

The Oxford Handbook of Social Cognition, Second Edition (2nd edn)

Donal E. Carlston (ed.) et al.

<https://doi.org/10.1093/oxfordhb/9780197763414.001.0001>

Published: 2024

Online ISBN: 9780197763445

Print ISBN: 9780197763414

Search in this book

CHAPTER

20 Social Cognition and the Reduction of Intergroup Bias

Laura K. Hildebrand, Margo J. Monteith, Elisabeth S. Noland

<https://doi.org/10.1093/oxfordhb/9780197763414.013.20> Pages 560–588

Published: 21 August 2024

Abstract

This chapter provides an overview of social cognitive bias-reduction strategies. These strategies are organized around three broad categories. The strategies in the first category, which includes personalization- and recategorization-based strategies, reduce bias by changing group-based mental representation. The second category, which includes counterstereotyping and evaluative conditioning, “trains away bias” by changing the traits and affect that are automatically activated by target group members. The final category, which includes the self-regulation of prejudice model and the prejudice habit-breaking intervention, considers how one can exert control over bias to override automatically activated, biased responses with controlled, nonbiased responses. This chapter concludes by discussing the need to complement social cognitive bias-reduction strategies with other approaches to tackle contemporary issues (e.g., individual- versus systemic-level bias) involving prejudice, stereotyping, and discrimination.

Keywords: [stereotyping](#), [prejudice](#), [social cognition](#), [bias reduction](#), [self-regulation](#), [malleability](#), [automaticity](#), [control](#), [categorization](#)

Subject: [Cognitive Psychology](#), [Social Psychology](#), [Psychology](#)

Series: [Oxford Library of Psychology](#)

Collection: [Oxford Handbooks Online](#)

Social cognition involves studying how people think about people (Wegner & Vallacher, 1977), and the study of intergroup bias reduction focuses on reducing (often negative) evaluations of people that are based on their group membership. Thus, social cognitive approaches to bias reduction may be defined as how we change what people think about others based on their group memberships. While addressing intergroup bias reduction in this chapter, we will be concerned with the reduction of stereotyping (ascription of a trait to its group members), prejudice (group-based evaluation), and discriminatory responses (judgments and behaviors linked to people’s group memberships). For instance, what strategies can lead people to change

their belief that women make poor leaders, their feeling that gay people are bad, or their more positive evaluations of White versus Black job applicants?

p. 561

Consistent with the social cognition approach (Fiske & Taylor, 1991; Hamilton & Carlston, 2013), much intergroup bias-reduction research focuses on mental processes. Moreover, because the social cognition approach revealed new ways of conceptualizing and understanding prejudice (Monteith et al., 2013), it greatly influenced the types of biases that researchers attempt to change. For instance, social cognitive theories on cognitive structure gave prejudice researchers new ways of conceptualizing mental representations of groups (Carlston, 2010; Wyer & Carlston, 1994). These cognitive structure theories then provided the foundation of several strategies aimed at changing group-based mental representations so as to reduce intergroup bias. In addition, the discovery that intergroup bias can be automatically activated (i.e., with little conscious input, intention, or control; Bargh, 1994) led researchers to study strategies for reducing this form of bias. This discovery in turn led to the finding that, even when automatically activated, biased responses can be overridden by controlled, nonprejudiced responses (Devine, 1989). With this finding, researchers turned to studying how one can exert control over biased responses.

We organize the present chapter around three broad categories of bias reduction: changing intergroup categorization, changing stereotypic and affective responses to target members, and exerting control over bias. It is important to note that stereotyping, prejudice, and discrimination can be manifested as implicit bias or explicit bias, depending on whether more automatic or controlled processing is involved in their production. For instance, implicit bias presumably is operative if a stereotypic inference such as “That woman must be the physician’s assistant” (rather than the physician) is generated with little conscious awareness or intention. In contrast, the same inference resulting from the conscious and intentional use of stereotypes constitutes explicit bias. Although our chapter is not organized around these different forms of bias, we highlight whether strategies are primarily aimed at reducing implicit bias, explicit bias, or both. Finally, we conclude the chapter by considering some limitations of social cognitive bias-reduction strategies and the necessity of complementing them with other approaches.

Changing Intergroup Categorization

The strategies described in this first section are direct descendants of Muzafer Sherif and colleagues’ (1961) classic field research conducted among boys attending a summer camp near Robbers Cave State Park in Oklahoma. In this research, dubbed the “Robbers Cave study,” Sherif and colleagues found that when boys who were randomly assigned to one of two groups within the camp engaged in competitive activities, they quickly developed stereotypes and intergroup animosity. They often engaged in open conflict, calling each other derogatory names and fighting. However, when the researchers introduced tasks that required the boys to cooperate to achieve superordinate goals (e.g., fixing the camp’s water supply), the boys became amicable and the intergroup conflict was abated. This classic study stimulated decades of research examining categorization-based processes in relation to intergroup bias (Gaertner et al., 2000).

How we feel about other people is closely tied to how we feel about ourselves (McConnell et al., 2013). We typically view in-group members, or people with whom we share a social category, positively and bestow them with tangible and intangible benefits (i.e., in-group favoritism; Tajfel et al., 1971; Tajfel & Turner, 1979). In contrast, out-group members, or people who are not part of our social category, do not receive such benefits. This differential treatment of in-group and out-group members can lead to prejudice and discrimination, even when one is not actively hostile toward the out-group (Brewer, 1999; Greenwald & Pettigrew, 2014; Nier & Gaertner, 2012).

Research suggests that differential treatment of in-groups and out-groups stems from an engrained process that results from the human need to categorize (Allport, 1954; Tajfel et al., 1971; Tajfel & Turner,

1979). Use of the minimal group paradigm is especially helpful for illustrating this engrained process (Tajfel et al., 1971). In this paradigm, researchers assign participants to groups based on arbitrary criteria, so that whether someone is in one's in-group or an out-group carries little meaning. Even with such arbitrary group distinctions, in-groups end up being treated more favorably than out-groups. Yet, whether we categorize others as in-group or out-group members is easily swayed by a variety of factors (e.g., Brewer & Miller, 1984). In one situation we may consider a person to be part of the out-group, while in another situation we may consider that same person to be part of our in-group. The strategies described in this section leverage this malleability to change the mental representation of out-group members. Personalization-based strategies break down the categorization, so that the out-group member is seen as an individual, whereas recategorization-based strategies reshape group boundaries so that the out-group member becomes an in-group member. Finally, indirect contact draws on the cognitive underpinnings of classic contact work to reshape mental representations when actual contact is not possible or desirable.

Personalization

Personalization builds on a fundamental component of the contact hypothesis, which is that interpersonal interaction with out-group members facilitates the perception of out-group members as individuals with unique characteristics, rather than simply as out-group members (e.g., Brewer & Miller, 1984). Personalization can include a number of different components that operate independently or in tandem (Ensari et al., 2012). The first, most basic component is individuation, or when participants are exposed to unique traits or attributes of an outgroup member (Ensari et al., 2012; Miller, 2002). In a typical individuation experiment, participants meet an out-group member, who shares individuating information about themselves that is unrelated to group membership, such as a piece of happy news they received in the past year (e.g., Ensari & Miller, 2005). This individuating information reduces out-group homogeneity and in-group favoritism (Brewer et al., 1981; Ensari et al., 2012; Miller, 2002; Wilder, 1978), decreases attributional bias in explaining an out-group job applicant's success (Ensari & Miller, 2005), increases out-group acceptance (Miller et al., 1985), and increases the favorability of self-reported out-group attitudes (Ensari & Miller, 2001; Ensari et al., 2012).

A second component of personalization is self-disclosure. Sharing and listening to personal information during self-disclosure increases trust, interpersonal liking, and friendliness toward the out-group member (Ensari et al., 2012; Ensari & Miller, 2001, 2002). It also reduces stereotypes, prejudice, and anxiety about interacting with out-group members (Ensari et al., 2012; Turner, Hewstone, & Voci, 2007).

A third component of personalization is self-other comparison, where people connect individuating information about an out-group member with the self. For instance, in one study that encouraged self-other comparison, a Republican student profile was presented next to the liberal participant's own profile, thus encouraging the identification of self-other similarities and differences. In contrast, the individuation-only condition simply presented the Republican student profile. Results showed that self-other comparison increased perceived similarity between the out-group and the in-group, which in turn reduced recall of negative memories involving the out-group (Ensari et al., 2012).

Self-other comparison may also promote empathy, the fourth component of personalization. Empathy, which is often evoked by perspective taking, decreases stereotyping and intergroup bias (Galinsky & Moskowitz, 2000; Vescio et al., 2003; but see Vorauer & Sasaki, 2009). Although perspective-taking paradigms can take a number of forms (e.g., watching a video, reading a vignette; Dovidio et al., 2004; Finlay & Stephan, 2000), a common version involves introducing participants to their study partner (an out-group member) and then instructing them to imagine how they would feel if they were in their study partner's position (Todd et al., 2011; Vorauer et al., 2009; Vorauer & Sasaki, 2014). Compared to control conditions, participants who took their partner's perspective showed less explicit (e.g., self-reported

negative attitudes; Batson et al., 2002; Dovidio et al., 2004; Galinsky & Ku, 2004; Galinsky & Moskowitz, 2000; for a review, see Todd & Galinsky, 2014) and implicit bias (e.g., as assessed with the Implicit Association Test [IAT] and the Affect Misattribution Procedure; ↵ Greenwald et al., 1998; Payne et al., 2005; Todd et al., 2011; Todd & Burgmer, 2013; for a review, see Todd & Galinsky, 2014). Perspective taking also increases positive behaviors, such as helping out-group members (Dovidio et al., 1997; Mallett et al., 2008). Finally, it led White participants to engage in more approach-oriented nonverbal behaviors during an interracial interaction (e.g., leaning toward versus away from one's partner), which improved their Black partner's ratings of the interaction (Todd et al., 2011).

A potential pitfall of the personalization components discussed thus far is that positive effects may not generalize to other out-group members (Brewer & Miller, 1984; Miller, 2002). The participant does not cognitively connect the individual out-group member to the out-group as a whole; as a result, attitudes toward the general out-group may remain unchanged (Ensari & Miller, 2002; Hewstone & Brown, 1986). The fifth potential component of personalization, called decategorization, makes generalization to the out-group more likely (Brewer & Miller, 1984). Decategorization involves considering individuating information about out-group members while differentiating them from the out-group stereotype. For instance, a confederate mentioned when interacting with a liberal participant that they belong to a Republican student club (Ensari & Miller, 2006). This detail made out-group membership salient and activated cognitive representations about the out-group (Brewer & Miller, 1984; Ensari et al., 2012). The confederate then shared personal, individuating information with the participant that undermined the validity of out-group stereotypes (Brewer & Miller, 1984; Ensari et al., 2012). Consequently, by differentiating between out-group membership and individuating information, decategorization changes attitudes not only toward certain out-group members but also toward the out-group as a whole (Brewer & Miller, 1984; Ensari & Miller, 2001, 2002, 2006; Miller, 2002). Decategorization also reduces perceived out-group homogeneity (Miller, 2002) and the number of negative traits people remember about the out-group (Ensari et al., 2012). Additionally, it increases friendliness toward the out-group (Ensari & Miller, 2005) and perceived similarity between the self and the out-group (Ensari et al., 2012).

Personalization can be achieved not only through direct contact with an out-group member, but also through indirect contact. Imagined contact is by far the most well-studied indirect contact strategy (though see also vicarious, or extended, contact; Wright et al., 1997). In a typical imagined contact paradigm, participants are asked to imagine a positive interaction with an out-group member (Crisp & Turner, 2009). Importantly, the imagined contact must involve an interaction (versus just thinking about an out-group member; Turner, Crisp, & Lambert, 2007) and must have a positive, or promotion-oriented tone (versus a negative, prevention-focused, or even neutral tone; Stathi & Crisp, 2008; West & Greenland, 2016). Although other factors may enhance the effect (e.g., more detailed mental imagery, a third- versus first-person perspective; Husnu & Crisp, 2010), they are not necessary. Simply imagining a positive interaction with an out-group member reduces bias toward the out-group (Crisp & Turner, 2009).

Impressively, the positive effects of imagined contact have been demonstrated across a wide variety of explicit and implicit bias measures: Imagined contact improves self-reported attitudes toward the out-group (Turner et al., 2007), increases the number of positive traits perceived for the out-group (Stathi & Crisp, 2008), reduces out-group homogeneity (Turner et al., 2007), reduces bias on the IAT (Turner & Crisp, 2010), decreases dehumanization (Vezzali et al., 2012), reduces endorsement of negative stereotypes (Stathi et al., 2012), and increases intentions for future contact (Husnu & Crisp, 2010). Beyond attitudes and behavioral intentions, imagined contact reduces biased behavior: After imagining themselves interacting with an out-group member, participants sat closer to an out-group member (Turner & ↵ West, 2012) and cooperated more with an out-group member during a prisoner's dilemma game (Meleady & Seger, 2017). Imagined contact primarily works by reducing anxiety (Husnu & Crisp, 2010; Stathi et al., 2012; Turner, Crisp, & Lambert, 2007) and increasing trust (Meleady & Seger, 2017; Pagotto et al., 2012; Vezzali et al.,

2012) toward the out-group. Furthermore, these positive effects have held across a wide variety of out-groups, including elderly people (e.g., Turner, Crisp, & Lambert, 2007), Muslim people (Husnu & Crisp, 2010), Indians (Meleady & Seger, 2017), obese people (e.g., Turner & West, 2012), and gay men (e.g., Turner, Crisp, & Lambert, 2007). A meta-analysis revealed that imagined contact has a small-to-medium effect size ($d = 0.35$) and works equally well for explicit and implicit measures of bias (Miles & Crisp, 2014).

Although imagined contact does not always successfully reduce prejudice (e.g., Klein et al., 2014), overall the literature paints a picture of a robust prejudice-reduction strategy. Direct contact experiences (including those involving personalization and recategorization strategies) are likely to have stronger effects. However, researchers suggest that imagined contact can be used as a strategic “warm-up” to direct contact (Crisp et al., 2010; Crisp & Turner, 2009) or when direct contact is not possible (e.g., because of segregation; Crisp & Turner, 2009). We add that in some cases imagined contact may be better than actual contact, where subordinate or minoritized group members are more likely to experience microaggressions or overt bias, or when they may be pressured to “educate” majority group members about bias (e.g., Feagin, 1991; Meyers et al., 2020). Indeed, Tropp and Pettigrew’s (2005) meta-analysis indicated that the relation between intergroup contact and prejudice is weaker for minority than for majority group members. However, we also note that much of the imagined contact literature has focused on its various positive effects on attitudes and beliefs; more work is needed to examine both the underlying processes and whether imagined contact actually creates more positive intergroup contact outcomes down the road, including for minority group individuals.

Recategorization

Whereas the personalization strategies outlined above reduce bias by breaking down category lines, recategorization-based strategies reduce bias by redrawing category lines. The most common approach, based on the common in-group identity model (CIIM), involves creating a shared superordinate group (Gaertner et al., 1989, 1993, 2000). The CIIM works by restructuring mental representations from “us” versus “them” to “we,” so that in-group favoritism is extended toward former out-group members (Dovidio et al., 1997, 2009). For instance, a CIIM-based intervention would re-engineer an interracial interaction to highlight a shared identity, such as nationality or university membership.

Recategorization consistently and reliably reduces bias. In one early study, Gaertner and colleagues (1989) initially assigned participants to an in-group or out-group based on arbitrary distinctions, or participants were not assigned to a group at all (i.e., individual condition). Each group was taken to a separate room, decided on a name for their group, and completed a problem-solving task with their group. Participants in the individual condition completed the task alone and in separate rooms. Next, for participants who had been assigned to groups, the two groups were brought to one room to work on the problem-solving task again. Critically, although everyone participated in the task together, either the original two-group distinction was maintained (e.g., having participants sit together with their in-group and separate from the out-group; groups kept their original names) or steps were taken to recategorize the two groups into one inclusive ingroup (e.g., with arranged seating that intermixed participants in the two groups; a new name was generated for the superordinate group). In the individual condition, participants just worked on the problem together in the same room but only with their individual identities salient. In line with the CIIM, participants who were recategorized into one inclusive in-group reported more positive evaluations of other participants than participants in either the two-group or individual conditions.

Other studies have investigated whether intergroup cooperation can cause people to view former out-group members as in-group members. Participants who worked with out-group members on a common problem evaluated these individuals more positively and exhibited less biased nonverbal facial expressions (Gaertner

et al., 1990, 1999). Creating a common in-group identity via cooperation also increases helping (Dovidio et al., 1997), self-disclosure (Dovidio et al., 1997), and intergroup forgiveness (Wohl & Branscombe, 2005).

Yet recategorization may backfire and increase bias among people who are highly identified with their subgroup (Crisp et al., 2006; Hornsey & Hogg, 2000). This occurs because recategorization may minimize existing subgroup identities, which threatens the distinctiveness provided by those identities (Crisp et al., 2006; Hewstone, 1996; Hewstone & Brown, 1986; Hornsey & Hogg, 2000). For instance, a British Muslim person whose national (i.e., superordinate) identity is emphasized may feel that their religious (i.e., subgroup) identity is not being recognized. The mutual intergroup differentiation model (Hewstone, 1996; Hewstone & Brown, 1986) emphasizes that one can maintain distinctiveness and reduce bias by simultaneously emphasizing both the subgroup identity (e.g., people's religion) and a superordinate identity (e.g., a shared national identity). Empirically, this idea has typically been examined by manipulating whether the subgroups, the superordinate group, or all groups are made salient to participants (e.g., Hornsey & Hogg, 2000). Participants for whom both the subgroup and the superordinate group were made salient (i.e., the mutual differentiation condition) reported more positive explicit attitudes toward former out-group members (Crisp et al., 2001, 2006; Dovidio et al., 1998; Gaertner et al., 1994; Hornsey & Hogg, 2000). Reduced threat to distinctiveness (Ioannou et al., 2017) and more inclusive group representation (Dovidio et al., 1998) mediated this effect. Importantly, both the subgroup and the subordinate group must be equally valued for this effect to succeed, and the emphasized group differences and similarities should be positive, rather than negative (Hewstone & Brown, 1986).

A final variation of recategorization strategies is self-anchoring. Rather than recategorizing the in-group and out-group into a superordinate group, as in the common in-group identity model and the mutual intergroup differentiation model, the out-group is linked to the self. As a result, in-group favoritism is then extended to the out-group (Otten & Wentura, 2001). Gulker and Monteith (2013) tested the effect of self-anchoring by having White participants imagine that they adopted a baby from another country. Three experiments showed that this single link between the self and the out-group fostered perceived overlap between the self and out-group, which in turn was associated with reduced explicit prejudice toward the out-group. Yet, this single potent link was not enough to influence prejudice on implicit measures; implicit prejudice was only reduced for participants who repeatedly practiced this link between the self and the baby.

Woodcock and Monteith (2013) further demonstrated the effect of self-out-group conditioning on implicit measures. White participants were placed in a group including four Black people and instructed to study photographs of their group members so that they could accurately identify them. Then, in a speeded reaction time task, participants categorized photographs of people as belonging either to "MY GROUP" or to "OTHER GROUP." In all, participants paired Black people with "MY GROUP" 163 times and White people with "OTHER GROUP" 149 times. Compared to a neutral condition, participants who repeatedly paired Black people with the self subsequently showed less racial bias on an evaluative IAT. In another, related study, participants repeatedly moved a circle representing the self to overlap completely with a circle representing a Black student (Phills et al., 2011). Compared to participants who moved the self-circle away from the Black student, participants who moved the self-circle to overlap with the Black student circle automatically viewed Black individuals as more similar to the self on a self-Black IAT. Like the other strategies described in this section, self-anchoring changes how the out-group is cognitively categorized. Self-anchoring also shows how the associations between two mental representations (in this case, the self and the out-group) can be changed by repeatedly practicing new associations.

Overall, the personalization and recategorization strategies reviewed in this section provide powerful tools for improving intergroup attitudes and behaviors with procedures that influence how people think about out-groups. We believe that an especially compelling aspect of this work includes some demonstrations where the strategies produce positive outcomes even when there are long-standing intergroup animosities. For instance, indirect contact reduces prejudice even when the groups studied have a protracted history of

severe intergroup conflict (e.g., in the Middle East; Lemmer & Wagner, 2015). Still, we wonder about boundary conditions. For instance, are people with preexisting prejudiced attitudes likely to engage in subtyping when provided with individuating information, rather than changing their beliefs about the out-group? Some evidence suggests that this is indeed the case (Riek et al., 2013), and thus more work investigating whether motivational factors moderate the effect of these cognitive strategies on bias reduction is needed.

This section describes strategies that indirectly change the stereotypic and affective content associated with out-group members: Rather than changing the associations that are activated by target group members, these strategies change whether the target group member is viewed as an in-group member, as out-group member, or not as a group member at all. But what if, rather than changing the category that is associated with out-group members, we directly change the traits and affect that are associated with out-group members? Such strategies are described in the next section.

Changing the Stereotypic and Affective Associations Activated by Target Group Members

In this section, we review prejudice-reduction strategies that directly target the mental representations automatically activated by target group members. Once researchers discovered that the automatic activation of implicit stereotypic and evaluative bias could be influenced by a variety of factors (e.g., the context and attributes of targets; the state and situation of perceivers; see Monteith et al., 2013), they set out to understand whether certain factors might be used strategically to change the automatic activation of bias. These strategies—subcategorized here as counterstereotyping and evaluative conditioning strategies—work by changing the stereotypic and affective associations with target group members.

Counterstereotyping

In counterstereotyping, a target group is repeatedly associated with the opposite of stereotypes (e.g., Black targets are repeatedly associated with “motivated” rather than “lazy”). The idea is that repeated practice with counterstereotyping changes what content is associated with the targeted group and, as a result, reduces the automatic activation of stereotypic information when presented with that target group.

In the earliest research examining counterstereotype training, participants completed hundreds of trials where they practiced responding “NO” when photographs of target group members (e.g., women) were presented with stereotypic traits (e.g., “weak”) and “YES” when presented with a counterstereotype trait (e.g., “powerful”; Kawakami et al., 2000; see also Kawakami et al., 2005; Kawakami, Dovidio, & van Kamp, 2007). Participants then completed a measure of stereotype activation, in which they stated the color of a filler word following a stereotypic or nonstereotypic trait prime. Participants who did not complete counterstereotype training were slower to state colors following stereotypic versus nonstereotypic primes, indicating that automatic activation of the stereotypic prime interfered with the color statements. Conversely, participants who completed counterstereotype training did not show speed differences between the stereotypic and nonstereotypic primes, an effect that persisted 24 hr later (Kawakami et al., 2000). These results illustrate that counterstereotype training reduced the automatic activation of stereotypes. Later work revealed that responding “YES,” or affirming the counterstereotype, is what drives the counterstereotyping effect; participants who simply responded “NO” to stereotypes actually showed increased activation of stereotypes (Gawronski et al., 2008). The meaningful negation of stereotypes (i.e., “That’s wrong!” versus mere negation, or “NO”) can also reduce automatic activation (Johnson et al., 2018). In other iterations of counterstereotype training, participants are presented with two traits and

instructed to choose the trait that is counterstereotypic for the presented person (e.g., “weak” versus “powerful” when presented with a woman) across hundreds of trials (Kawakami et al., 2005; Kawakami, Dovidio, & Van Camp, 2007).

Counterstereotype training effects are robust and have successfully reduced stereotype activation in many laboratory studies. Yet, one may wonder how well this strategy maps onto daily life, where opportunities to practice counterstereotyping in such an intense and concentrated way may not regularly arise. Other strategies draw on counterstereotyping principles but may be more easily applied in everyday life. For example, with mental imagery training, participants are instructed to imagine a counterstereotypic scenario (e.g., a scenario where a Black man rescues the participant from a White mugger; Blair et al., 2001; Lai et al., 2014). Compared to control participants, participants who imagined the counterstereotypic scenario showed less bias on both evaluative (Lai et al., 2014) and stereotype (Blair et al., 2001) IATs, were slower to associate target group members with stereotypic traits and faster to associate them with counterstereotypic traits (Blair et al., 2001), and had fewer false memories of stereotypic traits (Blair et al., 2001). Counterstereotypic exemplar training has also been explored. Compared to control participants, participants who view admired Black exemplars (e.g., Barack Obama, Oprah Winfrey) and disliked White exemplars (e.g., Hitler, Jeffrey Dahmer) showed less pro-White bias on both evaluative (Dasgupta & Greenwald, 2001; Gonsalkorale et al., 2010; Lai et al., 2014) and stereotype (Dasgupta & Asgari, 2004) IATs.

Seeing counterstereotypic exemplars in everyday life also reduces the activation of stereotypic associations. In a field study, Dasgupta and Asgari (2004) examined female students’ gender-leadership IAT scores between their first and second years of college. Half of the students attended a women’s college; the other half attended a coeducational college. Students at the women’s college were exposed to more women in counterstereotypic positions (i.e., female science, technology, engineering, and math [STEM] faculty and female deans), which in turn reduced the magnitude of IAT bias they showed during their second year compared to the first year. In contrast, students at the coeducational college, who were exposed to fewer women in counterstereotypic positions, responded with more bias during their second year compared to their first year. This finding illustrates that, even outside the lab, naturally occurring exposure to counterstereotypic exemplars can reduce stereotypic associations.

In an effort to understand the processes responsible for reducing automatic stereotype activation through counterstereotyping, Calanchini and colleagues (2020) applied the Quadruple process model (or Quad model; Conrey et al., 2005). Work with the Quad model starts with the important assertion that responses on implicit measures (e.g., the IAT) are influenced by multiple processes that can be estimated separately by employing certain techniques when analyzing one’s data. Of interest here, the Quad model provides a way of parsing an association-oriented process that reflects activation of cognitive associations from a control-oriented process that involves inhibition of activated associations that conflict with detected correct responses, or “overcoming bias.” Calanchini and colleagues (2020) found that both engaging in counterstereotypic imagery and exposure to counterstereotypic exemplars reduced association activation processes and increased the overcoming bias process.

Evaluative Conditioning

In evaluative conditioning training, target group members are paired with pleasant versus unpleasant words (Calanchini et al., 2013; Lai et al., 2014; Olson & Fazio, 2006) or images (Kurdi & Banaji, 2017; Olson & Fazio, 2006). In some paradigms, participants are instructed to negate biased evaluative pairings (e.g., “Black-bad”) and affirm nonbiased evaluative pairings (e.g., “Black-good”; e.g., Calanchini et al., 2013); in other paradigms, participants are merely exposed to the evaluative pairings (e.g., Kurdi & Banaji, 2017). Compared to control conditions, evaluative conditioning reduced participants’ prejudice scores on evaluative priming measures (Fazio et al., 1995; Olson & Fazio, 2001, 2006) and evaluative IATs (Calanchini et al., 2013; Kurdi & Banaji, 2017; Lai et al., 2014). These evaluative conditioning effects occurred even when participants completed as few as 37 trials (Kurdi & Banaji, 2017). This research consistently suggests that evaluative conditioning reduces the automatic activation of prejudice.

Approach–avoidance strategies provide a variation on evaluative conditioning. These strategies build on the go/no-go association task, in which participants are instructed to press a key when target members are presented with positive words and to do nothing for all other pairings (Nosek & Banaji, 2001). Although originally used as a measure of automatic activated evaluations, the go/no-go task has been adapted for training by having participants complete many trials (between 45 and 100) where they press a key for target–positive word pairings and do nothing on other trials (Calanchini et al., 2020; Lai et al., 2014). Similar trainings have taken a literal approach to the approach–avoidance strategy by having participants move a joystick toward oneself when presented with an out-group member (i.e., “approach”) and push a joystick away from oneself (e.g., “avoid”) when presented with an in-group member (Kawakami, Phills, et al., 2007; Phills et al., 2011; Van Dessel et al., 2016, 2020). In such trainings, the “positive” concept being paired with the out-group is the self (see also Woodcock & Monteith, 2013). Although weak or null effects have at times been observed (see Van Dessel et al., 2020), approach–avoidance strategies typically reduce biased responding on evaluative IATs (Calanchini et al., 2020; Kawakami, Phills, et al., 2007; Lai et al., 2014; Phills et al., 2011; Van Dessel et al., 2020), evaluative priming tasks (Van Dessel et al., 2016), brain activity associated with detecting self–other differences (Phills et al., 2011), and nonverbal behavior (Kawakami, Phills, et al., 2007).

Evaluative conditioning strategies were originally believed to change the affective associations activated by exposure to the target group directly (Olson & Fazio, 2006; Van Dessel et al., 2016). However, recent research suggests that evaluative conditioning may also change automatic associations indirectly, by first changing explicit attitudes. Across six studies, Kurdi and Banaji (2017) found that evaluative statements, or verbally signaling that one will pair out-group members with positive constructs and in-group members with negative constructs in an upcoming task, reduced IAT scores to the same extent as traditional evaluative conditioning. Such statements may reduce IAT scores because they make participants aware of the conditioned pairs; indeed, research indicates that evaluative conditioning effects only occur when participants are aware of the conditioned contingencies (e.g., Pleyers et al., 2007). In line with such work, Kurdi and Banaji (2017) found the effect of evaluative statements was mediated by changes in explicit attitudes toward the same group. When evaluative statements were framed in terms of approach–avoidance instructions (i.e., “You will press the ‘down’ arrow to approach Luupites and the ‘up’ arrow to avoid Niffites”), only automatic, not explicit, evaluations were affected (Van Dessel et al., 2016). Yet other studies have failed to replicate the effect of evaluative statements on automatic evaluations altogether (see Van Dessel et al., 2020). These studies present mixed evidence on whether repeated conditioning is necessary to change evaluative associations or whether one-shot, language-based training is enough.

Even more recent work suggests that evaluative conditioning may be driven not by association-oriented processes, but by control-oriented processes. Using the Quad model, Calanchini and colleagues (2020) found that evaluative conditioning training (both traditional and approach–avoidance training) did not change the extent to which Black-bad and White-good associations were activated in an evaluative IAT.

However, evaluative conditioning did increase control-oriented processes to prevent the expression of bias. Thus, rather than changing automatic associations, evaluative conditioning may instead reduce bias through control-oriented processes.

Potential for Long-Term and Behavioral Change

Early evidence suggested that the bias-reducing power of both counterstereotyping and evaluative conditioning persisted beyond the immediate training experience (Kawakami et al., 2000; Olson & Fazio, 2006). Yet in recent years, both counterstereotyping and evaluative conditioning have been plagued by failures to produce long-term change in automatically activated bias. In two studies, both counterstereotyping and evaluative conditioning reduced bias on implicit measures initially, but bias bounced back to initial levels when tested again 24 to 96 hr later (Lai et al., 2016). Furthermore, a meta-analysis of 492 studies revealed that, though responses on implicit measures can be changed, the effect is small ($|ds| < 0.30$; Forscher et al., 2019; see also Van Dessel et al., 2020).

Other researchers (Kawakami et al., 2017) have pointed out numerous explanations for why the strategies tested by Lai and colleagues (2016) may have failed to reduce long-term change in IAT scores. The strategies were brief (5 min or less), online interventions. In contrast, other research testing these strategies included trainings that can last up to an hour or were even broken into multiple sessions (Kawakami et al., 2000, 2017). The trainings shown to be effective in past work have also typically occurred in a controlled, laboratory setting. While laboratory settings certainly come with their own set of limitations, these differences in study length and setting (i.e., online versus laboratory) could certainly influence the long-term effectiveness of counterstereotyping and evaluative conditioning. Furthermore, many of the studies in the meta-analysis by Forscher and colleagues (2019) employed IATs, and Lai and colleagues (2016) used evaluative IATs as their only implicit measure of change. This makes sense, because many counterstereotyping and evaluative conditioning studies also rely on the IAT as the primary measure for whether the training reduces prejudiced or stereotypic associations. Yet, some researchers contend there are validity issues with the IAT (e.g., Buttrick et al., 2020; Oswald et al., 2013) that call into question the sole reliance on this measure.

To these points, we add that biased associations may also bounce to initial levels after training because we are frequently exposed to situational and cultural reminders of stereotypes. Indeed, the bias of crowds theory (Payne et al., 2017) suggests that implicit bias is determined by situational factors present at a given point in time. To the extent that bias is present in our culture, neighborhood, media exposures, social interactions, institutional practices, and so on at a given point in time, it will be reflected in measures of people's implicit bias (Payne et al., 2017). A single counterstereotyping or evaluative conditioning training is unlikely to withstand the effects of biased environments that constantly reinforce biased associations. However, if environments are changed, there may be a greater chance of long-lasting change (see Dasgupta and Asgari's 2004 research, summarized above).

Even if counterstereotyping and evaluative conditioning can change automatic associations, they may not change biased judgments and behavior. A few studies found that counterstereotype training reduced stereotype application (i.e., choices of men versus women leaders), but only if the application task was dissociated from the training (e.g., by time; Kawakami et al., 2005; Kawakami, Dovidio, & Van Camp, 2007). In another study, approach-avoidance training caused White participants to sit closer and orient their body more directly to a Black confederate (Kawakami, Phillips, et al., 2007). However, other research has found that counterstereotyping did not reduce application of bias (e.g., evaluation of racist jokes), even when it reduced activation (Burns et al., 2017, Experiments 1 and 2); when application was reduced in this work, the effect was not mediated by the reduction in implicit bias (Experiment 3). Indeed, Forscher and colleagues' (2019) meta-analysis revealed that strategies that reduce bias on implicit measures do not necessarily

produce corresponding changes in explicit measures or behavior. They described the effect of implicit bias-reduction strategies on behavior as “trivial” in magnitude and found that any observed reductions for explicit or behavioral measures were not mediated by changes on implicit measures. In other words, counterstereotyping and evaluative conditioning trainings can reduce the extent to which prejudiced associations are automatically activated, at least weakly; however, this reduction does not lead to changes in prejudiced behavior.

One reason that counterstereotyping conditioning may not lead to behavioral change is that these trainings typically focus on changing only one or two associations among a larger set of biased associations. Consequently, even after training, other associations may still be activated (Burns et al., 2017; Casper et al., 2010; Gawronski & Bodenhausen, 2006). For example, a White person who has learned to associate “Black people” with “smart” through counterstereotype training may still have a number of other stereotypic associations (e.g., unmotivated, lazy) that can be activated and influence judgments and behaviors.

More generally, we question whether researchers have focused too much on training directly to the implicit bias task (e.g., the IAT). To be sure, practicing “Black-good” and “White-bad” associations can immediately reduce the ease of pairing Black with good and White with bad relative to the reverse pairings on an IAT. However, when one enters a different context (e.g., a classroom setting), will a target group member cue the activation of “good”? We have our doubts, especially given decades of research in cognitive psychology on the encoding specificity principle (Tulving & Thomson, 1973).

Exerting Control Over Bias

Especially in light of the challenges associated with “training away” our biases, bias reduction may depend crucially on strategies that enable people to exert control over their biases. Strategies that involve exerting control take as a given that bias is automatically activated, but then build on the social cognitive idea that our complex social behaviors are the product of multiple, coacting processes.

To begin, we note a couple of strategies that, at first glance, seem to provide straightforward remedies to thwarting biased responses. One strategy is suppression, whereby people attempt to banish stereotypic and prejudicial thoughts from their mind so that they cannot influence judgments and behaviors. However, the stereotype suppression literature shows that trying not to think about stereotypes actually primes them, resulting in greater use or “rebound” (Wegner, 2011) of stereotypic responses (Follenfant & Ric, 2010; Macrae et al., 1994; Monteith et al., 1998). Another strategy with lay appeal involves efforts not to notice or “see” group membership, such as when people attempt to be “colorblind.” However, colorblindness can also result in rebound, with greater subsequent implicit and explicit bias (Apfelbaum et al., 2008; Richeson & Nussbaum, 2004). In addition, the colorblindness strategy introduces a host of problematic outcomes, including decreased recognition of discrimination (Apfelbaum et al., 2010) and reduced support for programs to foster equality (Mazzocco et al., 2012).

As the work discussed below will underscore, there is no easy “fix” that will eliminate the influence of automatically activated biases on our responses. However, with the use of initially intentional and effortful processes, control may be exerted over one’s biased responses. Indeed, these strategies allow for the possibility that, with practice, control itself may become reflexive and automatic.

Implementation Intentions

Implementation intentions are *if-then* statements that describe what a person will think or do when encountering a specific situation that is relevant to achieving their goal (Gollwitzer, 1999; Gollwitzer & Sheeran, 2006). For instance, with a goal to be environmentally conscious, an implementation intention could be, “If I see trash on the ground, I will pick it up and throw it out.” A great deal of research has shown that establishing implementation intentions helps people attain their goals by increasing the successful completion of goal-relevant behaviors and creating persistence in the face of failures (Gollwitzer & Sheeran, 2006).

Implementation intentions are successful because they take advantage of both automatic and controlled processing. People form the implementation intention (i.e., link a situational cue to a specific behavior) with deliberative, conscious processes, and then the behavior can be reflexively carried out in the presence of that situational cue (Lengfelder & Gollwitzer, 2001). As a result, when the situational cue is present, the goal-relevant behavior can be implemented without much thought or conscious processing (Brandstatter et al., 2001; Gollwitzer & Sheeran, 2006; Webb & Sheeran, 2007).

Stewart and Payne (2008) investigated whether implementation intentions can be used to exert control over implicit bias. In Experiments 1 and 2, participants were introduced to the weapons identification task (Payne, 2001). Across many trials, a face prime of either a Black or a White man (presented for 200 ms) preceded a picture of either a tool or a weapon (presented for 100 ms), and participants used specified keys to indicate whether the object was either a tool or a weapon while responding as quickly as possible. At first, participants were informed that the faces would just orient them, but after initial trials, the experimenter explained that the race of faces influences people’s responses. Then participants in a “think-safe” condition were told, “We would like you to commit yourself to responding to the Black face by thinking the word ‘safe.’ By thinking the word ‘safe,’ you are reminding yourself on each trial that you are just as safe interacting with a Black individual as with a White individual” (Payne, 2001, p. 1337). Then participants were instructed to say to themselves silently, “I definitely want to respond as accurately as possible by thinking the word, safe. Whenever I see a Black face on the screen, I will think the word, safe” (Payne, 2001, p. 1337).

Replicating past research (Payne, 2001), participants in two control conditions that did not receive the think-safe implementation intention mistakenly identified tools as guns more often when the tools were preceded by Black compared to White faces. In contrast, this misidentification effect did not occur in the think-safe condition. A third experiment conceptually replicated these effects, but with the intention to think “good” when seeing a Black face while completing a White/Black-good/bad IAT (see also Lai et al., 2014). Finally, application of the process dissociation procedure (Jacoby, 1991), which, like the Quad model noted previously, can parse out the extent of contribution of automatic and controlled responses, suggested that implementation intentions reduced bias by decreasing automatic contributions during the task.

Mendoza and colleagues (2010) conceptually replicated the positive effect of implementation intentions while participants performed the shooter bias task (Correll et al., 2002; similar to the weapons identification task). Both a “distraction-inhibiting” implementation intention (i.e., “If I see a Black person, then I will ignore his race!”) and a “response-facilitating” implementation intention (i.e., “If I see a person with a gun, then I will shoot!”) reduced racial bias. Process dissociation procedure analyses indicated that reduced bias was associated with both reduced automatic and increased controlled processing for the distraction-inhibiting implementation intention and increased controlled processing for the response-facilitating implementation intention.

More research is needed to examine the potential utility of implementation intentions for exerting control over bias. Do these effects generalize beyond reaction time tasks to affect judgmental and behavior bias? How much of these effects are tied to participants’ task performance goal to follow the experimenter’s

instructions (e.g., the experimenter instruction to “think ‘safe’ when you see a Black face”), rather than to a goal to respond in egalitarian ways? Given the demonstrated deleterious effects of having the goal to ignoring race (i.e., colorblindness), might the positive effects of the distraction-inhibiting implementation intention be unique to this particular paradigm? Finally, research is needed to examine whether implementation intentions that apply to a variety of contexts (e.g., “If I see a Black person, I will think twice before acting on the basis of stereotypes”) can increase control and reduce stereotypic judgments and biased behavior. If not, the utility of implementation intentions may be confined to instances in which specific situational cues are linked to specific behaviors.

In addition to these questions for future research, an important consideration is just how people initially learn to identify situations where the exertion of control over bias is needed. We turn now to discuss a strategy that considers this issue while postulating the processes and mechanisms involved in learning to self-regulate one’s biased responses.

Self-Regulation of Prejudice

The self-regulation of prejudice (SRP) model describes the process through which individuals’ initial awareness of their propensity to respond in biased ways leads to a variety of affective, motivational, and cognitive outcomes that ultimately reduce future biased responses (e.g., Monteith, 1993; Monteith et al., 2002). Importantly, for these self-regulatory processes to occur, the initial awareness of one’s biases must be accompanied by sufficient motivation to do something about them. This argument harkens back to Gordon Allport’s (1954) seminal observation that people must be “dissatisfied” with their prejudice to be motivated to change. Work on the SRP is also rooted in Devine’s (1989) argument that people who wish to “break the prejudice habit” can do so by using controlled processing to replace biased responses. However, SRP research goes beyond this earlier work by detailing precisely how control can be exerted and with what effects, and the model posits that control can be increasingly automatized with practice at self-regulating prejudice responses (Monteith et al., 2009).

Figure 20.1 shows the SRP model. The model begins with the idea that automatic processes may result in biased responses that conflict with people’s personal standards that they should respond in less biased ways. This results in discrepant responses. Consistent with other work on self-discrepancies (for seminal work, see Higgins, 1987), when people become aware of the discrepancies between desired and actual responses, they experience negative self-directed affect, such as guilt and disappointment with the self. In the lab, these affective consequences have been observed following various manipulations to create discrepancy salience, including task feedback to participants indicating they engaged in biased responses (e.g., Amodio et al., 2007; Monteith, 1993; Monteith et al. 2002); completion of the Should–Would Discrepancy Questionnaire, which primes people’s prejudice-related discrepancies (e.g., Burns et al., 2017; Devine et al., 1991; Monteith & Voils, 1998); IAT completion, which can cause people’s biases to become evident to them (Fehr & Sassenberg, 2010; Monteith et al., 2001); and having one’s biases pointed out by others (i.e., confrontation; e.g., Czopp et al., 2006). Beyond the lab, these discrepancies correspond to actual lived experiences: Interviews revealed that White participants catch themselves thinking or behaving in biased ways toward Black people in everyday situations and that this leads to negative self-directed affect (Monteith et al., 2010).

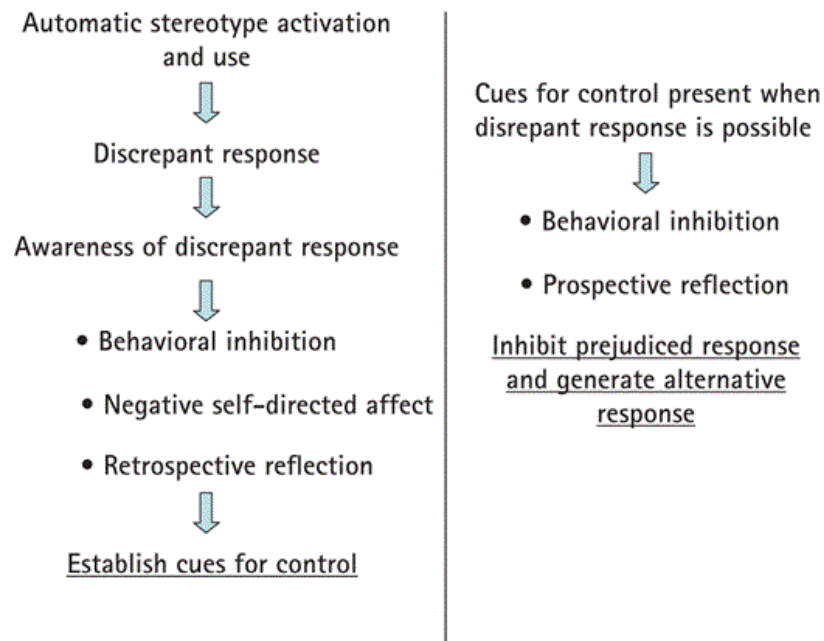


Figure 20.1 The self-regulation of prejudice model.

According to the SRP model, negative self-directed affect elicited by awareness of discrepant responses triggers a cascading sequence of events that ultimately increases one's ability to respond without bias in the future. First, it triggers the behavioral inhibition system, which monitors for mismatches between expected and actual events (Carver, 2003; Gray, 1982) and is linked to neural mechanisms associated with conflict monitoring and response control (Amodio, Devine, & Harmon-Jones, 2008; Botvinick et al., 2001). Behavioral inhibition system activation is manifested behaviorally in a brief pause in ongoing responding (i.e., behavioral inhibition) and retrospective reflection. This reflection involves paying greater attention to the discrepant response and the stimuli (e.g., a certain environment, a certain stereotype) that accompanied the discrepant response. As a result, associations are built between the stimuli that predict the occurrence of the discrepant response, the discrepant response itself, and the negative affect resulting from awareness of the response. Such associations result in the establishment of cues for control (Monteith, 1993; Monteith et al., 2002).

p. 574 For instance, if I catch myself assuming that a Black person in the store works there rather than being a shopper, I will experience negative self-directed affect, and my ongoing behavior will be interrupted briefly. Then, brief retrospective reflection will allow me to build associations between being in a store, Black people, the erroneous assumption I made, and my negative affect; in other words, I will establish cues for control. In one empirical demonstration of this process (Monteith et al., 2002; Studies 1 and 2), researchers provided White participants with fixed feedback that they experienced negative physiological arousal to either pictures of Black people (the racial condition) or nonracial pictures. Compared to participants in the nonracial condition, participants in the racial condition experienced discrepancies between their nonprejudiced goals and their supposedly prejudiced response. As a result, participants in the racial condition paused briefly (i.e., behavioral inhibition) after each instance of elevated arousal feedback to pictures of Black people, reported heightened negative self-directed affect, and showed preoccupation with their discrepant responses in a subsequent thought-listed task (retrospective reflection). Furthermore, this negative self-directed affect was positively correlated with a measure of retrospective reflection, suggesting cues for control had been established.

Initial discrepancy experiences and establishing cues for control are critical for exerting control over bias in the future. When cues for control are present in the future (e.g., I am in a store and see a Black person), the

behavioral inhibition system is triggered again, resulting in a brief pause in ongoing behavior that allows one to engage in prospective reflection (“Wait, how do I want to respond here?”). As a result, nonbiased responses can be generated instead of biased responses (e.g., through replacement, such as when one does not laugh at a stereotypical joke; by gathering individuating information). For example, in one experiment (Monteith et al., 2002, Experiment 4), White participants initially completed an IAT. The palpable difficulty of pairing Black with pleasant and White with unpleasant, compared to the reverse pairings, made biases salient to participants (Monteith et al., 2001). Participants then reported their affect, including negative self-directed affect. In a later, supposedly unrelated task, participants were presented with a series of words and instructed to indicate quickly (with specified keystrokes) whether they liked or disliked each word. Some words were historically Black names (e.g., Jamal). As participants’ reports of negative self-directed affect increased, they took longer to respond on Black name trials (prospective reflection), and they were more likely to provide “like” responses to these words (i.e., generating a nonbiased response). These results are consistent with the initial establishment of cues for control that later facilitate nonbiased responding. Other studies provide additional support for the self-regulatory steps posited by the SRP model (e.g., Amodio, Devine, & Harmon-Jones, 2007; Burns et al., 2017; Monteith, 1993; Monteith et al., 2002).

Self-regulation as we have described it thus far applies to people who consciously endorse egalitarian attitudes or who are internally motivated to respond without bias (Plant & Devine, 1998). In contrast, high-prejudiced people experience little negative self-directed affect when they respond in biased ways (e.g., Monteith, 1993; Monteith & Voils, 1998). Yet, even these individuals can be motivated to avoid biases resulting from external forces, such as nonprejudiced social norms (i.e., external motivation, Plant & Devine, 1998). More externally motivated individuals may experience discrepancies when they respond with greater bias than others think is appropriate (Monteith et al., 1993). This creates general discomfort, followed by self-regulation to avoid biased responses (Monteith et al., 2010). However, self-regulation among high-prejudiced or externally motivated individuals is situationally determined, such as occurring in public but not private contexts (Plant & Devine, 1998).

People who are personally committed to egalitarian responding may continually improve their regulatory abilities by repeatedly establishing cues for control across time and situations. As a result of this practice, people should eventually be able to exert control over their biases early in the processing stream and reflexively, reducing bias even on implicit measures. Consistent with this possibility are findings that people who are highly internally (but not externally) motivated to respond without bias showed less implicit bias than other participants, even when they performed the implicit bias task under cognitive load (Devine et al., 2002). Such results are interpreted as evidence of highly efficient self-regulation. Furthermore, Amodio and colleagues (2004) have uncovered neurocognitive evidence that conflict monitoring and detection facilitates response control on implicit racial bias measures. Using process dissociation procedure estimates, Amodio, Devine, and Harmon-Jones (2008) found that controlled processing on implicit measures is higher for internally (but not externally) motivated participants, and these controlled processing estimates correspond with neurocognitive evidence of conflict monitoring (Amodio et al., 2006). Finally, in research using the Quad model (Sherman et al., 2008), Gonsalkorale et al. (2011) found evidence of enhanced activity of a controlled process that involves the detection of conflicting responses among people higher in internal motivation to respond without bias. This detection process allows for greater success at determining nonbiased responses.

Although research supports the SRP and even early-stage control of bias, these demonstrations beg an important question: If people can effectively self-regulate and become increasingly efficient at exerting control of automatic processes, why then does automatically activated bias and its negative consequences appear to be so pervasive, even among egalitarian-minded people? We believe this is because people wish to see themselves as unbiased and egalitarian (e.g., Dovidio & Gaertner, 2004), and thus even egalitarian-minded people may be reticent and defensive when recognizing their own biases (e.g., Howell et al., 2015;

Monteith et al., 2001). When presented with information about or evidence of bias, people may try to explain or justify that bias in terms of factors having nothing to do with bias. For instance, after receiving IAT feedback indicating that they were biased, people responded by blaming the researchers and delegitimizing the IAT task (e.g., Howell et al., 2015). People may also simply lack awareness of all the ways in which bias can influence responses or be uncertain about how exactly to respond in nonbiased ways (Banaji & Greenwald, 2013). Thus, raising awareness of and concern about bias are important first steps to exerting control over it. Providing concrete evidence of biased behavior may also reduce the extent to which people try to justify or “explain away” their biased behavior (e.g., Parker et al., 2018). Finally, teaching people what they can do about their bias may increase self-efficacy and reduce defensive responses to bias awareness. We turn now to a discussion of an intervention designed to increase bias awareness without creating defensiveness and, in turn, foster bias reduction.

Prejudice Habit-Breaking Intervention

Faced with the realities of bias within organizations, diversity training programs are often called to the rescue. For instance, after high-profile protests against police violence, many police departments implemented diversity training programs (Kaste, 2020). Starbucks closed all of their stores for a day of implicit bias training after a Starbucks employee called the police on two Black men who were accused of “trespassing” when they were sitting in the store just having a discussion (Scheiber & Abrams, 2018). Yet, for a variety of reasons, these trainings often fail to have the intended effects (Dobbins & Kalev, 2013; Pendry et al., 2007). Answering the call to develop and validate theoretically and empirically driven diversity interventions (e.g., Paluck, 2012), Devine and colleagues have developed and tested a “prejudice habit-breaking intervention” (Devine et al., 2012, 2017; for a review, see Cox & Devine, 2019). This intervention is designed to increase recognition of often-unintentional bias without creating defensiveness, to increase motivation to combat bias, and to provide concrete strategies to reduce bias.

p. 576 Consistent with the SRP model, the prejudice habit-breaking intervention begins by making participants aware of the discrepancies between their egalitarian personal standards for responding and their prejudiced behavior. Participants complete an IAT (the Black-White IAT for interventions focused on race bias; the gender-leadership IAT for interventions focused on gender bias) and receive feedback on their performance. To discourage dismissiveness of IAT feedback, the intervention facilitator dispels common lay explanations for IAT performance (e.g., that the order of the blocks is responsible for one’s scores). The intervention facilitator then highlights how implicit bias can lead to unintentional biased behavior, including the presentation of relatable scenarios that involve the operation of unintentional bias. Then the facilitator presents empirical evidence of unintentional bias and its important discriminatory consequences. This evidence also underscores that these biases are pervasive habits, automatically learned as a result of socialization, and do not represent a “moral failing” (Cox & Devine, 2019). Altogether, this portion of the intervention is intended to foster participants’ awareness of their potential biases without creating defensiveness and to increase their concern about discrimination.

The second part of the intervention presents a variety of empirically informed strategies that people may voluntarily adopt “as a toolkit” to help them to reduce their bias (e.g., stereotype replacement, counterstereotyping, individuation; see Cox & Devine, 2019). The intervention also presents backfiring strategies to avoid (e.g., stereotype suppression and colorblindness). The facilitator and the participants discuss situations in which these strategies should prove helpful, along with their potentially synergistic effects. Participants are told that using these strategies will help them to reach their goal of reducing bias and that, like all habits, the prejudice habit can only be broken through sustained, conscious effort that, over time, should become more automatic.

Though this is a time-intensive intervention (often a 2.5-hr intervention session) that requires effort and sustained commitment for people to put what they have been taught to use, tests of its efficacy suggest that the effort pays off. Compared to participants who completed a control intervention, participants who completed the prejudice habit-breaking intervention reported greater awareness of bias (Carnes et al., 2015; Devine et al., 2012; Forscher et al., 2017) and concern about racial discrimination (Devine et al., 2012; Forscher et al., 2017), showed reduced automatic activation of bias on the IAT (though results have been somewhat mixed; Devine et al., 2012; Forscher et al., 2017), reported more self-efficacy to recognize and replace stereotypic thoughts and behaviors (Carnes et al., 2015), and were more likely to confront a prejudiced social media post (Forscher et al., 2017). Impressively, many of these effects have persisted for up to 2 years (Forscher et al., 2017). In addition, positive effects for organizational climate have been found (Carnes et al., 2015), and one large-scale study connected the intervention with a reduction of institutional-level bias (Devine et al., 2017). Specifically, science, technology, engineering, mathematics, and medical departments were randomly assigned either to receive the habit-breaking intervention tailored to address gender bias or not to receive an intervention. In a follow-up 2 years after the intervention, female faculty hiring in the intervention cluster of departments had increased by 15 percentage points, from 32% to 47%. In contrast, the control cluster started at 33% and remained about the same at follow-up. We suspect one reason for this lasting success is that the intervention establishes nonprejudiced social norms. Even people who lack internal motivation to respond without prejudice may feel external pressure to respond without prejudice. Once established, such nonprejudiced social norms may even influence others in the relevant environment who did not participate in the intervention (see, for example, Paluck, 2011).

p. 577 Tapping into the precise processes underlying this type of intervention work is no doubt challenging, because the precise processes likely vary across individuals, situations, and the specific strategies people use to respond in nonbiased ways. However, we wonder how important such precision is. When an intervention is grounded in theory and builds on a variety of past empirical work in a smartly assembled package, a more pressing aim is arguably to establish replicability and generalizability. For instance, the research-heavy focus in this intervention worked well for undergraduate and faculty participants, and now researchers might focus on whether results generalize to other people and organizational contexts.

Conclusions

Research conducted through a social cognition lens has provided invaluable insight into how intergroup bias may be reduced. First, personalization and recategorization strategies influence how we draw group boundaries in ways that can reduce our bias toward people who initially were not part of our favored in-groups. Second, using social cognition methods to reduce well-learned and automatically activated associations provides insight into how these associations can be modified to reduce implicit bias. Counterstereotyping and evaluative conditioning training paradigms do not appear to produce lasting change or change to judgments and behavior. However, this work underscores the potential for modifying environments, media expressions, readily observable exemplars, and social learning experiences to alter people's mental representations of target groups. Third, although such sweeping and deep changes are difficult and take time (Charlesworth & Banaji, 2019), research examining the exertion of control over activated biases helps us to understand how people may make personal and more immediate progress in reducing their intergroup biases. Through cognitive, motivational, and learning processes, people can learn to identify when they are at risk for responding in biased ways so that they can self-regulate and reduce biased feelings, judgments, and behaviors. This research also underscores the importance of raising awareness of bias, creating concern and motivation to reduce one's biases, and learning what tools can be used to avoid bias, all of which may be addressed in practical interventions such as the prejudice habit-breaking intervention. As we hope our review has underscored, none of these strategies is purely cognitive, but they also involve affective and motivational processes that are important in bias reduction.

Careful consideration of which bias-reduction strategy may work best in particular contexts is important for anyone wishing to implement these strategies in everyday life. For instance, reducing the salience of group memberships or avoiding them entirely, as with decategorization and recategorization, may be useful in some contexts; however, these strategies may be inappropriate in other contexts because they also encourage a colorblind approach that neglects the importance of group-based identities (Rattan & Ambady, 2013). Another issue to consider is whether strategies will be beneficial and appropriate for everyone who participates. Care must be taken to ensure that strategy-based interventions benefit both majority and minority group members, rather than having unintended and negative outcomes for minority group members (Pietri et al., 2019). Indeed, prejudice-reduction research often focuses on the biases of majority group members without considering how minority group members may respond to such interventions. Although there are exceptions (e.g., Holooin & Shelton, 2012), this majority group focus is a current limitation for many prejudice-reduction strategies. Nevertheless, some strategies may be beneficial for both majority and minority groups with tailoring. For instance, in recent work in our lab, we developed an intervention for women in STEM that highlighted how gender stereotypes may be automatically activated and applied to *the self* and consequently undermine feelings of belonging and perceived aptitude in STEM fields. The intervention, though based on the SRP, was tailored to interrupt the process of applying these stereotypes to the self.

Finally, we wish to note two largely overlooked but important issues related to prejudice reduction in social cognition research (and in social psychological research more generally). First, research is needed to understand how intergroup bias can be reduced among people who lack any sort of internal or external motivation to respond without prejudice. These individuals may be motivated to express their prejudices (Forscher et al., 2015), and they likely endorse group-based dominance (i.e., social dominance orientation, Ho et al., 2015) and right-wing authoritarianism (Altemeyer, 1996). We are doubtful that any of the strategies we have reviewed has great promise for reducing intergroup bias among such individuals.

Second, the goal of all of the intergroup bias-reduction strategies we have reviewed has been to reduce individual-level bias, which is characteristic of current social psychological research on prejudice reduction. However, systemic bias toward Black people in the United States has become increasingly apparent to many people because of recently publicized fatal police violence toward Black people, including George Floyd, Breonna Taylor, Ahmaud Arbery, and many others. The Black Lives Matter movement against police brutality (<https://blacklivesmatter.com/>) has gained considerable momentum, and protests against police brutality continue to occur in many cities. The problem of systemic bias and the need to combat it surfaces not only in relation to law enforcement, but also much more broadly in terms of cultural artifacts, dominant ideologies, biased policies, and structural and institutional practices (see Salter et al., 2018). The reduction of intergroup bias must be approached not only in terms of individual bias, which is one contributor to systemic bias, but also by studying how intergroup bias can be reduced beyond individuals. Indeed, the type of individual-level bias-reduction strategies discussed in this chapter are unlikely to be sustained when, for instance, environments continue to promote and reproduce bias and when policies and practices continue to favor dominant groups.

How the study of prejudice is approached has long been shaped by zeitgeist and significant societal events at particular points in time (Duckitt, 2010). Moving forward, the current racial pandemic will no doubt lead researchers to devote more effort to understanding what strategies may encourage individuals' recognition of the problem of systemic bias and motivation and action to combat it. We also expect greater focus on understanding how bias can be dismantled structurally.

Furthermore, studying not only how individuals change systems but also how systems change individuals is important. According to the bias of crowds perspective (Payne et al., 2017), implicit bias is determined by people's exposure to biased environments and situational contexts. Although we agree with Gawronski and Bodenhausen's (2017) perspective that person-level factors are also influential (i.e., a Person \times Situation

interaction), the enhanced situational emphasis in the bias of crowds perspective and related research is important. With respect to prejudice reduction, this perspective suggests that altering environments to have less biased laws, policies, and practices will lead to changes among individuals. Indeed, in a compelling demonstration of this process, researchers examined whether the passing of state-by-state same-sex legislation in the United States was followed by decreased (state-clustered) implicit and explicit antigay prejudice in those states (Ofosu et al., 2019). Results indicated that antigay bias showed a sharp reduction after states adopted same-sex marriage legislation. However, states that never passed legislation showed an increase in antigay bias following federal legalization, indicating backlash.

p. 579 A greater emphasis on studying prejudice in terms of systems, environments, and situations places less singular emphasis on individual-level social cognitive processes. The individual-level strategies for encouraging prejudice reduction reviewed in this chapter have dominated research agendas for decades. Widening the lens to bring system-level factors into the picture not only will bring greater balance, but also presents the opportunity for improved theoretical frameworks with practical applications for reducing prejudice.

References

Allport, G. W. (1954). *The nature of prejudice*. Addison–Wesley.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Altemeyer, B. (1996). *The authoritarian specter*. Harvard University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2007). A dynamic model of guilt: Implications for motivation and self-regulation in the context of prejudice. *Psychological Science*, *18*(6), 524–530. <https://doi.org/10.1111/j.1467-9280.2007.01933.x>

[Google Scholar](#) [WorldCat](#)

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*, 60–74.

[Google Scholar](#) [WorldCat](#)

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science*, *15*(2), 88–93. <https://doi.org/10.1111/j.0963-7214.2004.01502003.x>

[Google Scholar](#) [WorldCat](#)

Amodio, D. M., Kubota, J. T., Harmon-Jones, E., & Devine, P. G. (2006). Alternative mechanisms for regulating racial responses according to internal vs external cues. *Social Cognitive and Affective Neuroscience*, *1*(1), 26–36.

<https://doi.org/10.1093/scan/nsl002>

[Google Scholar](#) [WorldCat](#)

Amodio, D. M., Master, S. L., Yee, C. M., & Taylor, S. E. (2008). Neurocognitive components of behavioral inhibition and activation systems: Implications for theories of self-regulation. *Psychophysiology*, *45*(1), 11–19. <https://doi.org/10.1111/j.1469-8986.2007.00609.x>

[Google Scholar](#) [WorldCat](#)

Apfelbaum, E. P., Pauker, K., Sommers, S. R., & Ambady, N. (2010). In blind pursuit of racial equality? *Psychological Science*, *21*(11), 1587–1592. <https://doi.org/10.1177/0956797610384741>

[Google Scholar](#) [WorldCat](#)

Apfelbaum, E. P., Sommers, S. R., & Norton, M. I. (2008). Seeing race and seeming racist? Evaluating strategic colorblindness in social interaction. *Journal of Personality and Social Psychology*, *95*(4), 918–932. <https://doi.org/10.1037/a0011990>

[Google Scholar](#) [WorldCat](#)

Banaji, M. R., & Greenwald, A. G. (2013). *Blindspot: Hidden biases of good people*. Delacorte Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer Jr. & T. K. Srull (Eds.), *Handbook of social cognition: Basic processes* (pp. 1–40). Erlbaum.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Batson, C. D., Chang, J., Orr, R., & Rowland, J. (2002). Empathy, attitudes and action: Can feeling for a member of a stigmatized group motivate one to help the group. *Personality and Social Psychology Bulletin*, *28*(12), 1656–1666.

<https://doi.org/10.1177/014616702237647>

[Google Scholar](#) [WorldCat](#)

Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology*, *81*(5), 828–841. <https://doi.org/10.1037/0022-3514.81.5.828>

[Google Scholar](#) [WorldCat](#)

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108(3), 624–652. <https://doi.org/10.1037/0033-295X.108.3.624>
[Google Scholar](#) [WorldCat](#)

Brandstätter, V., Lengfelder, A., & Gollwitzer, P. M. (2001). Implementation intentions and efficient action initiation. *Journal of Personality and Social Psychology*, 81(5), 946–960. <https://doi.org/10.1037/0022-3514.81.5.946>
[Google Scholar](#) [WorldCat](#)

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues*, 55(3), 429–444. <https://doi.org/10.1111/0022-4537.00126>
[Google Scholar](#) [WorldCat](#)

Brewer, M. B., Dull, V., & Lui, L. (1981). Perceptions of the elderly: Stereotypes as prototypes. *Journal of Personality and Social Psychology*, 41(4), 656–670. <https://doi.org/10.1037/0022-3514.41.4.656>
[Google Scholar](#) [WorldCat](#)

Brewer, M. B., & Miller, N. (1984). Beyond the contact hypothesis: Theoretical perspectives on desegregation. In N. Miller & M. B. Brewer (Eds.), *Groups in contact: The psychology of desegregation* (pp. 281–302). Academic Press.
[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Burns, M. D., Monteith, M. J., & Parker, L. R. (2017). Training away bias: The differential effects of counterstereotype training and self-regulation on stereotype activation and application. *Journal of Experimental Social Psychology*, 73, 97–110. <https://doi.org/10.1016/j.jesp.2017.06.003>
[Google Scholar](#) [WorldCat](#)

Buttrick, N., Axt, J., Ebersole, C. R., & Huband, J. (2020). Re-assessing the incremental predictive validity of Implicit Associations Test. *Journal of Experimental Social Psychology*, 88, 103941. <https://doi.org/10.1016/j.jesp.2019.103941>
[Google Scholar](#) [WorldCat](#)

Calanchini, J., Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2013). Counter-prejudicial training reduces activation of biased associations and enhances response monitoring. *European Journal of Social Psychology*, 43, 321–325. <https://doi.org/10.1002/ejsp.1941>
[Google Scholar](#) [WorldCat](#)

p. 580 Calanchini, J., Lai, C. K., & Klauer, K. C. (2020). Reducing implicit racial preferences: III. A process-level examination of changes in implicit preferences. *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/pspi0000339>
[Google Scholar](#) [WorldCat](#)

Carlston, D. E. (2010). On the nature of social cognition: My defining moment. In D. Carlston (Ed.), *Oxford handbook of social cognition* (pp. 3–15). Oxford University Press
[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Carnes, M., Devine, P. G., Manwell, L. B., Byars-Winston, A., Fine, E., Ford, C. E., Forscher, P., Isaac, C., Kaatz, A., Magua, W., Palta, M., & Sheridan, J. (2015). The effect of an intervention to break the gender bias habit for faculty at one institution: A cluster randomized, controlled trial. *Academic Medicine*, 90(2), 221–230. <https://doi.org/10.1097/ACM.0000000000000552>
[Google Scholar](#) [WorldCat](#)

Carver, C. (2003). Pleasure as a sign you can attend to something else: Placing positive feelings within a general model of affect. *Cognition and Emotion*, 17(2), 241–261. <https://doi.org/10.1080/026999303022294>
[Google Scholar](#) [WorldCat](#)

Casper, C., Rothermund, K., & Wentura, D. (2010). Automatic stereotype activation is context dependent. *Social Psychology*, 41(3), 131–136. <https://doi.org/10.1027/1864-9335/a000019>
[Google Scholar](#) [WorldCat](#)

Charlesworth, T. E. S., & Banaji, M. R. (2019). Patterns of implicit and explicit attitudes: I. Long-term change and stability from 2007 to 2016. *Psychological Science*, 30(2), 174–192. <https://doi.org/10.1177/0956797618813087>
[Google Scholar](#) [WorldCat](#)

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: The quadruple process model of implicit task performance. *Journal of Personality and Social Psychology*, 89(4), 469–487. <https://doi.org/10.1037/0022-3514.89.4.469>
[Google Scholar](#) [WorldCat](#)

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, 83(6), 1314–1329. <https://doi.org/10.1037/0022-3514.83.6.1314>
[Google Scholar](#) [WorldCat](#)

Cox, W. L. T., & Devine, P. G. (2019). The prejudice habit-breaking intervention: An empowerment-based confrontation approach. In R. Mallet & M. J. Monteith (Eds.), *Confronting prejudice and discrimination: The science of changing minds and behaviors* (pp. 250–274). Academic Press.
[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Crisp, R. J., Husnu, S., Meleady, R., Stathi, S., & Turner, R. N. (2010). From imagery to intention: A dual route model of imagined contact effects. *European Review of Social Psychology*, 21(1), 188–236. <https://doi.org/10.1080/10463283.2010.543312>
[Google Scholar](#) [WorldCat](#)

Crisp, R. J., & Turner, R. N. (2009). Can imagined interactions produce positive perceptions? Reducing prejudice through simulated social contact. *American Psychologist*, 64(4), 231–240. <https://doi.org/10.1037/a0014718>
[Google Scholar](#) [WorldCat](#)

Crisp, R. J., Hewstone, M., & Rubin, M. (2001). Does multiple categorization reduce intergroup bias? *Personality and Social Psychology Bulletin*, 27(1), 76–89. <https://doi.org/10.1177/0146167201271007>
[Google Scholar](#) [WorldCat](#)

Crisp, R. J., Stone, C. H., & Hall, N. R. (2006). Recategorization and subgroup identification: Predicting and preventing threats from common ingroups. *Personality and Social Psychology Bulletin*, 32(2), 230–243. <https://doi.org/10.1177/0146167205280908>
[Google Scholar](#) [WorldCat](#)

Czopp, A. M., Monteith, M. J., & Mark, A. Y. (2006). Standing up for a change: Reducing bias through interpersonal confrontation. *Journal of Personality and Social Psychology*, 90(5), 784–803. <https://doi.org/10.1037/0022-3514.90.5.784>
[Google Scholar](#) [WorldCat](#)

Dasgupta, N., & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology*, 40(5), 642–658. <https://doi.org/10.1016/j.jesp.2004.02.003>
[Google Scholar](#) [WorldCat](#)

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, 81(5), 800–814. <https://doi.org/10.1037/0022-3514.81.5.800>
[Google Scholar](#) [WorldCat](#)

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56(1), 5–18. <https://doi.org/10.1037/0022-3514.56.1.5>
[Google Scholar](#) [WorldCat](#)

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. L. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, 48(6), 1267–1278. <https://doi.org/j.jesp.2012.06.003>

Devine, P. G., Forscher, P. S., Cox, W. T. L., Kaatz, A., Sheridan, J., & Carnes, M. (2017). A gender bias habit-breaking intervention led to increased hiring of female faculty in STEMM departments. *Journal of Experimental Social Psychology, 73*, 211–215. <https://doi.org/10.1016/j.jesp.2017.07.002>

Google Scholar WorldCat

Devine, P. G., Monteith, M. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology, 60*(6), 817–830. <https://doi.org/10.1037/0022-3514.60.6.817>

Google Scholar WorldCat

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. <https://doi.org/10.1037/0022-3514.82.5.835>

Google Scholar WorldCat

p. 581 Dobbin, F., & Kalev, A. (2013). The origins and effects of corporate diversity programs. In Q. M. Roberson (Ed.), *The Oxford handbook of diversity and work* (pp. 253–281). Oxford University Press.

Google Scholar Google Preview WorldCat COPAC

Dovidio, J. F., & Gaertner, S. L. (2004). Aversive racism. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 36, pp. 1–52). Elsevier Academic Press. [https://doi.org/10.1016/S0065-2601\(04\)36001-6](https://doi.org/10.1016/S0065-2601(04)36001-6)

Google Scholar Google Preview WorldCat COPAC

Dovidio, J. F., Gaertner, S. L., & Validzic, A. (1998). Intergroup bias: Status, differentiation, and a common in-group identity. *Journal of Personality and Social Psychology, 75*(1), 109–120. <https://doi.org/10.1037/0022-3514.75.1.109>

Google Scholar WorldCat

Dovidio, J. F., Gaertner, S. L., Validzic, A., Makota, K., Johnson, B., & Frazier, S. (1997). Extending the benefits of recategorization: Evaluations, self-disclosure, and helping. *Journal of Experimental Social Psychology, 4*(33), 401–420.

<https://doi.org/10.1006/jesp.1997.1327>

Google Scholar WorldCat

Dovidio, J., Gaertner, S., & Saguy, T. (2009). Commonality and the complexity of “we”: Social attitudes and social change. *Society for Personality and Social Psychology, 13*(1), 3–20. <https://doi.org/10.1177/1088868308326751>

Google Scholar WorldCat

Dovidio, J. F., Ten Vergert, M., Stewart, T. L., Gaertner, S. L., Johnson, J. D., Esses, V. M., Riek, B. M., & Pearson, A. R. (2004). Perspective and prejudice: Antecedents and mediating mechanisms. *Personality and Social Psychology Bulletin, 30*(12), 1537–1549. <https://doi.org/10.1177/0146167204271177>

Google Scholar WorldCat

Duckitt, J. (2010). Historical overview. In J. F. Dovidio, M. Hewstone, P. Glick, & V. M. Esses (Eds.), *The SAGE handbook of prejudice, stereotype, and discrimination* (pp. 29–44). SAGE.

Google Scholar Google Preview WorldCat COPAC

Ensari, N., Christian, J., Kuriyama, D. M., & Miller, N. (2012). The personalization model revisited: An experimental investigation of the role of five personalization strategies on prejudice reduction. *Group Processes and Intergroup Relations, 15*, 503–522.

<https://doi.org/10.1177/1368430211434576>

Google Scholar WorldCat

Ensari, N., & Miller, N. (2001). Decategorization and the reduction of bias in the crossed categorization paradigm. *European Journal of Social Psychology, 31*(2), 193–216. <https://doi.org/10.1002/ejsp.42>

Google Scholar WorldCat

Ensari, N., & Miller, N. (2002). The out-group must not be so bad after all: The effects of disclosure, typicality, and salience on intergroup bias. *Journal of Personality and Social Psychology*, 83(2), 313–329. <https://doi.org/10.1037/0022-3514.83.2.313>
[Google Scholar](#) [WorldCat](#)

Ensari, N., & Miller, N. (2005). Prejudice and intergroup attributions: The role of personalization and performance feedback. *Group Processes and Intergroup Relations*, 8(4), 391–410. <https://doi.org/10.1177/1368430205056467>
[Google Scholar](#) [WorldCat](#)

Ensari, N. K., & Miller, N. (2006). The application of personalization model in diversity management. *Group Processes and Intergroup Relations*, 9(4), 589–607. <https://doi.org/10.1177/1368430206067679>
[Google Scholar](#) [WorldCat](#)

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(6), 1013–1027. <https://doi.org/10.1037/0022-3514.69.6.1013>
[Google Scholar](#) [WorldCat](#)

Feagin, J. R. (1991). The continuing significance of race: Anti-Black discrimination in public places. *American Sociological Review*, 56(1), 101–116. <https://doi.org/10.2307/2095676>
[Google Scholar](#) [WorldCat](#)

Fehr, J., & Sassenberg, K. (2010). Willing and able: How internal motivation and failure help to overcome prejudice. *Group Processes & Intergroup Relations*, 13(2), 167–181. <https://doi.org/10.1177/1368430209343116>
[Google Scholar](#) [WorldCat](#)

Finlay, K. A., & Stephan, W. G. (2000). Improving intergroup relations: The effects of empathy on racial attitudes. *Journal of Applied Social Psychology*, 30(8), 1720–1737. <https://doi.org/10.1111/j.1559-1816.2000.tb02464.x>
[Google Scholar](#) [WorldCat](#)

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. McGraw–Hill.
[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Follenfant, A., & Ric, F. (2010). Behavioral rebound following stereotype suppression. *European Journal of Social Psychology*, 40(5), 774–782. <https://doi.org/10.1002/ejsp.649>
[Google Scholar](#) [WorldCat](#)

Forscher, P. S., Cox, W. T. L., Graetz, N., & Devine, P. G. (2015). The motivation to express prejudice. *Journal of Personality and Social Psychology*, 109(5), 791–812. <https://doi.org/10.1037/pspi0000030>
[Google Scholar](#) [WorldCat](#)

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*, 117(3), 522–559. <https://doi.org/10.1037/pspa0000160>
[Google Scholar](#) [WorldCat](#)

Forscher, P. S., Mitamura, C., Dix, E. L., Cox, W. T. L., & Devine, P. G. (2017). Breaking the prejudice habit: Mechanisms, timecourse, and longevity. *Journal of Experimental Social Psychology*, 72, 133–146. <https://doi.org/10.1016/j.jesp.2017.04.009>
[Google Scholar](#) [WorldCat](#)

Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A., & Rust, M. C. (1993). The common intergroup identity model: Recategorization and the reduction of intergroup bias. *European Review of Social Psychology*, 4(1), 1–26. <https://doi.org/10.1080/14792779343000004>
[Google Scholar](#) [WorldCat](#)

Gaertner, S. L., Dovidio, J. F., Banker, B. S., Houlette, M., Johnson, K. M., & McGlynn, E. A. (2000). Reducing intergroup conflict:

From superordinate goals to decategorization, recategorization, and mutual differentiation. *Group Dynamics: Theory, Research, and practice*, 4(1), 98–114. <https://doi.org/10.1037/1089-2699.4.1.98>

WorldCat

Gaertner, S. L., Dovidio, J. F., Rust, M. C., Nier, J. A., Banker, B. S., Ward, C. M., Mottola, G. R., & Houlette, M. (1999). Reducing intergroup bias: Elements of intergroup cooperation. *Journal of Personality and Social Psychology*, 76(3), 388–402.

<https://doi.org/10.1037/0022-3514.76.3.388>

Google Scholar WorldCat

p. 582 Gaertner, S. L., Mann, J. A., Dovidio, J. F., Murrell, A. J., & Pomare, M. (1990). How does cooperation reduce intergroup bias? *Journal of Personality and Social Psychology*, 59(4), 692–704. <http://dx.doi.org/10.1037/0022-3514.59.4.692>

Google Scholar WorldCat

Gaertner, S. L., Mann, J., Murrell, A., & Dovidio, J. F. (1989). Reducing intergroup bias: The benefits of recategorization. *Journal of Personality and Social Psychology*, 57(2), 239–249. <https://doi.org/10.1037/0022-3514.57.2.239>

Google Scholar WorldCat

Gaertner, S. L., Rust, M. C., Dovidio, J. F., Bachman, B. A., & Anastasio, P. A. (1994). The contact hypothesis: The role of common ingroup identity on reducing intergroup bias. *Small Group Research*, 25(2), 224–249. <https://doi.org/10.1177/1046496494252005>

Google Scholar WorldCat

Galinsky, A. D., & Ku, G. (2004). The effects of perspective-taking on prejudice: The moderating role of self-evaluation. *Personality and Social Psychology Bulletin*, 30(5), 594–604. <https://doi.org/10.1177/0146167203262802>

Google Scholar WorldCat

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, 78(4), 708–724. <https://doi.org/10.1037/0022-3514.78.4.708>

Google Scholar WorldCat

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132(5), 692–731. <https://doi.org/10.1037/0033-2909.132.5.692>

Google Scholar WorldCat

Gawronski, B., & Bodenhausen, G. V. (2017). Beyond persons and situations: An interactionist approach to understanding implicit bias. *Psychological Inquiry*, 28, 268–272. <https://doi.org/10.1080/1047840X.2017.1373546>

Google Scholar WorldCat

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. <https://doi.org/10.1016/j.jesp.2006.12.004>

Google Scholar WorldCat

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, 54(7), 493–503.

<https://doi.org/10.1037/0003-066X.54.7.493>

Google Scholar WorldCat

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 69–119). Elsevier Academic Press.

Google Scholar Google Preview WorldCat COPAC

Gonsalkorale, K., Allen, T. J., Sherman, J. W., & Klauer, K. C. (2011). Mechanisms of group membership and exemplar exposure effects on implicit attitudes. *Social Psychology*, 41(3), 158–168. <https://doi.org/10.1027/1864-9335/a000023>

Google Scholar WorldCat

Gray, J. A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. Oxford

University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

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. <https://doi.org/10.1037/0022-3514.74.6.1464>

[Google Scholar](#) [WorldCat](#)

Greenwald, A. G., & Pettigrew, T. F. (2014). With malice towards none and charity for some: Ingroup favoritism enables discrimination. *American Psychologist*, 69(7), 669–684. <https://doi.org/10.1037/a0036056>

[Google Scholar](#) [WorldCat](#)

Gulker, J. E., & Monteith, M. J. (2013). Intergroup boundaries and attitudes: The power of a single potent link. *Personality and Social Psychology Bulletin*, 39(7), 943–955. <https://doi.org/10.1177/0146167213485444>

[Google Scholar](#) [WorldCat](#)

Hamilton, D. L., & Carlston, D. E. (2013). *The emergence of social cognition*. In D. E. Carlston (Ed.), *Oxford handbook of social cognition* (pp. 16–32). Oxford University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Hewstone, M. (1996). *Stereotypes and stereotyping*. Guilford Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Hewstone, M., & Brown, R. (1986). Contact is not enough: An intergroup perspective on the “contact hypothesis.” In M. Hewstone & R. Brown (Eds.), *Contact and conflict in intergroup encounters* (pp. 1–44). Basil Blackwell.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, 94(3), 319–340.

<https://doi.org/10.1037/0033-295X.94.3.319>

[Google Scholar](#) [WorldCat](#)

Ho, A. K., Sidanius, J., Kteily, N., Sheehy-Skeffington, J., Pratto, F., Henkel, K. E., Foels, R., & Stewart, A. L. (2015). The nature of social dominance orientation: Theorizing and measuring preferences for intergroup inequality using the new SDO₇ Scale. *Journal of Personality and Social Psychology*, 109(6), 1003–1028. <https://doi.org/10.1037/pspi0000033>

Holoien, D. S., & Shelton, J. N. (2012). You deplete me: The cognitive costs of colorblindness on ethnic minorities. *Journal of Experimental Social Psychology*, 48, 562–565. <https://doi.org/10.1016/j.jesp.2011.09.010>

[Google Scholar](#) [WorldCat](#)

Hornsey, M. J., & Hogg, M. A. (2000). Assimilation and diversity: An integrative model of subgroup relations. *Personality and Social Psychology Review*, 4(2), 143–156. https://doi.org/10.1207/S15327957PSPR0402_03

[Google Scholar](#) [WorldCat](#)

Howell, J. L., Gaither, S. E., & Ratliff, K. A. (2015). Caught in the middle: Defensive responses to IAT feedback among Whites, Blacks, and biracial Black/Whites. *Social Psychological and Personality Science*, 6(4), 373–381.

<https://doi.org/10.1177/1948550614561127>

[Google Scholar](#) [WorldCat](#)

p. 583 Husnu, S., & Crisp, R. J. (2010). Elaboration enhances the imagined contact effect. *Journal of Experimental Social Psychology*, 46(6), 943–950. <https://doi.org/10.1016/j.jesp.2010.05.014>

[Google Scholar](#) [WorldCat](#)

Ioannou, M., Hewstone, M., & Al Ramiah, A. (2017). Inducing similarities and differences in imagined contact: A mutual intergroup differentiation approach. *Group Processes & Intergroup Relations*, 20(4), 427–446. <https://doi.org/10.1177/1368430215612221>

[Google Scholar](#) [WorldCat](#)

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30(5), 513–541. [https://doi.org/10.1016/0749-596X\(91\)90025-F](https://doi.org/10.1016/0749-596X(91)90025-F)

[Google Scholar](#) [WorldCat](#)

Johnson, I. R., Kopp, B. M., & Petty, R. E. (2018). Just say no! (and mean it): Meaningful negation as a tool to modify automatic racial attitudes. *Group Processes and Intergroup Relations*, 21(1), 88–110. <https://doi.org/10.1177/1368430216647189>

[Google Scholar](#) [WorldCat](#)

Kaste, M. (2020, September 10). *NYPD study: Implicit bias training changes minds, not necessarily behavior*. NPR.

<https://www.npr.org/2020/09/10/909380525/nypd-study-implicit-bias-training-changes-minds-not-necessarily-behavior>

Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). Intergroup perception and cognition: An integrative framework for understanding the causes and consequences of social categorization. In J. M. Olson (Ed.), *Advances in experimental social psychology* (pp. 1–80). Elsevier Academic Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

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.

<https://doi.org/10.1037/0022-3514.78.5.871>

[Google Scholar](#) [WorldCat](#)

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.

<https://doi.org/10.1016/j.jesp.2004.05.004>

[Google Scholar](#) [WorldCat](#)

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.

<https://doi.org/10.1177/1368430207074725>

[Google Scholar](#) [WorldCat](#)

Kawakami, K., Phills, C. E., Steele, J. R., & Dovidio, J. F. (2007). (Close) distance makes the heart grow fonder: Improving implicit racial attitudes and interracial interactions through approach behaviors. *Journal of Personality and Social Psychology*, 92(6), 957–971. <https://doi.org/10.1037/0022-3514.92.6.957>

[Google Scholar](#) [WorldCat](#)

Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Jr., Bahnik, S., Bernstein, M. J., Bocian, K., Brandt, M. J., Brooks, B., Brumbaugh, C. C., Cemailcar, Z., Chandler, J., Cheong, W., Davis, W. E., Devos, T., Eisner, M., Frankowska, N., Furrow, D., Galliani, E. M., ... Nosek, B. A. (2014.) Investigating variation in replicability: A “Many Labs” replication project. *Social Psychology*, 45(3), 142–152. <https://doi.org/10.1027/1864-9335/a000178>.

[Google Scholar](#) [WorldCat](#)

Kurdi, B., & Banaji, M. R. (2017). Repeated evaluative pairings and evaluative statements: How effectively do they shift implicit attitudes? *Journal of Experimental Psychology: General*, 146(2), 194–213. <https://doi.org/10.1037/xge0000239>

[Google Scholar](#) [WorldCat](#)

Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. L., Joy-Gaba, J., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E. E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., ...

Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143(4), 1765–1785. <https://doi.org/10.1037/a0036260>

[Google Scholar](#) [WorldCat](#)

Lai, C. K., Skinner, A. L., Cooley, E., Murrari, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., Marshburn, C. K., Simon, S., Blanchar, J. C., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhass, F. M. H., Burns, M., ...

- Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, 145(8), 1001–1016. <http://dx.doi.org/10.1037/xge0000179>
[Google Scholar](#) [WorldCat](#)
- Lemmer, G., & Wagner, U. (2015). Can we really reduce ethnic prejudice outside the lab? A meta-analysis of direct and indirect contact interventions. *European Journal of Social Psychology*, 45, 152–168.
[Google Scholar](#) [WorldCat](#)
- Lengfelder, A., & Gollwitzer, P. M. (2001). Reflective and reflexive action control in patients with frontal brain lesions. *Neuropsychology*, 15(1), 80–100. <https://doi.org/10.1037/0894-4105.15.1.80>
[Google Scholar](#) [WorldCat](#)
- Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology*, 67(5), 808–817. <https://doi.org/10.1037/0022-3514.67.5.808>
[Google Scholar](#) [WorldCat](#)
- Mallett, R. K., Wilson, T. D., & Gilbert, D. T. (2008). Expect the unexpected: Failure to anticipate similarities leads to an intergroup forecasting error. *Journal of Personality and Social Psychology*, 94(2), 265–277. <https://doi.org/10.1037/0022-3514.94.2.265>
[Google Scholar](#) [WorldCat](#)
- p. 584 Mazzocco, P. J., Cooper, L. W., & Flint, M. (2012). Different shades of racial colorblindness: The role of prejudice. *Group Processes & Intergroup Relations*, 15(2), 167–178. <https://doi.org/10.1177/1368430211424763>
[Google Scholar](#) [WorldCat](#)
- McConnell, A. R., Brown, C. M., & Shoda, T. M. (2013). The social cognition of the self. In D. E. Carlston (Ed.), *The Oxford handbook of social cognition* (pp. 497–516). Oxford University Press.
[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)
- Meleady, R., & Seger, C. R. (2017). Imagined contact encourages prosocial behavior towards outgroup members. *Group Processes & Intergroup Relations*, 20(4), 447–464. <https://doi.org/10.1177/1368430215612225>
[Google Scholar](#) [WorldCat](#)
- 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*, 34(4), 512–523.
<https://doi.org/10.1177/0146167210362789>
[Google Scholar](#) [WorldCat](#)
- Meyers, C., Aumer, K., Schoniwitz, A., Janicki, C., Pauker, K., Chang, E. C., Gaither, S. E., & Williams, A. (2020). Experiences with microaggressions and discrimination in racially diverse and homogeneously white contexts. *Cultural Diversity and Ethnic Minority Psychology*, 26(2), 250–259. <https://doi.org/10.1037/cdp0000293>
[Google Scholar](#) [WorldCat](#)
- Miles, E., & Crisp, R. J. (2014). A meta-analytic test of the imagined contact hypothesis. *Group Process & Intergroup Relations*, 17(1), 3–26. <https://doi.org/10.1177/1368430213510573>
[Google Scholar](#) [WorldCat](#)
- Miller, N. (2002). Personalization and the promise of contact theory. *Journal of Social Issues*, 58(2), 387–410.
<https://doi.org/10.1111/1540-4560.00267>
[Google Scholar](#) [WorldCat](#)
- Miller, N., Brewer, M. B., & Edwards, K. (1985). Cooperative interaction in desegregated settings: A laboratory analogue. *Journal of Social Issues*, 41(3), 63–79. <https://doi.org/10.1111/j.1540-4560.1985.tb01129.x>
[Google Scholar](#) [WorldCat](#)
- Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice-reduction efforts. *Journal*

of *Personality and Social Psychology*, 65(3), 469–485. <https://doi.org/10.1037/0022-3514.65.3.469>

[Google Scholar](#) [WorldCat](#)

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. <https://doi.org/10.1037/0022-3514.83.5.1029>

[Google Scholar](#) [WorldCat](#)

Monteith, M. J., Devine, P. G., & Zuwerink, J. R. (1993). Self-directed versus other-directed affect as a consequence of prejudice-related discrepancies. *Journal of Personality and Social Psychology*, 64(2), 198–210. <https://doi.org/10.1037/0022-3514.64.2.198>

[Google Scholar](#) [WorldCat](#)

Monteith, M. J., Lybarger, J. E., & Woodcock, A. (2009). Schooling the cognitive monster: The role of motivation in the regulation and control of prejudice. *Social and Personality Psychology Compass*, 3(3), 211–226. <https://doi.org/10.1111/j.1751-9004.2009.00177.x>

[Google Scholar](#) [WorldCat](#)

Monteith, M. J., Mark, A. Y., & Ashburn-Nardo, L. (2010). The self-regulation of prejudice: Toward understanding its lived character. *Group Processes & Intergroup Relations*, 13(2), 183–200. <https://doi.org/10.1177/1368430209353633>

[Google Scholar](#) [WorldCat](#)

Monteith, M. J., Sherman, J. W., & Devine, P. G. (1998). Suppression as a stereotype control strategy. *Personality and Social Psychology Review*, 2(1), 63–82. https://doi.org/10.1207/s15327957pspr0201_4

[Google Scholar](#) [WorldCat](#)

Monteith, M. J., & Voils, C. I. (1998). Proneness to prejudiced responses: Toward understanding the authenticity of self-reported discrepancies. *Journal of Personality and Social Psychology*, 75(4), 901–916. <https://doi.org/10.1037//0022-3514.75.4.901>

[Google Scholar](#) [WorldCat](#)

Monteith, M. J., Voils, C. I., & Ashburn-Nardo, L. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit racial bias. *Social Cognition*, 19(4), 395–417. <https://doi.org/10.1521/soco.19.4.395.20759>

[Google Scholar](#) [WorldCat](#)

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). Oxford University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Nier, J. A., & Gaertner, S. L. (2012). The challenge of detecting contemporary forms of discrimination. *Journal of Social Issues*, 68(2), 207–220. <https://doi.org/10.1111/j.1540-4560.2012.01745.x>

[Google Scholar](#) [WorldCat](#)

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition*, 19(6), 625–666.

<https://doi.org/10.1521/soco.19.6.625.20886>

[Google Scholar](#) [WorldCat](#)

Ofose, E. K., Chambers, M. K., Chen, J. M., & Hehman, E. (2019). Same-sex marriage legalization associated with reduced implicit and explicit antigay bias. *Proceedings of the National Academy of Sciences*, 116(18), 8846–8851.

<https://doi.org/10.1073/pnas.1806000116>

[Google Scholar](#) [WorldCat](#)

Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science*, 12(5), 413–417. <https://doi.org/10.1111/1467-9280.00376>

[Google Scholar](#) [WorldCat](#)

- Olson, M. A., & Fazio, R. H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin*, 32(4), 421–433. <https://doi.org/10.1177/0146167205284004>
[Google Scholar](#) [WorldCat](#)
- Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, 105(2), 171–192. <https://doi.org/10.1037/a0032734>
[Google Scholar](#) [WorldCat](#)
- p. 585 Otten, S., & Wentura, D. (2001). Self-anchoring and in-group favoritism: An individual profiles analysis. *Journal of Experimental Social Psychology*, 37, 525–532. <https://doi.org/10.1006/jesp.2001.1479>
[Google Scholar](#) [WorldCat](#)
- Pagotto, L., Visintin, E. P., De Iorio, G., & Voci, A. (2013). Imagined intergroup contact promotes cooperation through outgroup trust. *Group Processes & Intergroup Relations*, 16(2), 209–216. <https://doi.org/10.1177/1368430212450057>
- Paluck, E. L. (2011). Peer pressure against prejudice: A high school field experiment examining social network change. *Journal of Experimental Social Psychology*, 47(2), 350–358. <https://doi.org/10.1016/j.jesp.2010.11.017>
[Google Scholar](#) [WorldCat](#)
- Paluck, E. L. (2012). Interventions aimed at the reduction of prejudice and conflict. In L. R. Tropp (Ed.), *The Oxford handbook of intergroup conflict* (pp. 179–192). Oxford University Press.
[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)
- Parker, L. R., Monteith, M. J., Moss-Racusin, C. A., & Van Camp, A. R. (2018). Promoting concern about gender bias with evidence-based confrontation. *Journal of Experimental Social Psychology*, 74, 8–23. <https://doi.org/10.1016/j.jesp.2017.07.009>
[Google Scholar](#) [WorldCat](#)
- Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, 81(2), 181–192. <https://doi.org/10.1037/0022-3514.81.2.181>
[Google Scholar](#) [WorldCat](#)
- Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, 89(3), 277–293. <https://doi.org/10.1037/0022-3514.89.3.277>
[Google Scholar](#) [WorldCat](#)
- Payne, B. K., Vuletich, H. A., & Lundberg, K. B. (2017). The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry*, 28(4), 233–248. <https://doi.org/10.1080/1047840X.2017.1335568>
[Google Scholar](#) [WorldCat](#)
- Pendry, L. F., Driscoll, D. M., & Field, S. C. (2007). Diversity training: Putting theory into practice. *Journal of Occupational and Organizational Psychology*, 80(1), 27–50. <https://doi.org/10.1348/096317906X118397>
[Google Scholar](#) [WorldCat](#)
- Phills, C. E., Kawakami, K., Tabi, E., Nadolny, D., & Inzlicht, M. (2011). Mind the gap: Increasing associations between the self and Blacks with approach behaviors. *Journal of Personality and Social Psychology*, 100(2), 197–210.
<https://doi.org/10.1037/a0022159>
[Google Scholar](#) [WorldCat](#)
- Pietri, E. S., Hennes, E. P., Dovidio, J. F., Brescoll, V. L., Bailey, A. H., Moss-Racusin, C. A., & Handelsman, J. (2019). Addressing unintended consequences of gender diversity interventions on women's sense of belonging in STEM. *Sex Roles*, 80(9–10), 527–547. <https://doi.org/10.1007/s11199-018-0952-2>
[Google Scholar](#) [WorldCat](#)
- Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, 75(3), 811–832. <https://doi.org/10.1037/0022-3514.75.3.811>

Pleyers, G., Corneille, O., Luminet, O., & Yzerbyt, V. (2007). Aware and (dis)liking: Item-based analyses reveal that valence acquisition via evaluative conditioning only emerges when there is contingency awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(1), 130–144. <https://doi.org/10.1037/0278-7393.33.1.130>
[WorldCat](#)

Rattan, A., & Ambady, N. (2013). Diversity ideologies and intergroup relations: An examination of colorblindness and multiculturalism. *European Journal of Social Psychology*, 43(1), 12–21. <https://doi.org/10.1002/ejsp.1892>
[Google Scholar](#) [WorldCat](#)

Richeson, J. A., & Nussbaum, R. J. (2004). The impact of multiculturalism versus color-blindness on racial bias. *Journal of Experimental Social Psychology*, 40(3), 417–423. <https://doi.org/10.1016/j.jesp.2003.09.002>
[Google Scholar](#) [WorldCat](#)

Riek, B. M., Mania, E. W., & Gaertner, S. L. (2013). Reverse subtyping: The effects of prejudice level on the subtyping of counterstereotypic outgroup members. *Basic and Applied Social Psychology*, 35, 409–417. <https://doi.org/10.1080/01973533.2013.823616>
[Google Scholar](#) [WorldCat](#)

Salter, P. S., Adams, G., & Perez, M. J. (2018). Racism in the structure of everyday worlds: A cultural-psychological perspective. *Current Directions in Psychological Science*, 27(3), 150–155. <https://doi.org/10.1177/0963721417724239>
[Google Scholar](#) [WorldCat](#)

Scheiber, N., & Abrams, R. (2018, April 18). Can training eliminate biases? Starbucks will test the thesis. *New York Times*. <https://www.nytimes.com/2018/04/18/business/starbucks-racial-bias-training.html>

Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1961). *The Robbers Cave experiment: Intergroup conflict and cooperation*. Wesleyan University Press.
[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

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. <https://doi.org/10.1037/0033-295X.115.2.314>
[Google Scholar](#) [WorldCat](#)

Stathi, S., & Crisp, R. J. (2008). Imagining intergroup contact promotes projection to outgroups. *Journal of Experimental Social Psychology*, 44(4), 943–957. <https://doi.org/10.1016/j.jesp.2008.02.003>
[Google Scholar](#) [WorldCat](#)

Stathi, S., Tsantila, K., & Crisp, R. J. (2012). Imagining intergroup contact can combat mental health stigma by reducing anxiety, avoidance and negative stereotyping. *The Journal of Social Psychology*, 152(6), 746–757. <https://doi.org/10.1080/00224545.2012.697080>
[Google Scholar](#) [WorldCat](#)

p. 586 Stewart, B. D., & Payne, B. K. (2008). Bringing automatic stereotyping under control: Implementation intentions as efficient means of thought control. *Personality and Social Psychology Bulletin*, 34(10), 1332–1345. <https://doi.org/10.1177/0146167208321269>
[Google Scholar](#) [WorldCat](#)

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149–178. <https://doi.org/10.1002/ejsp.2420010202>
[Google Scholar](#) [WorldCat](#)

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). Brooks/Cole.

Todd, A. R., Bodenhausen, G. V., Richeson, J. A., & Galinsky, A. D. (2011). Perspective taking combats automatic expressions of racial bias. *Journal of Personality and Social Psychology*, *100*(6), 1027–1042. <https://doi.org/10.1037/a0022308>

[Google Scholar](#) [WorldCat](#)

Todd, A. R., & Burgmer, P. (2013). Perspective taking and automatic intergroup evaluation change: Testing an associative self-anchoring account. *Journal of Personality and Social Psychology*, *104*(5), 786–802. <https://doi.org/10.1037/a0031999>

[Google Scholar](#) [WorldCat](#)

Todd, A. R., & Galinsky, A. D. (2014). Perspective-taking as a strategy for improving intergroup relations: Evidence, mechanisms, and qualifications. *Social and Personality Psychology Compass*, *8*(7), 374–397. <https://doi.org/10.1111/spc3.12116>

[Google Scholar](#) [WorldCat](#)

Tropp, L. R., & Pettigrew, T. F. (2005). Relationships between intergroup contact and prejudice among minority and majority status groups. *Psychological Science*, *16*(12), 951–957. <https://doi.org/10.1111/j.1467-9280.2005.01643.x>

[Google Scholar](#) [WorldCat](#)

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352–373. <https://doi.org/10.1037/h0020071>

[Google Scholar](#) [WorldCat](#)

Turner, R. N., & Crisp, R. J. (2010). Imagining intergroup contact reduces implicit prejudice. *British Journal of Social Psychology*, *49*(1), 129–142. <https://doi.org/10.1348/014466609X419901>

[Google Scholar](#) [WorldCat](#)

Turner, R. N., Crisp, R. J., & Lambert, E. (2007). Imagining intergroup contact can improve intergroup attitudes. *Group Processes & Intergroup Relations*, *10*(4), 427–441. <https://doi.org/10.1177/1368430207081533>

[Google Scholar](#) [WorldCat](#)

Turner, R. N., Hewstone, M., & Voci, A. (2007). Reducing explicit and implicit outgroup prejudice via direct and extended contact: The mediating role of self-disclosure and intergroup anxiety. *Journal of Personality and Social Psychology*, *93*(3), 369–388.

<https://doi.org/10.1037/0022-3514.93.3.369>

[Google Scholar](#) [WorldCat](#)

Turner, R. N., & West, K. (2012). Behavioural consequences of imagining intergroup contact with stigmatized outgroups. *Group Processes & Intergroup Relations*, *15*(2), 193–202. <https://doi.org/10.1177/1368430211418699>

[Google Scholar](#) [WorldCat](#)

Van Dessel, P., De Houwer, J., Gast, A., Roets, A., & Smith, C. T. (2020). On the effectiveness of approach–avoidance instructions and training for changing evaluations of social groups. *Journal of Personality and Social Psychology*, *119*(2), e1–e14.

<https://doi.org/10.1037/pspa0000189>

[Google Scholar](#) [WorldCat](#)

Van Dessel, P., De Houwer, J., Gast, A., Smith, C. T., & De Schryver, M. (2016). Instructing implicit processes: When instructions to approach or avoid influence implicit but not explicit evaluation. *Journal of Experimental Social Psychology*, *63*, 1–9.

<http://dx.doi.org/10.1016/j.jesp.2015.11.002> [0022-1031/](https://doi.org/10.1016/j.jesp.2015.11.002)

[Google Scholar](#) [WorldCat](#)

Vescio, T. K., Sechrist, G. B., & Paolucci, M. P. (2003). Perspective taking and prejudice reduction: The mediational role of empathy arousal and situational attributions. *European Journal of Social Psychology*, *33*(4), 455–472. <https://doi.org/10.1002/ejsp.163>

[Google Scholar](#) [WorldCat](#)

Vezzali, L., Capozza, D., Stathi, S., & Giovannini, D. (2012). Increasing outgroup trust, reducing inhumanization, and enhancing future contact intentions via imagined intergroup contact. *Journal of Experimental Social Psychology*, *48*(1), 437–440.

<https://doi.org/10.1016/j.jesp.2011.09.008>

[Google Scholar](#) [WorldCat](#)

Vorauer, J. D., Martens, V., & Sasaki, S. J. (2009). When trying to understand detracts from trying to behave: Effects of perspective taking in intergroup interaction. *Journal of Personality and Social Psychology*, 96(4), 811–827.

<http://dx.doi.org/10.1037/a0013411>

[Google Scholar](#) [WorldCat](#)

Vorauer, J. D., & Sasaki, S. J. (2009). Helpful only in the abstract? Ironic effects of empathy in intergroup interaction.

Psychological Science, 20(2), 191–197. <https://doi.org/10.1111/j.1467-9280.2009.02265.x>

[Google Scholar](#) [WorldCat](#)

Vorauer, J. D., & Sasaki, S. J. (2014). Distinct effects on imagine-other versus imagine-self perspective-taking on prejudice reduction. *Social Cognition*, 32(2), 130–147. <https://doi.org/10.1521/soco.2014.32.2.130>

[Google Scholar](#) [WorldCat](#)

Webb, T. L., & Sheeran, P. (2007). How do implementation intentions promote goal attainment? A test of component processes.

Journal of Experimental Social Psychology, 43(2), 295–302. <https://doi.org/10.1016/j.jesp.2006.02.001>

[Google Scholar](#) [WorldCat](#)

Wegner, D. M. (2011). Setting free the bears: Escape from thought suppression. *American Psychologist*, 66(8), 671–680.

<https://doi.org/10.1037/a0024985>

[Google Scholar](#) [WorldCat](#)

Wegner, D. M., & Vallacher, R. R. (1977). *Implicit psychology: An introduction to social cognition*. Oxford University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

West, K., & Greenland, K. (2016). Beware of “reducing prejudice”: Imagined contact may backfire if applied with a prevention focus. *Journal of Applied Social Psychology*, 46(10), 583–592. <https://doi.org/10.1111/jasp.12387>

[Google Scholar](#) [WorldCat](#)

p. 587 Wilder, D. A. (1978). Reduction of intergroup discrimination through individuation of the out-group. *Journal of Personality and Social Psychology*, 36(12), 1361–1374. <https://doi.org/10.1037/0022-3514.36.12.1361>

[Google Scholar](#) [WorldCat](#)

Wohl, M. J., & Branscomb, N. R. (2005) Forgiveness and collective guilt assignment to historical perpetrator groups depend on level of social category inclusiveness. *Journal of Personality and Social Psychology*, 88(2), 288–303.

<http://dx.doi.org/10.1037/0022-3514.88.2.288>

[Google Scholar](#) [WorldCat](#)

Woodcock, A., & Monteith, M. J. (2013). Forging links with the self to combat implicit bias. *Group Process & Intergroup Relations*, 16, 445–461. <https://doi.org/10.1177/1368430212459776>

[Google Scholar](#) [WorldCat](#)

Wright, S. C., Aron, A., Mclaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology*, 73(1), 73–90.

[Google Scholar](#) [WorldCat](#)

p. 588 Wyer, R. S., Jr., & Carlston, D. E. (1994). *The cognitive representation of persons and events*. In R. S. Wyer Jr. & T. K. Srull (Eds.), *Handbook of social cognition: Basic processes* (pp. 41–98). Erlbaum. ↵

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)