terms of morality helps to increase their (internal) motivation to reduce implicit prejudice (van Nunspeet, Ellemers, Derks, & Nieuwenhuis, 2014).

In this research, non-Muslim participants completed an IAT after they were informed about the possible implications of their performance. Specifically, it was emphasized that this task could give an indication of either their moral values concerning egalitarianism and discrimination or (as a control condition) their competence at learning new tasks and processing new information. Results revealed that participants in the moral values condition showed less implicit bias against Muslim women than participants in the competence condition. Importantly, no specific norms were made salient: participants only read that the test could give an indication of the value they attached to egalitarianism versus discrimination (which could thus be considered an opportunity to show one's own moral values). Furthermore, participants' associations were assessed while they performed the test in private and anonymously. The findings therefore suggest that people are internally motivated to act upon their own moral values, presumably because this is important for how they see themselves (van Nunspeet et al., 2014).

Additional research examined whether this (internal) motivation to behave according to one's own moral values is increased in an ingroup context (i.e., among people/group members with whom one identifies, shares similar values, and behaves according to the same moral group norms)—in other words, whether the above-mentioned effect of reduced implicit bias is enhanced when participants' behavior is

monitored by a member of their own group (van Nunspeet, Derks, Ellemers, & Nieuwenhuis, 2015).² To this end, the study design was extended: beside the two conditions in which either the test implications concerning one's morality or competence were emphasized, participants were also led to believe that their performance on the IAT was monitored by an in- or an outgroup member. Specifically, the evaluator was another (non-Muslim) individual, who was said to share the same minimal group membership as the participant or to hold another group membership. Findings revealed that participants' implicit bias is particularly reduced when the moral implications of their performance are emphasized and when they are being evaluated by another ingroup member (van Nunspeet et al., 2015). Interestingly, these groups were experimentally created using a minimal group paradigm (Tajfel, 1970), instead of relying on natural group memberships, excluding the possibility that alternative concerns (such as prior liking or familiarity) might induce participants' responses to different evaluators. Thus, this study revealed that introducing two groups, which had no meaning or known moral values outside of the laboratory, was sufficient to increase people's motivation to appear unbiased toward an alleged ingroup rather than an outgroup member. In other words, these findings suggest that evaluation by an ingroup member (which can be seen as representing an external source of social influence but from a person who is considered a part of people's social identity) may increase people's internal motivation to respond without prejudice, perhaps because of the importance people attach to sharing their personal moral values with members of groups that are relevant to the self.

Nevertheless, one could wonder how meaningful these results are for real-life situations and what would happen when people are faced with an actual interaction with a member of

² Please note that a social context (i.e., the presence of or evaluation by other individuals) can be seen to increase external (rather than internal) motivations to be moral and thus to respond without prejudice. However, we suggest that an ingroup context will enhance internal moral motivations because of the importance people attach to both their personal moral identity, as well as their social moral identity (i.e., being part of a moral group and a moral group member; Leach et al., 2007).

another group (e.g., a Muslim). Previous research has revealed that the presence of a member of the target group in the IAT (e.g., a Black experimenter or confederate) helps to reduce implicit bias (e.g., Lowery, Hardin, & Sinclair, 2001). Establishing such intergroup contact, however, may not always be feasible. Therefore, an additional study was conducted in which the effectiveness of an emphasis on one's own moral values (and introducing a minimal ingroup evaluator) was directly compared with the evaluation by a representative of the social target group. Consistent with previous findings, results showed that when participants thought that their IAT performance was monitored by a Muslim (compared to a non-Muslim) woman, this affected their behavioral responses on the IAT in such a way that their negative bias toward Muslims was reduced to nonsignificance. Moreover, an emphasis on the moral implications of participants' performance (or introducing the Muslim evaluator as a minimal in- or outgroup member) did not yield any (additional) effects. Interestingly, however, when the evaluator was not Muslim, a similar (although somewhat weaker) bias reduction was evident when the moral implications of participants' performance for the self were emphasized and when this emphasis was combined with presenting the non-Muslim evaluator as a minimal ingroup member (van Nunspeet, Ellemers, & Derks, 2015). These findings thus show that implicit prejudice can be reduced when people are confronted with evaluations by others who depend on them for equal treatment (e.g., Muslims). Importantly, however, even when it is not feasible to create such intergroup interactions, signaling the moral implications of one's behavior for the self may also instigate the desired responses. In addition, the effectiveness of emphasizing the moral implications of one's behavior can also be enhanced when in an ingroup context.

Comparing the Effectiveness of Different Interventions

To our knowledge, no other research than the study described above has directly compared the effectiveness of introducing a social context (evaluation by another individual) with more cognitive strategies or personal motivations (activated by the emphasis on moral implications)

on the reduction of implicit prejudice. In part, this may be related to the fact that the ease of implementation in real-life situations differs quite a lot between these interventions (i.e., creating intergroup contact vs. activating an internal motivation or cognitive strategy). When having to decide which of different possible interventions to apply, it is thus highly relevant to consider the effectiveness of these interventions, in relation to the ease of implementing them.

Lai et al. (2014) found that interventions that included counterstereotypical exemplars or strategies to inhibit or overcome implicit prejudice (such as implementation intentions) were highly effective ways to reduce implicit bias, with Cohen's *d*s between .30 and .55. However, as with creating actual intergroup contact, even such (cognitive) interventions may not be that easy to implement: they require effortful training or repeated exposure to different exemplars and strategies (and, in some instances, not only an increase in positive evaluations of outgroup members but also a decrease in positive evaluation of ingroup members, which may be undesirable). Thus, the general applicability of such strategies is limited, and the complexity of their implementation may undermine their effectiveness in practice. In contrast, interventions that aim at inspiring people to show their personal moral values—such as the ones reviewed previously—also seem to be quite effective (partial η^2 between .05 and .18³). But at the same time, such interventions are relatively easy to implement in everyday life and in company policies because they only require mentioning or putting an emphasis on the moral implications of one's behavior before an action is initiated. This may make them a favored method to reduce implicit bias. In fact, this may also be used as an ancillary strategy to support the effectiveness of more elaborate and complex bias reduction programs. For instance, the ability of people to learn new strategies (such as implementation intentions) relies on their willingness to invest effort in such a training. Emphasizing the moral implications of their performance may offer a way to enhance or ensure such motivation.

³ Based on Does et al. (2011), Legault et al. (2011), and van Nunspeet et al. (2014, 2015).

Underlying Cognitive Processes

Cognitive processes associated with the motivation to be moral during IAT performance have also been studied at a neural level. Examining such processes may help us to understand, for instance, whether people control or readjust prejudiced behavioral responses on a task or whether they approach such a task differently to initiate their responses in another (unprejudiced) way. Examination of these unconscious cognitive processes using neuroscientific research methods such as event-related brain potentials (ERPs, which can reveal what and how quickly cognitive processes are being engaged during task performance) may thus help us to gain a better understanding of how (implicit) prejudice can be reduced.

van Nunspeet et al. (2014, 2015) examined these issues using ERPs derived from an electroencephalogram collected while participants performed the IAT. ERPs are peaks or fluctuations in electrical activity in the brain occurring at specific points in time (i.e., after stimulus presentation or during response generation) and are thought to be associated with particular cognitive processes. By comparing the amplitude of these peaks in response to different stimuli, it is thus possible to examine which cognitive processes play a role during IAT performance. Research findings revealed that the moral motivation to reduce implicit bias is associated with the operation of multiple processes: enhanced response monitoring as well as increased perceptual attention to group-related cues.

Response and error monitoring have been associated with the error-related negativity (ERN) modulation. Measured within 100 ms after a response is given, the ERN appears particularly when participants give incorrect (compared to correct) responses on a task (e.g., Gehring, Goss, Coles, Meyer, & Donchin, 1993). Moreover, research has revealed that increases in the ERN are associated with the extent to which people subjectively worry about making errors (Hajcak, Moser, Yeung, & Simons, 2005). In the research by van Nunspeet et al. (2014, 2015), participants showed enhanced error monitoring (i.e., greater ERN modulations to incorrect responses) when the moral (rather than the competence) implications of the IAT were emphasized. This may thus suggest that people are more concerned about making mis-

takes when the task supposedly indicates their moral values than when it “merely” indicates their competence. Furthermore, the emphasis on moral implications in combination with being evaluated by an ingroup member increased overall response monitoring (i.e., for incorrect and correct responses; van Nunspeet et al., 2015), which may indicate that under these circumstances, participants may be more attentive to both avoiding responses that reveal bias as well as demonstrating their ability to show unbiased behavioral responses.

The emphasis on the moral implications of the task—and being monitored by an ingroup member—also affected the perceptual attention toward and categorization of ingroup and outgroup faces. The process of social categorization is associated with the N1 and P2 modulations (occurring around respectively 100 or 200 ms after stimulus-onset) when viewing ingroup versus outgroup faces (e.g., Ito & Urland, 2003). Interestingly, in the research of van Nunspeet et al. (2014, 2015), these modulations—associated with early attention to stimulus information—were enhanced in the morality compared to the competence condition. Participants thus seemed to attend more to the difference between ingroup and outgroup members when the moral implications of their task performance were emphasized. In addition, this increased perceptual attention emerged only when participants were evaluated by a (minimal) ingroup rather than an outgroup member.

At first glance, the behavioral responses on the task and the neuroscientific findings described above may seem to be contradictory. Emphasizing morality resulted in more equal responses toward ingroup and outgroup faces, whereas it increased differentiation (i.e., stimulus categorization) between these groups at the neural level. However, when considering the differences in assessment of both measures, we can see that the salience of morality increased participants’ perceptual attention toward the faces of in- and outgroup members, suggesting that this is what enabled them to behaviorally respond with decreased bias. In addition, the examination of the underlying cognitive processes revealed that interventions aimed at reducing prejudice may thus be based on two different kind of processes: initiating early cognitive processes that are crucial for preparing positive responses, as well as controlling or

redirecting negative responses. Furthermore, extending previous research that examined effects strategies aiming to activate (one of) these different processes (such as “if-then” implementation intentions to activate top-down control; e.g., *Mendoza, Gollwitzer, & Amodio, 2010*), the emphasis on the moral implications of people’s behavior was found to affect different processes simultaneously.

Societal Implications of the Research Findings

In line with the vast majority of the research on implicit prejudice (*Lai, Hoffman, & Nosek, 2013*), the interventions related to moral motivations reviewed here have only been examined as a method to reduce implicit bias. They do not show whether emphasizing the moral implications of one’s performance is also effective in changing actual behavior (e.g., reducing discrimination, improving intergroup interactions) and whether such effects can be long-lasting. Nevertheless, other research suggests that there should be a relation between implicit response patterns and more explicit behaviors. For instance, it has been demonstrated that priming people with equality decreased their seating distance from a Muslim individual (assessed immediately after task performance; *Zogmaister, Arcuri, Castelli, & Smith, 2008*). Furthermore, *Devine, Forscher, Austin, and Cox (2012)* conducted a longitudinal study in which they showed that a multifaceted intervention caused implicit bias reduction that lasted over a 12-week period. Importantly, part of this intervention was focused on people’s awareness of and concerns about the effects of bias. This is thus in line with our notion that emphasizing the (moral) implications of one’s behavior will help to reduce implicit prejudice, and this is a strategy that may be relatively easy to implement in everyday life. Although more research is needed to further establish such effects, we thus argue that it is worthwhile to consider the possible practical implications this may have in future work.

Emphasizing moral implications of people’s behavior has been shown to cause people to inhibit their (implicit) behavioral prejudice toward outgroup targets. This suggests that in real-life settings, people may also adjust their behavioral or verbal expressions of prejudice

when they are made aware of what such expressions might say about their own moral values. For example, an employer may reject an applicant merely because she would like to wear a headscarf at work. In this situation, the employer will probably be aware of the consequences for the applicant while failing to consider what the rejection of such applicants may reveal about himself and the company more generally. Increased awareness of the implications of such behavior in terms of one’s own morality and what this says about personal or company values regarding equality and intercultural respect may make people more careful to ensure equal treatment. In other words, moral motivation and related bias reduction may be enhanced by reminding people that their behavior displays their moral intentions and values. This is something that can be done in different ways, for instance, in the company policy or during interpersonal business meetings.

Although one could suggest that this may also be applied in public ads or campaigns, the additional research revealed that the social context in which moral implications of one’s behavior are emphasized also affects (implicit) prejudice. That is, people especially care about complying with moral ingroup norms and thus about being perceived as moral by their fellow group members. In contrast to being reminded about the moral implications of their actions by an agency with which people do not identify, being monitored by self-relevant others may thus enhance people’s internal motivation to control prejudice. In other words, inspiring people to inhibit their prejudice by relating this to their moral values (instead of their ability of doing so) will thus be particularly effective in an ingroup context.

This finding speaks to debates about how to best promote diversity policies in work settings. Previously, it has been suggested that organizations should motivate people to change their biased behavior by emphasizing the “business case for diversity.” Thus, a focus on competence tends to be recommended because this is thought to persuade the executive board of a company to work toward a more ethnically and gender diverse organization, as a way to improve the organization’s profit and success. Indeed, the “business case for diversity” tends to be promoted as a way to increase employee motivation and business performance (e.g., *Eu-*

ropean Commission, 2005; Robinson & Dechant, 1997). However, based on the research findings presented here, the effectiveness of such measures may be limited. Instead, a more effective way to achieve that employees embrace diversity and treat colleagues from other ethnic backgrounds with respect might be to emphasize that striving toward a diverse organization is the moral thing to do. This way, a company can disseminate its moral norms, which in turn may motivate its employees to behave according to these norms because it will enhance their motivation to be (perceived as) a moral individual and company member.

Nevertheless, the current findings also show that moral behavior (in terms of prejudice reduction) can be affected when one's actions are evaluated by an outgroup member. That is, people will generally be inclined to inhibit the expression of negative bias when they are being monitored by a member of the group that is the target of such bias. This implies that increasing diversity—to create a setting where people are aware that their actions are monitored by representatives of other groups—may prevent displays of prejudice and discrimination. For example, having a Muslim employee as a member of an evaluation committee (to monitor the decision-making process concerning candidates for the job) may help the committee to create equal opportunities for Muslim as well as non-Muslim applicants. Likewise, having women present in the board of directors of a company may help others control gender bias when considering applications for leadership positions.

Importantly, this recent research revealed that emphasizing the moral implications of people's behavior when they are alone can be just as effective in reducing prejudice as is the presence of a representative of the group that is the target of prejudice. This is important for contexts in which intergroup contact is not feasible. For instance, when employees do not yet have any colleagues with a different ethnic background or religion, it might encourage them to be more welcoming to such a colleague. Likewise, emphasizing moral implications of being unbiased might benefit the integration of newcomers in neighborhoods that primarily consist of people from the same social class or ethnicity. In social contexts such as these, people may be motivated to control their prejudice when this is emphasized as the moral thing to do,

which will give the new colleague or neighbor a fair chance to reveal his or her personal qualities rather than relying on biased expectations.

Standard communications regarding company policy or national campaigns to induce equal treatment of minority members tend to focus on the negative implications for the targets of prejudice, as a way to prevent people from expressing bias. The present research suggests that there is likely to be added value in communicating about moral implications of and adherence to equality goals of the perpetrators, as a way to help diminish prejudice.

References

- Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. *Nature Reviews Neuroscience*, *15*, 670–682. <http://dx.doi.org/10.1038/nrn3800>
- Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology*, *91*, 652–661. <http://dx.doi.org/10.1037/0022-3514.91.4.652>
- Bargh, J. A. (1994). The Four Horsemen of automaticity. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 1–40). Hillsdale, NJ: Erlbaum.
- Barreto, M., & Ellemers, N. (2013). Sexism in contemporary societies: How it is expressed, perceived, confirmed, and resisted. In M. K. Ryan & N. R. Branscombe (Eds.), *The Sage handbook of gender and psychology: Pt. 3. Gender: Social context, conflict, and coping* (pp. 289–305). London, UK: Sage. <http://dx.doi.org/10.4135/9781446269930.n18>
- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, *6*, 242–261. http://dx.doi.org/10.1207/S15327957PSPR0603_8
- Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, *48*, 1267–1278. <http://dx.doi.org/10.1016/j.jesp.2012.06.003>
- 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*, 835–848. <http://dx.doi.org/10.1037/0022-3514.82.5.835>
- Does, S., Derks, B., & Ellemers, N. (2011). Thou shalt not discriminate: How emphasizing moral

- ideals rather than obligations increases Whites' support for social equality. *Journal of Experimental Social Psychology*, 47, 562–571. <http://dx.doi.org/10.1016/j.jesp.2010.12.024>
- Does, S., Derks, B., Ellemers, N., & Scheepers, D. (2012). At the heart of egalitarianism: How morality framing shapes cardiovascular challenge versus threat in Whites. *Social Psychological and Personality Science*, 3, 747–753. <http://dx.doi.org/10.1177/1948550612438924>
- Ellemers, N., Pagliaro, S., Barreto, M., & Leach, C. W. (2008). Is it better to be moral than smart? The effects of morality and competence norms on the decision to work at group status improvement. *Journal of Personality and Social Psychology*, 95, 1397–1410. <http://dx.doi.org/10.1037/a0012628>
- European Commission. (2005). *The business case for diversity: Good practices in the workplace*. Retrieved from <http://ec.europa.eu/social/main.jsp?catId=370&featuresId=25&langId=nl>
- Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the implicit association tests. *Basic and Applied Social Psychology*, 27, 307–316. http://dx.doi.org/10.1207/s15324834basp2704_3
- Frantz, C. M., Cuddy, A. J. C., Burnett, M., Ray, H., & Hart, A. (2004). A threat in the computer: The race implicit association test as a stereotype threat experience. *Personality and Social Psychology Bulletin*, 30, 1611–1624. <http://dx.doi.org/10.1177/0146167204266650>
- Gaertner, S. L., & Dovidio, J. F. (1986). *The aversive form of racism*. New York, NY: Academic Press.
- Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, 4, 385–390. <http://dx.doi.org/10.1111/j.1467-9280.1993.tb00586.x>
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4–27. <http://dx.doi.org/10.1037/0033-295X.102.1.4>
- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, 109, 3–25. <http://dx.doi.org/10.1037/0033-295X.109.1.3>
- 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, 1464–1480. <http://dx.doi.org/10.1037/0022-3514.74.6.1464>
- Hajcak, G., Moser, J. S., Yeung, N., & Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology*, 42, 151–160. <http://dx.doi.org/10.1111/j.1469-8986.2005.00270.x>
- Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electro-cortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, 85, 616–626. <http://dx.doi.org/10.1037/0022-3514.85.4.616>
- Lai, C. K., Hoffman, K. M., & Nosek, B. A. (2013). Reducing implicit prejudice. *Social and Personality Psychology Compass*, 7, 315–330. <http://dx.doi.org/10.1111/spc3.12023>
- Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. E. L., Joy-Gaba, J. A., . . . Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143, 1765–1785. <http://dx.doi.org/10.1037/a0036260>
- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, 93, 234–249. <http://dx.doi.org/10.1037/0022-3514.93.2.234>
- Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of antiprejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science*, 22, 1472–1477. <http://dx.doi.org/10.1177/0956797611427918>
- Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology*, 81, 842–855. <http://dx.doi.org/10.1037/0022-3514.81.5.842>
- Mendoza, S. A., Gollwitzer, P. M., & Amodio, D. M. (2010). Reducing the expression of implicit stereotypes: Reflexive control through implementation intentions. *Personality and Social Psychology Bulletin*, 36, 512–523. <http://dx.doi.org/10.1177/0146167210362789>
- Monteith, M. J., Woodcock, A., & Gulker, J. E. (2013). Automaticity and control in stereotyping and prejudice: The revolutionary role of social cognition across three decades. In D. Carlston (Ed.), *Oxford handbook of social cognition* (pp. 74–94). New York, NY: Oxford University Press. <http://dx.doi.org/10.1093/oxfordhob/9780199730018.013.0005>
- Pagliaro, S., Ellemers, N., & Barreto, M. (2011). Sharing moral values: Anticipated ingroup respect as a determinant of adherence to morality-based (but not competence-based) group norms. *Personality and Social Psychology Bulletin*, 37, 1117–1129. <http://dx.doi.org/10.1177/0146167211406906>
- Plant, E. A., & Devine, P. G. (2009). The active control of prejudice: Unpacking the intentions guiding control efforts. *Journal of Personality and Social Psychology*, 96, 640–652. <http://dx.doi.org/10.1037/a0012960>

- Robinson, G., & Dechant, K. (1997). Building a business case for diversity. *The Academy of Management Executive*, 11, 21–31. <http://dx.doi.org/10.5465/AME.1997.9709231661>
- Rudman, L. A., Dohn, M. C., & Fairchild, K. (2007). Implicit self-esteem compensation: Automatic threat defense. *Journal of Personality and Social Psychology*, 93, 798–813. <http://dx.doi.org/10.1037/0022-3514.93.5.798>
- Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223, 96–102. <http://dx.doi.org/10.1038/scientificamerican1170-96>
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.
- van Nunspeet, F., Derks, B., Ellemers, N., & Nieuwenhuis, S. (2015). Moral impression management: Evaluation by an ingroup member during a moral IAT affects perceptual attention and conflict- and response monitoring. *Social Psychological and Personality Science*, 6, 183–192. <http://dx.doi.org/10.1177/1948550614548076>
- van Nunspeet, F., Ellemers, N., & Derks, B. (2015). *Reducing implicit prejudice against Muslim women: The effects of moral concerns, intra- and intergroup motives*. Unpublished manuscript.
- van Nunspeet, F., Ellemers, N., Derks, B., & Nieuwenhuis, S. (2014). Moral concerns increase attention and response monitoring during IAT performance: ERP evidence. *Social Cognitive and Affective Neuroscience*, 9, 141–149. <http://dx.doi.org/10.1093/scan/nss118>
- van Prooijen, A. M., & Ellemers, N. (2015). Does it pay to be moral? How indicators of morality and competence enhance organizational and work team attractiveness. *British Journal of Management*, 26, 225–236. <http://dx.doi.org/10.1111/1467-8551.12055>
- Zogmaister, C., Arcuri, L., Castelli, L., & Smith, E. R. (2008). The impact of loyalty and equality on implicit ingroup favoritism. *Group Processes & Intergroup Relations*, 11, 493–512. <http://dx.doi.org/10.1177/1368430208095402>

Received October 31, 2014

Revision received August 20, 2015

Accepted September 13, 2015 ■