

*Chapter 2*

**COGNITIVE CONSISTENCY AND THE RELATION  
BETWEEN IMPLICIT AND EXPLICIT PREJUDICE:  
RECONCEPTUALIZING OLD-FASHIONED, MODERN,  
AND AVERSIVE PREJUDICE**

*Paula M. Brochu, Bertram Gawronski and Victoria M. Esses*

The University of Western Ontario

**ABSTRACT**

Research on modern, subtle manifestations of prejudice is currently characterized by the existence of diverse concepts that are not well integrated from a general perspective. The present chapter reviews a new framework that integrates several of these concepts from a cognitive consistency perspective, and the data collected to date that support this framework. Specifically, it is argued that the impact of immediate affective reactions toward stigmatized groups on evaluative judgments about these groups depends on the (in)consistency of this evaluation with other relevant beliefs reflecting the central components of old-fashioned, modern, and aversive forms of prejudice. This conceptualization makes unique predictions regarding the correspondence between implicit and explicit prejudice; namely, that the relation between the two should be moderated by the interaction of (a) egalitarianism-related nonprejudicial goals (i.e., the extent to which one believes that negative evaluations of stigmatized groups are wrong) and (b) perceptions of discrimination (i.e., the extent to which one believes that a specific social group is a target of systematic discrimination). The strength of the proposed framework is that it articulates specific relations among various forms of prejudice, can be applied to a variety of target groups, and provides new insights into the nature of and the relations between implicit and explicit prejudice.

**INTRODUCTION**

Imagine for a moment that you have been raised in a society that holds negative emotions, beliefs, and values about Camarians, an immigrant group taking up residence in

your country because of constant threat of natural disaster in their home region (see Maio, Esses, and Bell, 1994). Camarians affectively elicit discomfort, tension, disrespect, and anxiety from others, and are believed to be unfriendly, dishonest, lazy, and unintelligent. Further, Camarians are perceived as violating important values such as equality, law and order, economic development, and family. From a very young age, you have been bombarded with such negative messages about Camarians from a variety of sources, including your family, your neighbours, and the mass media. As a result, whenever you encounter a Camarian, you experience a negative reaction. Would you openly express this negative reaction? If so, how would it be expressed, and what factors would increase or decrease the likelihood that you would do so?

In the present chapter, we review a new framework that aims at specifying the particular conditions under which negative affective reactions to social groups translate into negative judgments about these groups (see Gawronski, Peters, Brochu, and Strack, 2008). This framework integrates the central components of old-fashioned, modern, and aversive prejudice by attributing a significant role to the interplay of egalitarianism-related nonprejudicial goals and perceptions of discrimination in the expression of prejudice. In addition, the integrative framework provides new insights into the relation between implicit and explicit forms of prejudice by distinguishing between associative and propositional processes underlying group evaluations (see Gawronski and Bodenhausen, 2006; Strack and Deutsch, 2004). Specifically, we argue that the desire to maintain consistency within one's system of beliefs determines whether spontaneous negative reactions toward a particular social group (implicit prejudice) will be reflected in negative evaluative judgments of that group (explicit prejudice).

For this purpose, we will first review the concepts of old-fashioned, modern, and aversive prejudice in greater detail. Drawing on this discussion, we will illustrate the core concepts of the integrative model: the distinction between associative and propositional processes, along with the importance of cognitive consistency, and the interplay of egalitarianism-related nonprejudicial goals and perceptions of discrimination. These concepts then will be applied to formally conceptualize the relations among old-fashioned, modern, and aversive forms of prejudice. In the remainder of the chapter, we will review evidence in support of the integrative model and discuss implications for research on the expression of prejudice.

## FORMS OF PREJUDICE

Prejudice is commonly defined as an overall negative evaluation of a social group or its members based on one's perceptions of and reactions to the group (Esses, Haddock, and Zanna, 1993). Historically, such negative evaluations have been studied in the form of direct, blatant expressions of negativity, which are often described as old-fashioned prejudice. Over the last few decades, researchers have proposed more subtle forms of prejudice, such as modern and aversive prejudice, due to observations of changing societal and personal norms regarding the overt expression of prejudice (Dovidio and Gaertner, 2004; Gaertner and Dovidio, 1986; McConahay, 1986; McConahay, Hardee, and Batts, 1981; Swim, Aiken, Hall, and Hunter, 1995). In addition, methodological concerns regarding the use of self-report measures have led to significant advances in the area of indirect attitude measurement (e.g.,

Fazio, Jackson, Dunton, and Williams, 1995; Greenwald, McGhee, and Schwartz, 1998; Payne, Cheng, Govorun, and Stewart, 2005). These methodological advances have suggested potential dissociations between implicit and explicit prejudice, such that people may harbour negative implicit evaluations of social groups in the absence of negativity at the explicit level. For the most part, however, investigations into different forms and conceptualizations of prejudice have occurred relatively independently from each other.

Old-fashioned prejudice may best be conceptualized in terms of non-egalitarian beliefs, such as the endorsement of negative stereotypes, support for segregation and open discrimination, and belief in the inferiority of particular social groups (McConahay, 1986; McConahay et al., 1981; Swim et al., 1995). This form of prejudice has been labeled *old-fashioned* because, although it used to be acceptable, it is no longer fashionable or acceptable in most social circles (McConahay, 1986; McConahay et al., 1981). In fact, public opinion polls in North America have shown a steady decline in the negative evaluations of racial minority groups after World War II, mirrored by a steady increase in the endorsement of the goals of racial integration and equal treatment (Bobo, 2001; Dovidio and Gaertner, 2004). Interestingly, despite the significant decline in overtly expressed negativity, racial conflicts during this time showed only a moderate reduction (Dovidio and Gaertner, 2004). This disconnect is often attributed to individuals' inner conflict between egalitarian values and negative racial sentiments, which has given rise to theorizing of newer, more subtle forms of prejudice.

Modern prejudice is one influential concept that has been used to explain the subtle expression of prejudice, reflecting the conflict between egalitarianism and underlying negativity toward particular social groups. According to the theory of modern prejudice, negativity is only expressed overtly when it can be justified on non-prejudicial grounds, as this allows for the maintenance of an egalitarian and non-prejudiced self-image (McConahay, 1983, 1986). According to McConahay, such nonprejudiced justifications are based upon contextual factors such as ideological or situational ambiguity. Modern prejudice is exemplified in the case where people deny the existence of discrimination in the first place, and thus do not support public policies aimed at reducing discrimination (e.g., affirmative action). Those who endorse this position claim that they are not prejudiced, but that it is an empirical fact that discrimination no longer exists and that such policies are unfair and violate egalitarian principles (McConahay, 1986; Swim et al., 1995).

The concept of aversive prejudice also has been influential in explaining the subtle expression of prejudice, and is similarly based upon the conflict between egalitarianism and underlying negativity. The theory of aversive prejudice asserts that individuals generally acknowledge discrimination and believe that it is wrong, but nevertheless experience negative reactions toward particular social groups. This negativity is expressed, however, only when it can be justified on non-prejudicial grounds in order to protect one's egalitarian image (Dovidio and Gaertner, 2004; Gaertner and Dovidio, 1986). Despite its resemblance to modern prejudice, there are two important differences between these constructs (Dovidio and Gaertner, 2004). First, aversive prejudice is theorized to describe those who are politically liberal, whereas modern prejudice tends to describe those who are politically conservative. Second, although both modern and aversive prejudice are theorized to characterize those who endorse egalitarian, nonprejudiced beliefs, only modern prejudice is associated with a deliberate denial of discrimination.

Old-fashioned and modern prejudice are typically assessed via self-report. However, due to the susceptibility of such direct measures to socially desirable responding and introspective inaccessibility of attitudes, indirect measures have been developed to overcome issues of monitoring and control (e.g., Fazio et al., 1995; Greenwald et al., 1998; Payne et al., 2005). Whereas self-reported attitudes are expressed consciously with awareness and intent, indirectly assessed attitudes have been described as *implicit* in the sense that they may reflect “introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects” (Greenwald and Banaji, 1995, p. 8). Based on this definition, implicit prejudice is often conceptualized as an automatic association between social groups and negative valence (Rudman, 2004; see also Fazio, 1995). As such, implicit evaluations often dissociate from explicitly endorsed evaluations, in that people may show negative responses at the implicit level despite the absence of negativity at the explicit level. This claim is consistent with the findings of a recent meta-analysis, showing that the mean population correlation between implicit and explicit measures of group attitudes was only modest with an average, error-corrected correlation of 0.25 (Hofmann, Gawronski, Gschwendner, Le, and Schmitt, 2005). Such dissociations have been incorporated into the theory of aversive prejudice, such that indirect measures are assumed to tap the negative affective component of aversive prejudice, whereas direct measures are assumed to reflect the conscious, egalitarian component (Dovidio and Gaertner, 2004; Son Hing, Li, and Zanna, 2002).

Although each of the three forms of prejudice is concerned with the experience, expression, and consequences of prejudice, the conceptual and empirical relations between these different concepts has not been well specified. That is, to date, there has been no integrative model that incorporates all of these concepts into a general prejudice framework. In the present chapter, we review a new integrative model which hypothesizes that old-fashioned, modern, and aversive forms of prejudice share central components concerning spontaneously experienced negativity, egalitarianism-related nonprejudicial goals, and perceptions of discrimination (see Gawronski, Peters, et al., 2008). In integrating these forms of prejudice, however, it is important to first consider the general mechanisms underlying evaluations, namely the roles of associative (implicit) and propositional (explicit) processes, along with the notion of cognitive consistency.

## PROCESSES UNDERLYING EVALUATION

### Associative Versus Propositional Processes

In the integrative framework, two types of mental processes that underlie tendencies to respond positively or negatively to social groups are distinguished: associative and propositional processes (see Gawronski and Bodenhausen, 2006; Strack and Deutsch, 2004). The first kind of evaluative tendency is best described as a spontaneous affective reaction toward a given group, which depends on the particular associations that are activated in response to that group. The notion of affective reactions endorsed in this research is similar to Russell’s (2003) notion of core affect, in which affective reactions vary only in terms of valence (i.e., positivity or negativity) and arousal, but do not involve additional qualitative

distinctions (e.g., distinct emotions). The defining feature of such associative activation processes is that they are independent of what people believe to be true or false. That is, associations can be activated irrespective of whether a person considers the evaluations implied by these associations to be accurate or not. For example, being exposed to African Americans may activate negative associations even though an individual regards these associations to be inadequate or false (Devine, 1989).

The second type of evaluative tendency is reflected in endorsed evaluative judgments about a given social group, which represent the outcome of propositional validation processes. According to Gawronski and Bodenhausen (2006), processes of propositional reasoning aim to determine the validity of evaluations and beliefs by assessing their consistency with other relevant propositions (Jones and Gerard, 1967; Quine and Ullian, 1978). Drawing on a central assumption of Strack and Deutsch (2004), Gawronski and Bodenhausen (2006) proposed that people typically translate their spontaneous affective reaction to a given social group into propositional format (e.g., a negative affective reaction to a social group is translated into the proposition "I dislike that social group"). According to Gawronski and Bodenhausen (2006), the resulting proposition is then subject to syllogistic inferences that assess its validity. This assumption is in line with Zajonc's (1980) claim regarding the primacy of affect, in that affective responses are assumed to precede a cognitive appraisal of these responses (see also Russell, 2003). However, the most significant feature that distinguishes propositional from associative processes at a conceptual level is their dependency on truth values. Whereas the activation of associations can occur regardless of whether a person considers these associations to be true or false, processes of propositional reasoning are generally concerned with the validation of evaluations and beliefs. Moreover, whether or not the propositional evaluation implied by an affective reaction will be explicitly endorsed depends on the subjective validity of that evaluation, as determined by the consistency of this proposition with other salient, relevant propositions (Gawronski and Bodenhausen, 2006).

### **Cognitive Consistency**

One of the most important aspects of the present model is the notion of cognitive consistency. Gawronski and Bodenhausen (2006) argued that the desire to maintain a consistent system of beliefs determines whether people base their evaluative judgment of a social group on their spontaneous affective reaction to that group. According to Festinger (1957), two cognitions are inconsistent if one follows from the opposite of the other. Conversely, two cognitions are consistent if one does not follow from the opposite of the other. Since the logical relations presupposed in Festinger's (1957) definition require an assignment of truth values, cognitive (in)consistency can be regarded as inherently propositional (Gawronski and Strack, 2004; Gawronski, Strack, and Bodenhausen, in press). There is no logical relation between two cognitions unless they are regarded as either true or false. Such logical relations in conjunction with a given set of accepted propositions provide the basis for the (subjective) truth or falsity of other propositions, such as the evaluation implied by one's spontaneous affective reaction (Jones and Gerard, 1967; Quine and Ullian, 1978).

The perceived validity of a proposition—and thus of the evaluation implied by a spontaneous affective reaction—depends on the consistency of this proposition with other propositions that are considered to be relevant to the judgment at hand (Kruglanski, 1989). If the evaluation implied by a spontaneous affective reaction is consistent with other relevant propositions, it may be considered valid and thus may serve as the basis for an evaluative judgment. If, however, the evaluation implied by a spontaneous affective reaction is inconsistent with other relevant propositions, it may be considered invalid and thus may be rejected as a basis for an evaluative judgment. Most important, simply rejecting a spontaneous affective reaction for an evaluative judgment does not necessarily deactivate the associations that have led to this reaction (Deutsch, Gawronski, and Strack, 2006; Gawronski, Deutsch, Mbirkou, Seibt, and Strack, 2008; Gawronski and Strack, 2004). Accordingly, negative affective reactions to a particular social group tend to be unaffected by deliberate rejections of these reactions as a basis for evaluative judgments (see Butler et al., 2003; Gross, 1998). Hence, inconsistency-related rejections of spontaneous affective reactions often lead to dissociations, such that spontaneous affective reactions to a social group may be negative despite the absence of negativity in evaluative judgments (Gawronski and Strack, 2004; Gawronski et al., in press).

## RECONCEPTUALIZING FORMS OF PREJUDICE

The distinction between associative and propositional processes, along with the notion of cognitive consistency, is central to the general framework integrating different forms of prejudice (see Gawronski, Peters, et al., 2008). We equate explicit prejudice, as assessed through direct self-report measures, with the outcome of propositional processes (i.e., *endorsed evaluative judgments* of a social group). In contrast, implicit prejudice, as assessed by indirect attitude measures, is equated with the outcome of associative processes (i.e., *spontaneous affective reactions* to a social group). In addition, we argue that egalitarianism-related nonprejudicial goals and perceptions of discrimination—two integral components of old-fashioned, modern, and aversive prejudice—reflect propositions relevant to the endorsement of evaluative judgments about social groups. This theorizing allows us to integrate old-fashioned, modern, and aversive forms of prejudice, as outlined below.

Let us begin with the assumption that an individual is experiencing a negative affective reaction toward a member of a particular social group, which is driven by the valence of associations that are activated in response to this group. According to Gawronski and Bodenhausen's (2006) theorizing, whether this affective reaction results in a negative judgment about the social group member should depend on the consistency of this evaluation with other propositions that are considered to be relevant, such as egalitarianism-related nonprejudicial goals and perceived discrimination.<sup>1</sup>

---

<sup>1</sup> For the purpose of the present framework, nonprejudicial goals and perceived discrimination are viewed to be the most significant propositions relevant to the evaluation of social groups and their members. However, it is conceivable that there are other prejudice-related propositions that may contribute to (in)consistency within a person's system of beliefs in addition to the ones discussed in the present chapter.

More precisely, the resulting set of judgment-relevant propositions may include the following components (see Figure 1, Panel A):

- 1) "I dislike members of this social group",
- 2) "Negative evaluations of disadvantaged social groups are wrong", and
- 3) "Members of this social group are disadvantaged because of their group membership".

These three propositions are inconsistent with each other in that they cannot be endorsed simultaneously without violating the basic notion of cognitive consistency (Festinger, 1957). Proposition 1 is inconsistent with the joint implication of Propositions 2 and 3; Proposition 2 is inconsistent with the joint implication of Propositions 1 and 3; and Proposition 3 is inconsistent with the joint implication of Propositions 1 and 2. Thus, in order to avoid uncomfortable feelings arising from cognitive dissonance (Festinger, 1957), cognitive consistency may be maintained by rejecting (i.e., changing the truth value of) at least one of the three propositions (Gawronski and Strack, 2004; Gawronski et al., in press).<sup>2</sup> The three possible cases are illustrated in Figure 1 (Panels B, C, and D).

First, people may reject the proposition representing nonprejudicial goals (e.g., "I don't care about disadvantaged social groups"). In this case, people may still agree that the social group is disadvantaged. However, this belief does not result in a rejection of negative affective reactions as a basis for evaluative judgments, as negative judgments of disadvantaged social group members are considered acceptable. Thus, negative affective reactions to members of this social group may directly translate into negative judgments (see Figure 1, Panel B). This case represents central components of the theory of old-fashioned prejudice: people do not endorse nonprejudicial values and openly support discriminatory practices, resulting in overt negative evaluations (e.g., McConahay, 1986; Swim et al., 1995).

Second, people may reject the proposition representing perceptions of discrimination (e.g., "Members of this social group are not disadvantaged because of their group membership"). In this case, people may still hold strong nonprejudicial goals. However, these goals do not result in a rejection of negative affective reactions to members of disadvantaged social groups as a basis for evaluative judgments, as they are not considered to be targets of discrimination (e.g., Franco and Maass, 1999). Accordingly, negative affective reactions may directly translate into negative judgments (see Figure 1, Panel C). This case represents central components of the theory of modern prejudice: people deny the continued discrimination of disadvantaged social groups, thereby rationalizing negative reactions despite the endorsement of nonprejudicial goals (e.g., McConahay, 1986; Swim et al., 1995).

<sup>2</sup> An alternative strategy to resolve inconsistency is to "search for consonant information" (Festinger, 1957), which implies the addition of a new proposition rather than the rejection of an old proposition (Gawronski and Strack, 2004). This strategy is discussed in more detail in the section on Implications of the Model.

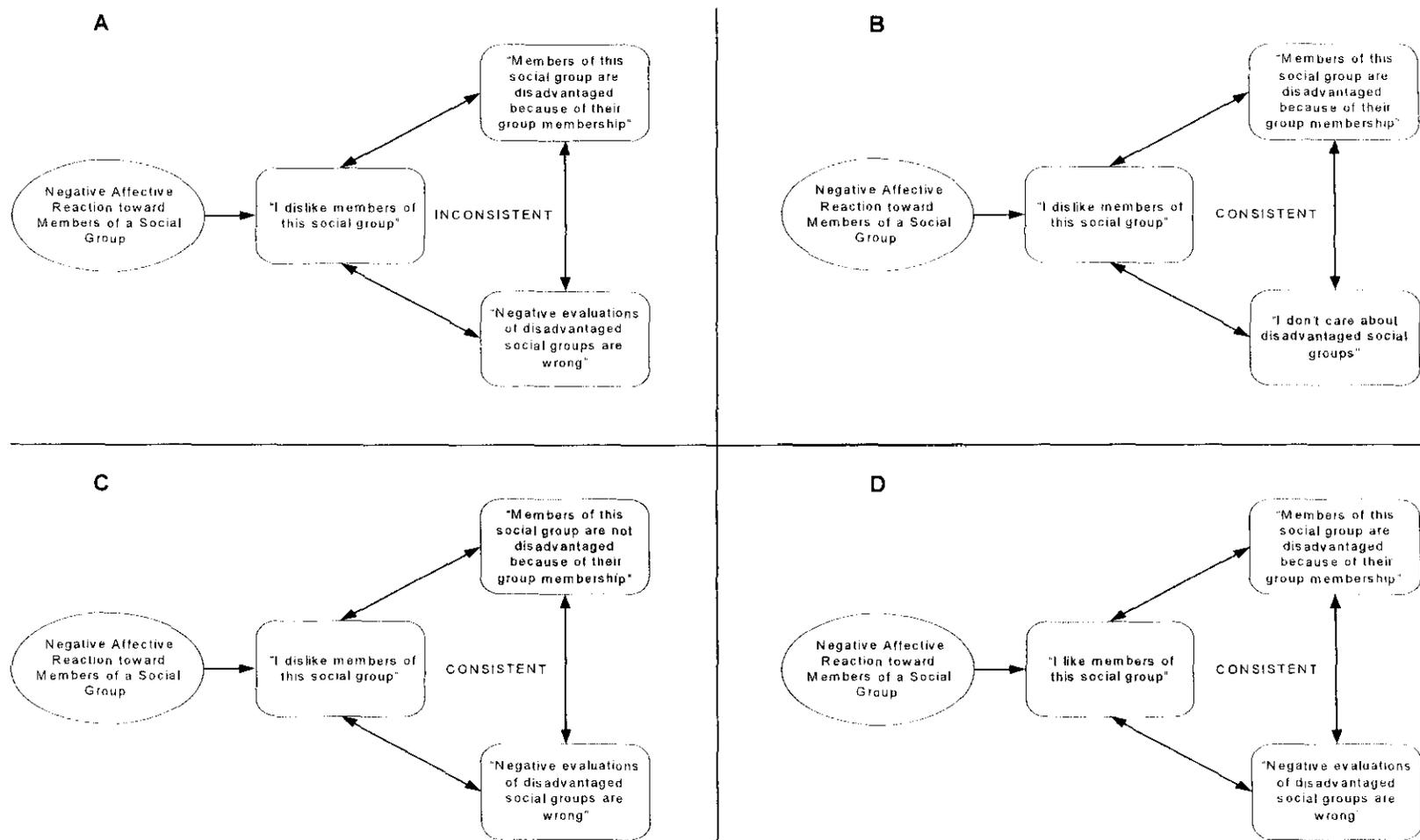


Figure 1. Interplay between spontaneous affective reactions (i.e., implicit prejudice), nonprejudicial goals, and perceptions of discrimination in the endorsement of negative evaluations of social groups (i.e., explicit prejudice). Panel A depicts the case of an inconsistent belief system resulting from negative affective reactions toward members of a social group; Panels B, C, and D depict consistent belief systems whereby the proposed components of the model reflect central components of the theories of old-fashioned, modern, and aversive forms of prejudice.

Third, people may reject their negative affective reactions as a valid basis for an evaluative judgment (e.g., "I like members of this social group"). Such a rejection may occur when people hold strong nonprejudicial goals and, at the same time, agree that members of this social group are disadvantaged by virtue of their group membership. In this case, negative affective reactions to members of this social group will *not* translate into negative judgments about this group (see Figure 1, Panel D). Rather, people's evaluative judgments should be relatively neutral (or even positive), irrespective of the valence of their affective reactions. Importantly, the mere rejection of affective reactions as a basis for evaluative judgments does not necessarily change their affective quality (Deutsch et al., 2006; Gawronski and Strack, 2004; Gawronski, Deutsch, et al., 2008; see also Butler et al., 2003; Gross, 1998). That is, even though negative affective reactions may not be reflected in evaluative judgments, affective reactions may still be negative. This case represents central components of the theory of aversive prejudice: people hold strong nonprejudiced values and believe that particular social groups are disadvantaged, but nevertheless experience negative feelings toward these groups even though these feelings are not reflected in negative judgments (e.g., Dovidio and Gaertner, 2004; Gaertner and Dovidio, 1986).

## EMPIRICAL EVIDENCE

Aside from integrating different forms of prejudice, the integrative model implies new predictions regarding the relation between implicit prejudice as assessed by indirect measures and explicit prejudice as assessed by direct measures. In particular, the model predicts that the endorsement of nonprejudicial goals and perceptions of discrimination interactively moderate the correspondence between implicit and explicit prejudice. More precisely, implicit and explicit prejudice should be highly correlated when either nonprejudicial goals or perceptions of discrimination are low. In these cases, negative affective reactions to a particular social group should serve as a basis for direct evaluative judgments of that group, resulting in high correlations between the two. Such direct translations should occur when people perceive a particular social group as facing discrimination, but do not care about disadvantaged social groups (central components of old-fashioned prejudice), or when people believe it is wrong to negatively evaluate disadvantaged social groups, but do not perceive members of a particular social group as facing discrimination (central components of modern prejudice). However, when both nonprejudicial goals and perceived discrimination are high, implicit and explicit prejudice should be unrelated (or perhaps negatively related). In this case, negative affective reactions to a particular social group should *not* serve as a basis for direct evaluative judgment of that group. Such dissociations between spontaneous affective reactions (implicit prejudice) and evaluative judgments (explicit prejudice) should occur when people believe it is wrong to negatively evaluate disadvantaged social groups and, at the same time, perceive the particular social group as facing discrimination (central components of aversive prejudice).

In summary, the integrative model implies that people should base their evaluative judgments about a particular social group on their spontaneous affective reactions toward this group when either (a) they do not endorse nonprejudicial goals, or (b) they do not believe that this group is a target of discrimination. In contrast, spontaneous affective reactions should *not*

be reflected (or may be negatively reflected) in evaluative judgments when people (a) hold strong nonprejudicial goals, and at the same time, (b) believe that the group is a target of discrimination.

Thus far, these predictions have been tested and supported in a series of three studies, using three different indirect attitude measures and two referent social groups. The first two studies tested the model within the domain of racial prejudice, utilizing two different indirect attitude measures: the Implicit Association Test (IAT; Greenwald et al., 1998; Olson and Fazio, 2004), and the Affect Misattribution Procedure (AMP; Payne et al., 2005). The third study extended the model to the domain of weight prejudice, utilizing a variant of Wittenbrink, Judd, and Park's (1997) sequential priming task (SPT) to assess implicit weight prejudice. In the following sections, we review the central findings of these studies.

### **Application to Racial Prejudice Using the Implicit Association Test**

The main goal of the first study (Gawronski, Peters, et al., 2008, Study 1) was to provide a test of the integrative framework by examining the interactive effects of nonprejudicial goals and perceptions of discrimination on the relation between implicit and explicit prejudice against Black people. In this study, implicit prejudice, or negative affective reactions resulting from spontaneously activated associations, was assessed with two variants of the Implicit Association Test (IAT): the standard variant proposed by Greenwald and colleagues (1998), and the personalized variant proposed by Olson and Fazio (2004). Explicit prejudice, or direct evaluative judgments, was assessed with a feeling thermometer scale (Esses et al., 1993). Both IAT scores and difference scores of the feeling thermometer ratings were calculated such that higher values indicated a stronger preference for Whites over Blacks. Nonprejudicial goals were assessed with Dunton and Fazio's (1997) Motivation to Control Prejudiced Reactions Scale (e.g., *"I get angry with myself when I have a thought or feeling that might be considered prejudiced"*). Perceptions of discrimination were assessed with McConahay's (1986; Table 2) Modern Racism Scale (e.g., *"How many Black people in Canada do you think miss out on jobs or promotions because of racial discrimination?"*). Both nonprejudicial goals and perceptions of discrimination were scored such that higher values indicated stronger nonprejudicial goals and higher perceived discrimination, respectively.<sup>3</sup>

As described earlier, the integrative model proposes that nonprejudicial goals and perceived discrimination of Blacks interactively determine whether negative affective reactions to Blacks (implicit prejudice) lead to the endorsement of negative evaluations of Blacks (explicit prejudice). More precisely, it is predicted that implicit prejudice should be directly related to explicit prejudice when either nonprejudicial goals or perceptions of discrimination are low. However, implicit prejudice should be unrelated, or possibly negatively related, to explicit prejudice when both nonprejudicial goals and perceptions of discrimination are high. In other words, the integrative framework implies a three-way interaction in the prediction of feeling thermometer scores, such that IAT scores should predict feeling thermometer scores when nonprejudicial goals or perceptions of

<sup>3</sup> Note that the present coding of perceived discrimination is opposite to the standard coding of the Modern Racism Scale, in which high scores typically reflect a low level of perceived discrimination.

discrimination are weak. IAT scores should be unrelated (or potentially negatively related) to feeling thermometer scores when nonprejudicial goals are strong, and, at the same time, perceived discrimination is high.

To test these predictions, explicit preference for Whites over Blacks was regressed on implicit preference for Whites over Blacks, nonprejudicial goals, perceived discrimination, and all of their possible interactions.

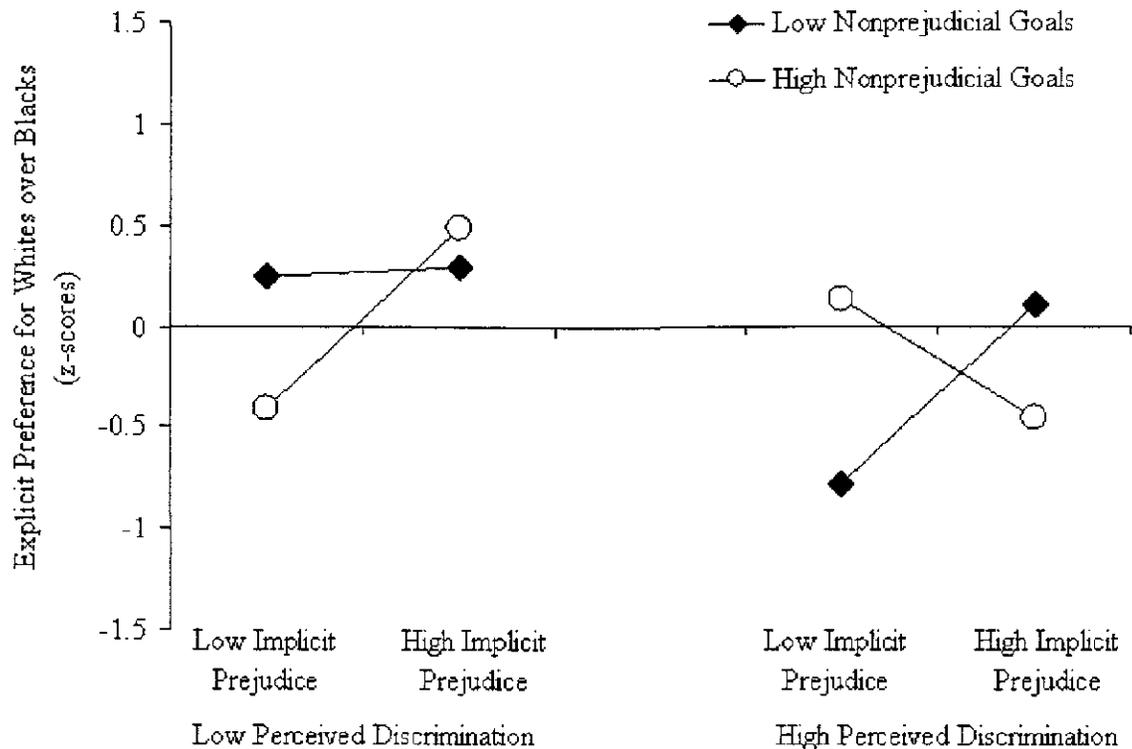


Figure 2. Relation between implicit racial prejudice assessed by the Implicit Association Test and explicit racial prejudice as a function of nonprejudicial goals and perceptions of discrimination (Gawronski, Peters, et al., 2008, Study 1).

This regression analysis revealed a significant three-way interaction, which is shown in Figure 2. Consistent with predictions, implicit prejudice was positively related to explicit prejudice when perceived discrimination was high but nonprejudicial goals were weak, reflecting the central components of old-fashioned prejudice. Moreover, implicit prejudice was positively related to explicit prejudice when nonprejudicial goals were strong but perceived discrimination was weak, reflecting the central components of modern prejudice.

In contrast, implicit prejudice showed a tendency for a negative relation to explicit prejudice when perceived discrimination was high and, at the same time, nonprejudicial goals were strong, reflecting the central components of aversive prejudice. Unexpectedly, participants low in perceived discrimination and holding weak nonprejudicial goals showed independence between their levels of implicit and explicit prejudice. That is, participants' implicit preference for Whites over Blacks was not significantly related to their explicit preference for Whites over Blacks. The pattern of findings did not differ between the standard and personalized variants of the IAT.

### **Application to Racial Prejudice Using the Affect Misattribution Procedure**

The main goal of Gawronski, Peters, and colleagues' (2008) second study was to determine whether the findings of their first study would replicate with a different measure of implicit prejudice. This goal was stimulated by two issues. First, it seemed important to determine whether the obtained independence of implicit and explicit prejudice for participants with low scores on both nonprejudicial goals and perceptions of discrimination was simply a random effect or a replicable, systematic effect. Second, even though similar effects using the standard and personalized IAT were obtained in Study 1, the general procedure of the IAT has been criticized on several grounds (e.g., Brendl, Markman, and Messner, 2001; Mierke and Klauer, 2003; Rothermund and Wentura, 2004), suggesting that replication with an alternative implicit measure would be beneficial.

In this study, implicit prejudice was assessed with Payne and colleagues' (2005) Affect Misattribution Procedure (AMP). For this task, participants were briefly presented with a Black or White face, which was followed by a Chinese character of neutral valence that was backward masked. Participants then indicated whether they considered the presented Chinese character as more pleasant or less pleasant than the average Chinese character (see also Murphy and Zajonc, 1993). Affective priming effects in this task are reflected in higher proportions of more pleasant responses after priming with positive (e.g., White face) as compared to negative (e.g., Black face) stimuli, and in higher proportions of less pleasant responses after priming with negative as compared to positive stimuli. AMP scores were calculated such that higher values indicated a stronger implicit preference for Whites over Blacks. Explicit prejudice, nonprejudicial goals, and perceived discrimination were assessed with the same measures as employed in the first study.

Explicit preference for Whites over Blacks was regressed onto implicit preference for Whites over Blacks, nonprejudicial goals, perceived discrimination, and all of their possible interactions. This regression analysis again revealed a significant three-way interaction. The specific pattern of this interaction is depicted in Figure 3.

Replicating the pattern obtained in Gawronski, Peters, et al.'s (2008) first study, implicit prejudice was positively related to explicit prejudice when perceived discrimination of Blacks was high but nonprejudicial goals were weak, reflecting the central components of old-fashioned prejudice. Implicit prejudice was positively related to explicit prejudice when nonprejudicial goals were strong but perceived discrimination of Blacks was low, reflecting the central components of modern prejudice. In contrast, implicit prejudice was unrelated to explicit prejudice when perceived discrimination of Blacks was high and, at the same time, nonprejudicial goals were strong, reflecting the central components of aversive prejudice. Replicating the unexpected pattern obtained in the first study, explicit and implicit prejudice were not significantly related for participants with low perceived discrimination of Blacks and weak nonprejudicial goals.

### **Application to Weight Prejudice Using the Sequential Priming Task**

The main goals of our third study (Brochu, Esses, and Gawronski, 2008) were to test whether the obtained results would replicate using another indirect measure of implicit

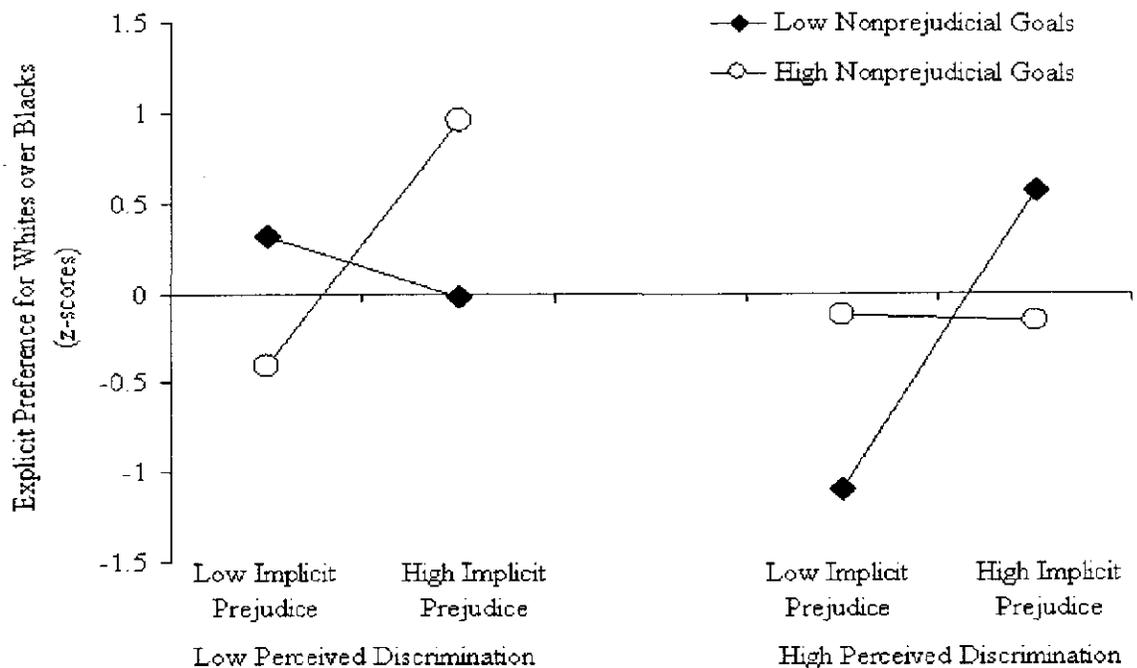


Figure 3. Relation between implicit racial prejudice assessed by the Affect Misattribution Procedure and explicit racial prejudice as a function of nonprejudicial goals and perceptions of discrimination (Gawronski, Peters, et al., 2008, Study 2).

prejudice, and to extend the model to other prejudicial domains. The theories of old-fashioned, modern, and aversive prejudice were all stimulated and tested primarily within the domain of race relations. Nonetheless, it is important to test the generalizability of the integrative framework to other targets of prejudice. In this instance, we examined weight prejudice because subtle forms of prejudice exhibited toward this social group have not been widely examined (for a notable exception, see Brochu, 2007), and weight bias may be considered one of the last acceptable forms of discrimination (Puhl and Brownell, 2001).

In this study, implicit weight prejudice was assessed using a variant of Wittenbrink and colleagues' (1997) sequential priming task (SPT). In this paradigm, participants were subliminally presented with the terms *OVER-WEIGHT* or *NORMAL-WEIGHT*, which were then backward masked. After the masking stimulus, participants were presented with a letter string which was a positive, negative, or neutral word, or a meaningless non-word. Participants were instructed to indicate as quickly as possible whether the presented letter string was a meaningful word or not. Priming effects in this task are reflected in faster response times in indicating that a negative word is a word, and slower response times in indicating that a positive word is a word, after being primed with *OVER-WEIGHT* than after being primed with *NORMAL-WEIGHT*. Scores on this task were calculated such that higher values indicated more negative implicit evaluations of the overweight.

Explicit prejudice was assessed with Crandall's (1994) Anti-fat Attitudes Questionnaire – Dislike subscale (e.g., “I really don't like overweight people much”), with the term fat replaced by overweight. Nonprejudicial goals were again assessed with Dunton and Fazio's (1997) Motivation to Control Prejudiced Reactions Scale. Perceptions of discrimination were assessed with 11 items developed for the purpose of this study (e.g., “Overweight people are victims of discrimination”; see Brochu et al., 2008). Higher scores on these scales indicate

greater endorsement of explicit negative evaluations of the overweight, nonprejudicial goals, and perceived discrimination of the overweight, respectively.

In order to test the model within the domain of weight prejudice, explicit evaluations of the overweight were regressed on implicit evaluations of the overweight, nonprejudicial goals, perceived discrimination, and all of their possible interactions.

This analysis again revealed a significant three-way interaction (see Figure 4). Replicating the pattern of the first two studies, implicit evaluations were positively related with explicit evaluations when nonprejudicial goals were low and perceived discrimination was high, reflecting the central components of old-fashioned prejudice. Implicit evaluations tended to be positively related to explicit evaluations when nonprejudicial goals were high and perceived discrimination was low, reflecting the central components of modern prejudice. In addition, implicit evaluations were negatively related with explicit evaluations when both nonprejudicial goals and perceived discrimination were high, reflecting the central components of aversive prejudice. Finally, replicating the unexpected finding of the previous studies, implicit evaluations were not significantly related to explicit evaluations when both nonprejudicial goals and perceived discrimination were low.

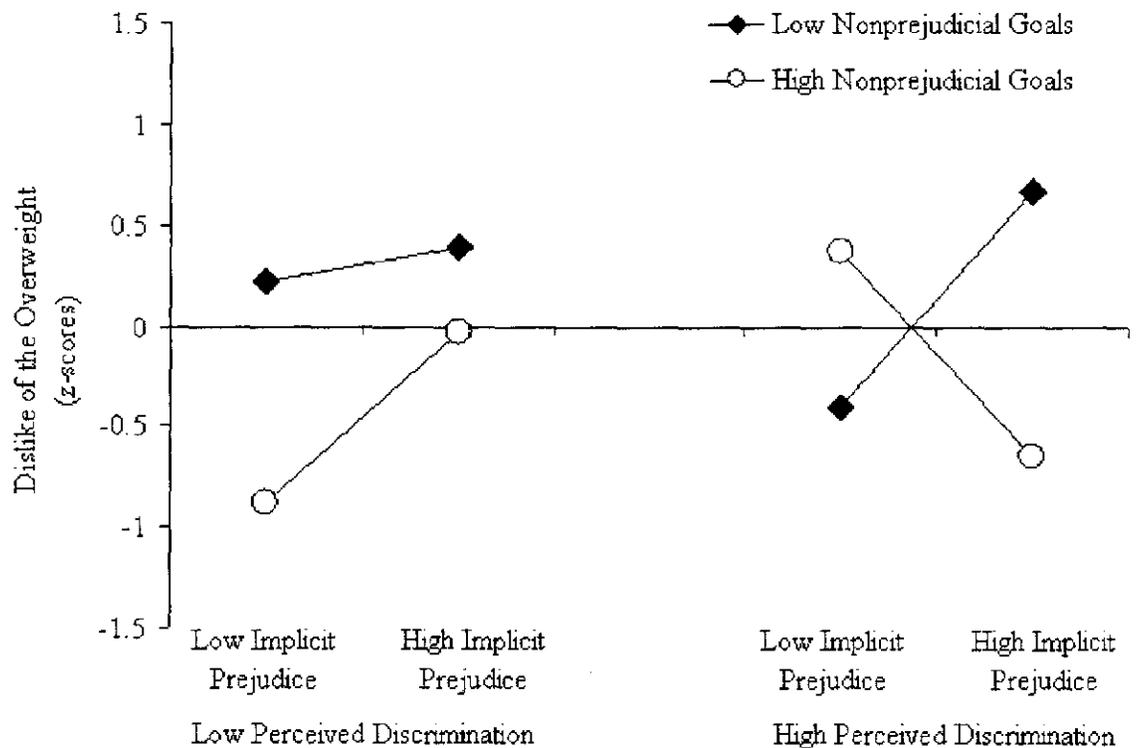


Figure 4. Relation between implicit weight prejudice assessed by the Sequential Priming Task and explicit weight prejudice as a function of nonprejudicial goals and perceptions of discrimination (Brochu et al., 2008).

## SUMMARY

In support of the integrated framework of prejudice (Gawronski, Peters, et al., 2008), the findings indicate that the correspondence between implicit and explicit prejudice is

determined by the interplay between (a) egalitarianism-related nonprejudicial goals (i.e., the extent to which one believes that negative evaluations of social groups are wrong) and (b) perceptions of discrimination (i.e., the extent to which one believes that a specific social group is a target of systematic discrimination). The primary predictions of this framework were supported across three studies employing three different indirect measures of attitudes and assessing prejudice toward two different social groups. Further, one unexpected finding involving participants who do not endorse nonprejudicial goals or perceive discrimination consistently revealed itself across the three studies. Taken together, the reviewed evidence indicates that: (1) explicit evaluations of a social group reflect implicit evaluations when nonprejudicial goals are weak and, at the same time, perceived discrimination is high; (2) explicit evaluations of a social group reflect implicit evaluations when nonprejudicial goals are strong and, at the same time, perceived discrimination is low; (3) explicit evaluations of a social group do not reflect, or may negatively reflect, implicit evaluations when nonprejudicial goals are strong and perceived discrimination is high; and (4) explicit evaluations of a social group do not reflect implicit evaluations when nonprejudicial goals are weak and perceived discrimination is low.

Each of these four response patterns has interesting theoretical implications and relations to different forms of prejudice. In the first case, the positive relation between implicit and explicit evaluations of a social group when nonprejudicial goals are weak and, at the same time, perceived discrimination is high, represents central components of old-fashioned prejudice, which is an overt and blatant form of prejudice that dictates open support for discrimination and segregation (McConahay, 1986; McConahay et al., 1981; Swim et al., 1995). In the second case, the positive relation between implicit and explicit evaluations when nonprejudicial goals are strong and perceived discrimination is low, represents central components of modern prejudice, which is a covert and subtle form of prejudice that denies discrimination while maintaining an image of egalitarianism (McConahay, 1986; Swim et al., 1995). In the third case, the unrelated or negative relation between implicit and explicit evaluations when nonprejudicial goals are strong and perceived discrimination is high, represents central components of aversive prejudice, which is a covert and subtle form of prejudice characterized by a conflict between negative affective reactions and egalitarian beliefs (Dovidio and Gaertner, 2004). According to Dovidio and Gaertner's theorizing, "aversive racists either will treat Blacks and Whites equally, or they will respond even more favorably to Blacks than to Whites" (p. 7), which is reflected in the tendency for negative relations between implicit and explicit prejudice in the current studies. However, when a justification is available that allows for the expression of the affective negativity on nonprejudicial grounds, it is expected that implicit and explicit prejudice will relate positively with each other, which is discussed further in the next section. Finally, the nonsignificant relation observed between implicit and explicit prejudice when nonprejudicial goals are weak and perceived discrimination is low does not directly map onto a previously described form of prejudice. Implications of this fourth case are also discussed in more detail in the next section.

## IMPLICATIONS OF THE MODEL

The reconceptualization of old-fashioned, modern, and aversive prejudice in terms of their underlying associative and propositional processes has many implications for prejudice research. Perhaps most interesting is the interactive influence of nonprejudicial goals and perceptions of discrimination on the relation between implicit and explicit evaluations, which integrates central components of the theories of old-fashioned, modern, and aversive forms of prejudice. However, the integrative framework also has many potential implications for research conducted in the areas of motivation to control prejudice, weight bias, and prejudice reduction. In addition, the model suggests other potential avenues for avoiding cognitive inconsistency in social group evaluation, and sheds light on the controversy surrounding theorizing regarding different forms of prejudice.

### Motivation to Control Prejudice

Previous studies using Dunton and Fazio's (1997) Motivation to Control Prejudiced Reactions Scale have demonstrated that implicit and explicit prejudice are significantly correlated only for participants low in motivation to control prejudice, but not for participants high in motivation to control prejudice (e.g., Fazio et al., 1995; Gawronski, Geschke, and Banse, 2003; Payne et al., 2005). The studies described in the present chapter extend such findings by focusing on the *interactive* impact of motivation to control prejudice, conceptualized here as nonprejudicial goals, and perceived discrimination on the relation between implicit and explicit prejudice. Most notably, we failed to replicate the simple moderation of motivation to control prejudice in all three studies reviewed in this chapter. Given the integrative framework of prejudice, however, this failed replication is not all that surprising. A central assumption of the model is that nonprejudicial goals and perceptions of discrimination *interactively* determine the relation between implicit and explicit prejudice. Specifically, the model suggests that high levels of motivation to control prejudice, or strong nonprejudicial goals, are not sufficient to reduce the impact of negative affective reactions on the overt endorsement of negative evaluations. Rather, this reduction requires high levels of both nonprejudicial goals *and* perceptions of discrimination. As such, the moderating influence of motivation to control prejudice found in previous studies may have depended on contingent characteristics of the sample, namely high levels of perceived discrimination.

### Alternative Strategies for Maintaining Cognitive Consistency

In the present chapter, we focused primarily on the maintenance of cognitive consistency by *rejecting* relevant propositions. However, as alluded to earlier, inconsistency also can be resolved by *searching* for consonant information (or an additional proposition) that resolves the inconsistency (Festinger, 1957). In research on cognitive dissonance, additional information of this kind is often represented by situational explanations for counterattitudinal behavior (e.g., Festinger and Carlsmith, 1959; Gawronski and Strack, 2004). In the current case of prejudice, one could argue that people may resolve the inconsistency between

negative evaluations, nonprejudicial goals, and perceived discrimination by searching for a situational explanation that could justify a negative evaluation. Indeed, situationally-based justification processes are theorized to play a crucial role in research on modern expressions of prejudice (Crandall and Eshleman, 2003; Dovidio and Gaertner, 2004; Gaertner and Dovidio, 1986; McConahay, 1983, 1986). These theories suggest that people are most likely to express their underlying negative attitudes toward various social groups when there is a readily available justification for doing so. For example, Dovidio and Gaertner (2000) found that Black and White job applicants were similarly recommended for a job when their qualifications were strong or weak; when their qualifications were ambiguous, however, Black job applicants were recommended significantly less strongly than White job applicants. In other words, participants in this study only expressed their negative reaction to Blacks when it could be justified on non-discriminatory grounds (i.e., ambiguous qualifications that could be interpreted in a variety of ways).

Applied to the integrative framework, the possibility of resolving cognitive inconsistency via consonant information should result in a correspondence between spontaneous affective reactions and direct evaluative judgments. That is, implicit negativity should result in explicit negativity when a potential inconsistency between negative affective reactions, nonprejudicial goals, and perceived discrimination can be resolved by means of another proposition (i.e., a justification that appears to be non-prejudicial is available). For example, in the case of weight bias, this framework suggests that implicit negativity is likely to be expressed explicitly when nonprejudicial goals and perceptions of discrimination are high, and, at the same time, the belief that overweight individuals are responsible for their excess weight is endorsed (an additional relevant proposition). However, negative affective reactions should be unrelated (or negatively related) to evaluative judgments about the group when such a proposition is not available or accessible, and inconsistency is resolved by rejecting the negative affective reaction.

### **The Fourth Case**

One unexpected finding revealed across all three studies was an observed independence between spontaneous affective reactions and endorsed evaluative judgments for participants who reported weak nonprejudicial goals and weak perceptions of discrimination. For this group of participants, implicit and explicit prejudice were generally unrelated. One immediate question raised by this finding is what may distinguish this group of participants in evidencing levels of implicit and explicit prejudice that did not correspond with each other.

In the third study reviewed in this chapter—which examined the integrative prejudice framework within the domain of weight prejudice (Brochu et al., 2008)—we had the opportunity to relate the observed dissociation to several other variables that may possibly account for this effect. However, none of the variables included in these analyses showed significant relations to this pattern, including social dominance orientation (Pratto, Sidanius, Stallworth, and Malle, 1994), right-wing authoritarianism (Altemeyer, 1996), Protestant work ethic (Katz and Hass, 1988), belief in a just world (Rubin and Peplau, 1975), and religious fundamentalism (Altemeyer and Hunsberger, 1992). Thus, future research is needed to examine additional features of this group of participants that may contribute to the observed dissociation.

One promising avenue of inquiry is to examine the structure of participants' attitudes, in particular whether they are more cognitively- or affectively-based. Huskinson and Haddock (2004) have demonstrated that there is wide variability in the relations between evaluations of attitude objects and affective and cognitive responses to the same attitude objects. Further, they demonstrated that this variability could be at least partially explained by individual differences in the Need for Affect (Maio and Esses, 2001) and the Need to Evaluate (Jarvis and Petty, 1996). That is, participants who evidenced a strong relation between their evaluative and affective responses tended to score higher on the Need for Affect, whereas participants who evidenced a strong relation between both their evaluative and affective responses, and evaluative and cognitive responses, tended to score higher on the Need to Evaluate. Thus, based upon this research, it is possible that our fourth group of participants, who showed a dissociation between their implicit and explicit attitudes, are characterized by low Need for Affect and/or Need to Evaluate. Future research should examine this possibility.

### **Modern Weight Prejudice**

The overt expression of weight bias is often described as the last socially accepted form of discrimination (Crandall, 1994; Puhl and Brownell, 2001). Although the pervasive and profound nature of weight bias is well-demonstrated (Brownell, Puhl, Schwartz, and Rudd, 2005; Puhl and Brownell, 2001, 2003), recent evidence has shown that the expression of weight bias may be fading (Brochu and Morrison, 2007; King, Shapiro, Hebl, Singletary, and Turner, 2006). For example, using a customer service paradigm, King and colleagues (2006) demonstrated that formal discrimination against overweight shoppers did not occur. That is, both average weight and overweight shoppers were greeted by sales representatives and offered help in searching for items in the store. However, interpersonal discrimination (operationalized as less smiling, eye contact, and friendliness, greater rudeness, and shorter interaction time) was only observed toward overweight shoppers who provided information that allowed the sales representatives to hold them responsible for their weight (e.g., not on a diet, did not exercise). This research demonstrates that people may need to justify the expression of weight prejudice, suggesting that its expression may not be as socially acceptable as it once was. The model outlined in the present chapter further suggests that nonprejudicial goals and perceptions of discrimination may play a significant role in this process. This assumption is consistent with Brochu and colleagues' (2008) findings which showed patterns of prejudicial responses corresponding to those obtained for racial prejudice.

### **A New Perspective on Subtle Forms of Prejudice**

Throughout this chapter, the constructs of interest were described as reflecting "different forms of prejudice." It is important to note, however, that this interpretation has been the subject of serious controversy in social psychology. For example, research in the tradition of modern prejudice has been criticized for inferring prejudice from measures that may simply reflect conservative political opinions (Sniderman and Tetlock, 1986). In a similar vein, research on implicit prejudice has been criticized for inferring prejudice from measures that may reflect cultural stereotypes rather than "personal animus" (Arkes and Tetlock, 2004). The

present framework offers a new perspective on these controversies by specifying the relation between several different concepts.

The integrative model agrees with previous criticism by arguing that measures of implicit prejudice and modern variants of prejudice do not tap the endorsement of negative evaluations of social groups. That is, measures of implicit prejudice and modern variants of prejudice do not directly assess individuals' evaluations of social groups and, thus, may be considered indirect in nature. In keeping with definitions proposed by De Houwer (2006), such an endorsement can only be assessed with measures implying a direct evaluation of these groups (e.g., attitude thermometers or likeability ratings). The present conceptualization also agrees with the previous criticism that neither the denial of discrimination nor implicit negativity is sufficient for the endorsement of negative evaluations. However, the present conceptualization disagrees with the previous criticism in arguing that both perceived discrimination and implicit negativity systematically contribute to the endorsement of negative evaluations of social groups. More precisely, the results of the research described in this chapter indicate that negative affective reactions (as assessed by measures of implicit prejudice) directly translate into negative judgments of social groups when either nonprejudicial goals are weak or perceived discrimination is low. Negative affective reactions do not translate into negative judgments when nonprejudicial goals are strong and, at the same time, perceived discrimination is high. Thus, theoretical controversies regarding the ontological nature of different forms of prejudice could possibly be resolved by focusing on the specific relations between the proposed concepts. In addition, incorporating the notion of cognitive consistency (Festinger, 1957; Gawronski et al., in press) and recent theorizing on associative and propositional processes (Gawronski and Bodenhausen, 2006; Strack and Deutsch, 2004) may provide deeper insights into the underlying dynamics of different forms of prejudice.

### **Strategies for Prejudice Reduction**

The integration of old-fashioned, modern, and aversive forms of prejudice also has important implications for prejudice reduction. The primary implication of the model is that attempts to reduce prejudice need to target three components *simultaneously*. That is, they should try to (a) enhance nonprejudicial goals, (b) increase awareness of discrimination, and (c) reduce automatic negative reactions. Strategies certainly exist that *independently* seek to directly or indirectly accomplish these goals. For example, work on the common ingroup identity model, which encourages individuals to see the common human bonds that connect people, may serve to foster growth of nonprejudicial goals (Gaertner et al., 2000). Further, providing individuals with information regarding the experiences of particular social groups may reduce ignorance about the existence of discrimination (Stephan and Stephan, 2000). Moreover, research has demonstrated the malleability of implicit attitudes by means of extended training in counterstereotypical responding (e.g., Gawronski, Deutsch, et al., 2008; Kawakami, Dovidio, Moll, Hermsen, and Russin, 2000) and evaluative conditioning (e.g., Olson and Fazio, 2006).

However, the model indicates that a caveat must be taken into consideration when employing such strategies. The findings reviewed in the present chapter suggest that attempts to reduce prejudice may be unsuccessful as long as they ignore the underlying dynamics of

associative and propositional processes (Gawronski and Bodenhausen, 2006; Strack and Deutsch, 2004). For example, simply enhancing nonprejudicial goals may leave negative judgments of social groups unaffected when these groups are not considered to be targets of discrimination. In a similar vein, enhancing people's awareness of the continued discrimination of particular social groups may be unsuccessful in reducing the endorsement of negative evaluations when nonprejudicial goals are weak. Attempts to change evaluative associations—and thus spontaneous affective reactions—may leave the endorsement of negative evaluations unaffected when nonprejudicial goals are weak and perceived discrimination is low. Thus, strategies that are directed at all of the relevant components simultaneously may be most successful in reducing prejudice (e.g., Rudman, Ashmore, and Gary, 2001).

## CONCLUSION

Research on prejudice is currently characterized by the existence of diverse concepts (e.g., implicit prejudice, old-fashioned prejudice, modern prejudice, aversive prejudice) that are not well integrated from a theoretical perspective. In this chapter, we reviewed a new theoretical framework that reconceptualizes these forms of prejudice in terms of their underlying associative and propositional processes (Gawronski, Peters, et al., 2008). A central implication of this model is that the correspondence between implicit and explicit prejudice is determined by the consistency between direct evaluative judgments, nonprejudicial goals, and perceptions of discrimination. A significant advance of this reconceptualization is that it (a) articulates the specific relations among various forms of prejudice, (b) can be applied to a variety of target groups, and (c) contributes to our understanding of the nature of, and the relations between, implicit and explicit prejudice. Needless to say, these issues are crucial when it comes to interpreting the findings of basic research so that they can be effectively applied to reduce the various forms of prejudice that are still prevalent in modern societies. Based on this conclusion, we hope that our theorizing will prove practical in reducing different forms of prejudice.

## REFERENCES

- Altemeyer, B. (1996). *The authoritarian specter*. Cambridge: Harvard University Press.
- Altemeyer, B., and Hunsberger, B. (1992). Authoritarianism, religious fundamentalism, quest, and prejudice. *The International Journal for the Psychology of Religion*, 2, 113-133.
- Arkes, H. R., and Tetlock, P. E. (2004). Attributions of implicit prejudice, or "Would Jesse Jackson 'fail' the Implicit Association Test?" *Psychological Inquiry*, 15, 257-278.
- Bobo, L. D. (2001). Racial attitudes and relations at the close of the twentieth century. In N. J. Smelser, W. J. Wilson, and F. M. Mitchell (Eds.), *America becoming: Racial trends and their consequences* (Vol. 1, pp. 264-301). Washington: National Academy Press.

- Brendl, C. M., Markman, A. B., and Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology, 81*, 760-773.
- Brochu, P. M. (2007). *Old-fashioned and modern prejudice against the overweight: Scale development and validation*. Unpublished master's thesis, University of Western Ontario, London, Ontario, Canada.
- Brochu, P. M., Esses, V. M., and Gawronski, B. (2008). *Cognitive consistency in prejudice against the overweight: The interplay of implicit evaluations, nonprejudicial goals, and perceived discrimination*. Manuscript in preparation.
- Brochu, P. M., and Morrison, M. A. (2007). Implicit and explicit prejudice toward overweight and average weight men and women: Testing their correspondence and relation to behavioural intentions. *Journal of Social Psychology, 147*, 681-706.
- Brownell, K. D., Puhl, R. M., Schwartz, M. B., and Rudd, L. (2005). *Weight bias: Nature, consequences, and remedies*. New York: Guilford Press.
- Butler, E. A., Egloff, B., Wilhelm, F. H., Smith, N. C., Erickson, E. A., and Gross, J. J. (2003). The social consequences of expressive suppression. *Emotion, 3*, 48-67.
- Crandall, C. S. (1994). Prejudice against fat people: Ideology and self-interest. *Journal of Personality and Social Psychology, 66*, 882-894.
- Crandall, C. S., and Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin, 129*, 414-446.
- De Houwer, J. (2006). What are implicit measures and why are we using them? In R. W. Wiers and A. W. Stacy (Eds.), *The handbook of implicit cognition and addiction* (pp. 11-28). Thousand Oaks, CA: Sage.
- Deutsch, R., Gawronski, B., and Strack, F. (2006). At the boundaries of automaticity: Negation as reflective operation. *Journal of Personality and Social Psychology, 91*, 385-405.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5-18.
- Dovidio, J. F., and Gaertner, S. L. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological Science, 11*, 315-319.
- Dovidio, J. F., and Gaertner, S. L. (2004). Aversive racism. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 36, pp. 1-52). New York: Elsevier.
- Dunton, B. C., and Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23*, 316-326.
- Esses, V. M., Haddock, G., and Zanna, M. P. (1993). Values, stereotypes, and emotions as determinants of intergroup attitudes. In D. M. Mackie and D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping: Interactive processes in group perception* (pp. 137-166). San Diego: Academic Press.
- Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty and J. A. Krosnick (Eds.), *Attitude strength* (pp. 247-282). Mahwah, NJ: Erlbaum.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., and Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013-1027.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston: Row Peterson.

- Festinger, L., and Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology, 58*, 203-210.
- Franco, F. M., and Maass, A. (1999). Intentional control over prejudice: When the choice of the measure matters. *European Journal of Social Psychology, 29*, 469-477.
- Gaertner, S. L., and Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio and S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61-89). Orlando: Academic Press.
- Gaertner, S. L., Dovidio, J. F., Nier, J. A., Banker, B. S., Ward, C. M., Houlette, M., et al. (2000). The common ingroup identity model for reducing intergroup bias: Progress and challenges. In D. Capozza and R. Brown (Eds.), *Social identity processes* (pp. 133-148). London: Sage.
- Gawronski, B., and Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 692-731.
- Gawronski, B., Deutsch, R., Mbirkou, S., Seibt, B., and Strack, F. (2008). When "just say no" is not enough: Affirmation versus negation training and the reduction of automatic stereotype activation. *Journal of Experimental Social Psychology, 44*, 370-377.
- Gawronski, B., Geschke, D., and Banse, R. (2003). Implicit bias in impression formation: Associations influence the construal of individuating information. *European Journal of Social Psychology, 33*, 573-589.
- Gawronski, B., Peters, K. R., Brochu, P. M., and Strack, F. (2008). Understanding the relations between different forms of racial prejudice: A cognitive consistency perspective. *Personality and Social Psychology Bulletin, 34*, 648-665.
- Gawronski, B., and Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology, 40*, 535-542.
- Gawronski, B., Strack, F., and Bodenhausen, G. V. (in press). Attitudes and cognitive consistency: The role of associative and propositional processes. In R. E. Petty, R. H. Fazio, and P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures*. Mahwah, NJ: Erlbaum.
- Greenwald, A. G., and Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27.
- Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464-1480.
- Gross, J. J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology, 74*, 224-237.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., and Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*, 1369-1385.
- Huskinson, T. L. H., and Haddock, G. (2004). Individual differences in attitude structure: Variance in the chronic reliance on affective and cognitive information. *Journal of Experimental Social Psychology, 40*, 82-90.
- Jarvis, W. B. G., and Petty, R. E. (1996). The need to evaluate. *Journal of Personality and Social Psychology, 70*, 172-194.

- Jones, E. E., and Gerard, H. B. (1967). *Foundations of social psychology*. New York: Wiley.
- Katz, I., and Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology*, 55, 893-905.
- Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., and Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, 78, 871-888.
- King, E. B., Shapiro, J. R., Hebl, M. R., Singletary, S. L., and Turner, S. (2006). The stigma of obesity in customer service: A mechanism for remediation and bottom-line consequences of interpersonal discrimination. *Journal of Applied Psychology*, 91, 579-593.
- Kruglanski, A. W. (1989). *Lay epistemics and human knowledge: Cognitive and motivational bases*. New York: Plenum Press.
- Maio, G. R., and Esses, V. M. (2001). The need for affect: Individual differences in the motivation to approach or avoid emotions. *Journal of Personality*, 69, 89-130.
- Maio, G. R., Esses, V. M., and Bell, D. W. (1994). The formation of attitudes toward new immigrant groups. *Journal of Applied Social Psychology*, 24, 1762-1776.
- McConahay, J. B. (1983). Modern racism and modern discrimination: The effects of race, racial attitudes, and context on simulated hiring decisions. *Personality and Social Psychology Bulletin*, 9, 551-558.
- McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. F. Dovidio and S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91-125). Orlando: Academic Press.
- McConahay, J. B., Hardee, B. B., and Batts, V. (1981). Has racism declined in America? It depends on who is asking and what is asked. *Journal of Conflict Resolution*, 25, 563-579.
- Mierke, J., and Klauer, K. C. (2003). Method-specific variance in the Implicit Association Test. *Journal of Personality and Social Psychology*, 85, 1180-1192.
- Murphy, S. T., and Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposure. *Journal of Personality and Social Psychology*, 64, 723-739.
- Olson, M. A., and Fazio, R. H. (2004). Reducing the influence of extra-personal associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology*, 86, 653-667.
- Olson, M. A., and Fazio, R. H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin*, 32, 421-433.
- Payne, B. K., Cheng, C. M., Govorun, O., and Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, 89, 277-293.
- Pratto, F., Sidanius, J., Stallworth, L. M., and Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, 67, 741-763.
- Puhl, R., and Brownell, K. D. (2001). Bias, discrimination, and obesity. *Obesity Research*, 9, 788-805.
- Puhl, R. M., and Brownell, K. D. (2003). Psychosocial origins of obesity stigma: Toward changing a powerful and pervasive bias. *Obesity Reviews*, 4, 213-227.

- Quine, W. V. O., and Ullian, J. S. (1978). *The web of belief* (2<sup>nd</sup> ed.). New York: McGraw-Hill.
- Rothermund, K., and Wentura, D. (2004). Underlying processes in the Implicit Association Test: Dissociating salience from associations. *Journal of Experimental Psychology: General*, *133*, 139-165.
- Rubin, Z., and Peplau, L. A. (1975). Who believes in a just world? *Journal of Social Issues*, *31*, 65-89.
- Rudman, L. A. (2004). Social justice in our minds, homes, and society: The nature, causes, and consequences of implicit bias. *Social Justice Research*, *17*, 129-142.
- Rudman, L. A., Ashmore, R. D., and Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. *Journal of Personality and Social Psychology*, *81*, 856-868.
- Russell, J. A. (2003). Core affect and the psychological construct of emotion. *Psychological Review*, *110*, 145-172.
- Sniderman, P. M., and Tetlock, P. E. (1986). Symbolic racism: Problems of motive attribution in political analysis. *Journal of Social Issues*, *42*, 129-150.
- Son Hing, L. S., Li, W., and Zanna, M. P. (2002). Inducing hypocrisy to reduce prejudicial responses among aversive racists. *Journal of Experimental Social Psychology*, *38*, 71-78.
- Stephan, W. G., and Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 23-45). Mahwah: Erlbaum.
- Strack, F., and Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220-247.
- Swim, J. K., Aikin, K. J., Hall, W. S., and Hunter, B. A. (1995). Sexism and racism: Old-fashioned and modern prejudices. *Journal of Personality and Social Psychology*, *68*, 199-214.
- Wittenbrink, B., Judd, C. M., and Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, *72*, 262-274.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*, 151-175.