CHAPTER 6

SELF-REGULATORY IDENTITY THEORY AND REACTIONS TOWARD FAIRNESS ENHANCING ORGANIZATIONAL POLICIES

**Karl Aquino, Americus Reed II,
Marcus M. Stewart, and Debra L. Shapiro**

We develop a model that explains how dual attitudes towards target members of out-groups influence fairness judgments of social policies that are designed to assist members of these groups. The model emphasizes the role of identity driven processes that either neutralize or reinforce the link between negative implicit attitudes and fairness judgments by influencing the expansiveness of a person's scope of justice. We operationalize the model using examples of different identities and conclude with theoretical and practical implications.

Organizations frequently implement policies designed to achieve social fairness amongst groups. Examples of this principle of macrojustice include affirmative action, family-friendly work policies, special mentoring for minority employees, and set-aside programs for hiring the handicapped or elderly. All of these policies are intended to improve the

circumstances of certain groups of employees, but by doing so they may negatively impact the outcomes and opportunities for other employee groups. As a result, such policies can sometimes generate feelings of injustice and resentment that can undermine their implementation. If managers and other stakeholders view the achievement of macrojustice as an important goal, then it seems important to understand how potential nonbeneficiaries of justice-enhancing policies evaluate their social fairness because such judgments can determine whether they will support or oppose such policies. In this chapter, we theorize that fairness judgments made by those who do not benefit and who also perceive themselves to be disadvantaged by fairness-enhancing organizational policies are partly influenced by their attitudes toward those groups who do benefit.

The idea that attitudes towards certain groups can influence peoples' reactions to policies that benefit other groups is not new. For example, peoples' attitudes towards African Americans and women have been found to affect fairness judgments about macrojustice-oriented policies like welfare or affirmative action (Gilens, 1999; Reid & Clayton, 1992). However, recent models of attitude formation suggest people can hold dual attitudes (Wilson, Lindsey, & Schooler, 2000). One attitude is *implicit,* resembling an enduring predisposition (cf. Greenwald & Banaji, 1995). Implicit attitudes are presumably difficult to change because they are automatic and routinized. Another type of attitude is *explicit*. Explicit attitudes are the result of conscious, reflective, and motivated cognitive processing. It displays the flexibility and responsiveness of contextually-based assessments.

According to Wilson et al.'s (2000) *dual attitude theory,* it is possible for people to hold a *negative* implicit and a *positive* explicit attitude toward the same object. The difference between them is that less cognitive effort is required to retrieve the negative implicit attitude since it is often activated at a preconscious-level. Furthermore, in many instances the negative implicit attitude will guide judgment and behavior if the positive explicit attitude is not accessed so that it may override the implicit attitude. One of the novel theoretical premises of dual attitude theory is that attitudes can become compartmentalized as opposed to integrated, such that either one may be retrieved to guide judgments and behavior. A key question posed by Wilson et al.'s (2000) theory with implications for organizational justice is: "Which of the two attitudes (the positive explicit or the negative implicit one in the above example) might dominate peoples' fairness judgments regarding a given organizational policy?"

This chapter offers one possible answer to this question and thereby extends dual attitude theory. Specifically, we propose that the self-importance of certain identities that are part of a person's working self-concept (Markus & Kunda, 1986) can influence the process of attitude-retrieval. In turn, these attitudes can influence judgments. This premise is the basis for

what we refer to as self-regulatory identification theory (SRIT). The theory incorporates the concept of "dueling identities" as a mechanism that drives the retrieval of coexisting attitudes (cf. Wilson et al., 2000). According to SRIT, different identities will facilitate the retrieval of either a negative implicit or a positive explicit attitude if these identities either expand or contract a person's *scope of justice*. The scope of justice refers to the moral rules and values governing peoples' conduct toward others and the extent to which they care about their rights and fair treatment (Opotow, 1990; Staub, 1990). We argue that an expansion of the scope of justice motivates the more effortful cognitive processing required to retrieve a positive explicit attitude. Conversely, a more restrictive scope of justice reduces cognitive effort, leading to the retrieval of the negative implicit attitude. As a result, a point we emphasize in this chapter is that the expansion or contraction of the scope of justice is one psychological mechanism that determines which of two attitudes—a negative implicit or a positive explicit—will be retrieved when people form fairness judgments.

To give a concrete example, consider a male employee who publicly proclaims that he holds egalitarian, tolerant, and favorable attitudes towards women, but who vigorously opposes organizational policies, like longer maternity leave, that can potentially benefit women more than men. An interesting theoretical explanation for the man's responses, and the one on which we focus in this chapter, is that his judgments about the fairness of policies that benefit women may reflect the oscillation of privately held dual attitudes. This will occur if, as Wilson et al. (2000) suggest, the man simultaneously holds a positive explicit and a negative implicit attitude towards women *as a group*. In this case, it may be that the negative implicit attitude primarily drives the man's fairness judgments about policies that benefit women. Although our example refers to the reaction by a man towards policies that benefit women, an assumption of SRIT is that its theoretical mechanisms can be broadly applied to *any* organizational policy designed to promote macrojustice in a way that can differentially impact members of various groups. We apply SRIT to explain fairness judgments regarding organizational policies designed to promote macrojustice in situations where these policies are likely to trigger negative stereotyping among members of nonbeneficiary groups. The following sections describe the theory and its predictions.

THEORETICAL OVERVIEW OF SRIT

The identity-regulation underpinnings of SRIT are drawn from a wide range of scholarly disciplines including personality theory (Rosenberg & Gara, 1985), self-concept and identity (Erikson, 1964), symbolic interac-

tionism (Goffman, 1959; Mead, 1934), impression management (Schlenker, 1980), social cognition (Markus, 1977), and social identity/social categorization theory (Tajfel, 1959; Turner, Hogg, Oaks, Reicher, & Wetherell, 1987). Therefore, there are several well-established assumptions of SRIT. First, SRIT assumes that peoples' working self-concept, or way of viewing and defining themselves, is comprised of multiple, hierarchically ordered identities (Brewer & Gardner, 1996; Deaux, Reid, Mizrahi, & Ethier, 1995; Stryker, 1980). Second, SRIT assumes that only a subset of these identities is activated in cognition at a given time depending on internal and external cues that make them salient (Carver & Scheier, 1998). This ordering and differential accessibility implies that the more importance a given identity holds within the person's working self-concept relative to other identities, the more likely that person's attitudes, cognitions, and behaviors are to be consistent with the attributes, traits, or standards associated with that identity. Third, SRIT assumes that the retrieval of people's coexisting implicit and explicit attitudes can be affected by which one of these identities is activated or *switched on* at the time fairness judgments are formed and expressed (cf. Forehand & Deshpande, 2001; LeBoeuf & Shafir, 2003).

Our assumption that identity-regulation processes influence *fairness judgments* is partly based on Skitka's (2003) accessibility identity model (AIM). Two theoretical assertions of AIM are relevant for explaining the relationship between identity activation and fairness judgments. The first is that people should be more concerned about justice in contexts that activate identity-relevant concerns (the *identity-relevance hypothesis*). The second is that people should devote more thought and analysis to whether a given situation is fair or unfair if aspects of that situation threaten rather than affirm the perceiver's currently activated identity and associated goals and values (the *identity-threat hypothesis*). These features of AIM provide a theoretical rationale for why different identities can influence the recruitment of either a negative implicit or a positive explicit attitude that, in turn, can influence fairness judgments.

SRIT departs from previous identity-related theorizing by identifying *when* people are likely to activate an identity that is more likely to retrieve one versus another of the dual attitudes they hold toward groups (whose membership characteristics exclude them) who are likely to benefit from a macrojustice policy; that is, an identity that is likely to be more rather than less supportive towards a macrojustice policy. More specifically, to explain fairness judgments, SRIT focuses on identity driven processes that are likely to lead to either a contraction or expansion of the psychological group boundaries that define the self's relationship to others and, by extension, the scope of justice. In turn, this process is presumed to drive the recruitment of compartmentalized dual attitudes toward the

group that is perceived to benefit from a fairness-enhancing organizational policy. SRIT is meant to explain fairness judgments of those who perceive themselves as being disadvantaged by the policy. For simplicity, we use the term "potentially disadvantaged group" (PDG) to refer to members of this nonbeneficiary group, and "potentially advantaged group" (PAG) to refer to members of groups that are perceived as benefiting from the macrojustice policy. Figure 6.1 depicts the psychological mechanisms of SRIT.

SRIT assumes that the different identities that a person uses to define the self can act as opposing motivational forces that facilitate the retrieval of either a positive explicit or negative implicit attitude towards PAGs. The model is restricted to this particular case of dual attitudes because if explicit and implicit attitudes are consistent (i.e., both are positive, or both are negative), there is only one valenced attitude available for recruitment. As it applies to fairness judgments, Figure 6.1 shows that at a preconscious cognitive level, a negative implicit attitude toward the PAG can be triggered by the activation of negative stereotypes held by a PDG (Fazio, Jackson, Dunton, & Williams, 1995). We assume that these stereotypes are initially triggered by a perception that the policy threatens the PDG members' attainment of valued goals. This assumption is consistent with a large body of research showing that people more systematically scrutinize the fairness of situations that threaten self-relevant goals (e.g., Higgins, 1987; Steele, 1999). For our purposes, we do not specify precisely what *all* possible triggering conditions might be, only that these conditions should provoke negative stereotyping among nonbeneficiaries (PDG members). An example would be an organizational policy that gives preferential treatment to certain ethnic groups and is perceived as decreasing the promotion chances of other ethnic groups. Another example would be a family-friendly work policy perceived as placing a greater productivity burden on single employees than those with families. In both cases, we propose a context is created in which SRIT can be applied to explain the formation of fairness judgments made by PDG group members who hold dual attitudes towards the PAG.

We expect negative implicit attitudes to *negatively* influence fairness judgments about the policy that benefits the stereotyped PAG (Path A). That is, the more negatively the person evaluates the PAG at a preconscious, automatic level the less fair he or she will perceive the policy to be. At another level that reflects higher order executive control (cf. Wilson et al., 2000), our model suggests that for those employees who also hold favorable explicit attitudes toward the PAG, such attitudes can *positively* influence fairness judgments (Path B). This path implies that a positive explicit attitude should lead the person to evaluate the policy as more fair. The model shows that the source of these explicit attitudes may be any
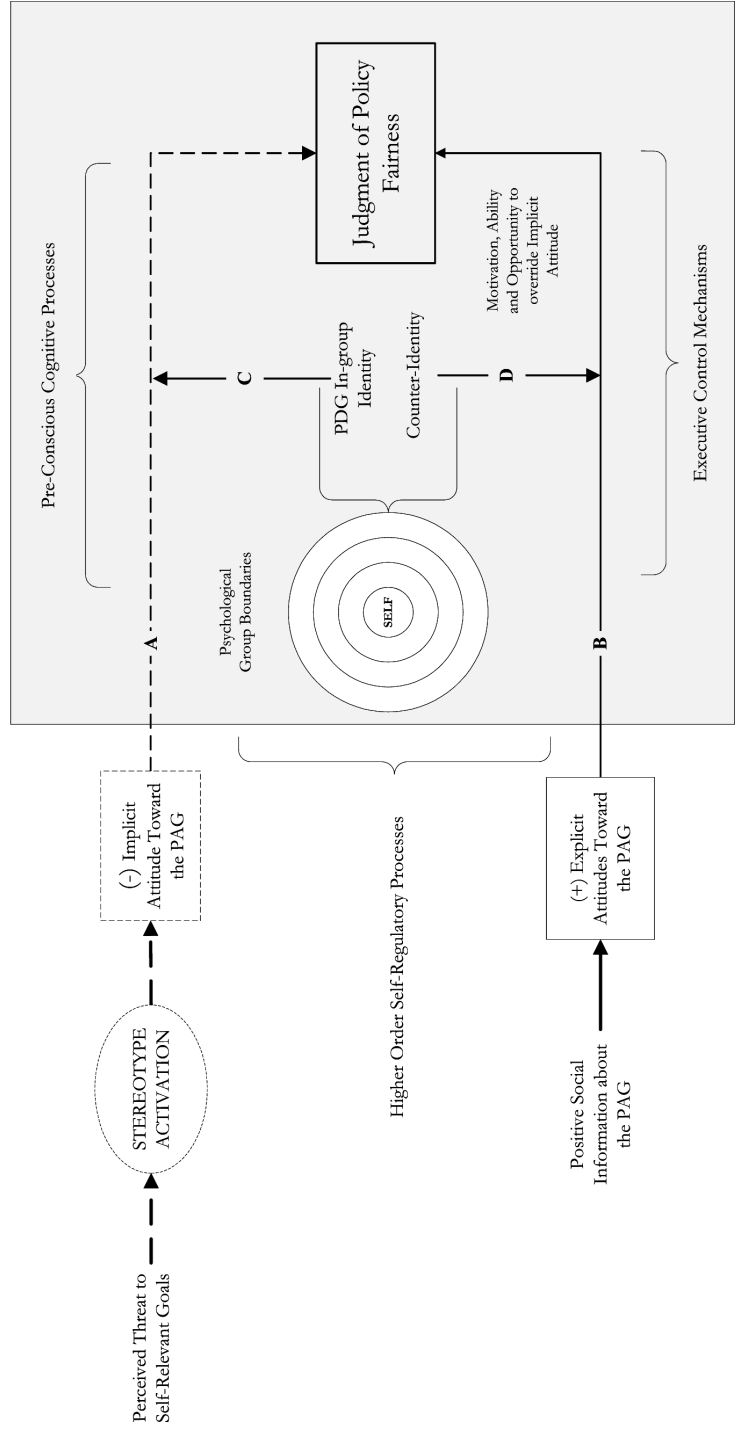
Figure 6.1   Self-regulatory identification theory (SRIT): implications for fairness judgments.

positively valenced social information that the person uses to construct favorable PAG attitudes. According to dual attitude theory (Wilson et al., 2000), it is possible for either the negative implicit or the positive explicit attitude to exert the *dominant* impact on fairness judgments, depending on which one is retrieved.

SRIT makes a specific prediction about how a key attribute of the various identities that people use to define the self influences the *retrieval* of dual attitudes; indeed, this prediction is what SRIT adds to AIM and prior theories of the self. The attribute identified as key by SRIT is whether an identity motivates an expansion or contraction of the psychological group boundaries that define the self's relationship to others. Moreover, SRIT defines two general concepts: in-group identity and counter-identity as two higher order self-regulatory mechanisms that will influence how expansive versus restrictive one's scope of justice will be. Next, we identify the circumstances in which people's identity is likely to motivate their psychological group boundary to expand versus contract.

Factors Likely to Influence Psychological Group Boundaries

Social psychological research and common experience shows that people routinely establish mental boundaries to distinguish those persons to whom they are expected to show moral concern from those to whom they are not. These boundaries delineate one's scope of justice (Opotow, 1996). For people included in the scope of justice, rights, and fair treatment become paramount concerns. Concerns about rights and fair treatment can seem irrelevant, however, for those who lie outside the scope of justice. Instead, these persons are more likely to be seen as nonentities, undeserving, or expendable (Opotow, 1990). For most people, the scope of justice almost always encompasses family and kin. But people also extend their scope of justice outward to include those who share the same neighborhood, organization, cultural practices, physical characteristics, nationality, ethnic heritage, or other salient, valued characteristics or distinguishing features. What is important for our purposes is the fact that some of the justice-defining boundaries are more exclusionary than others. For example, some people rigidly limit their scope of justice only to family, whereas others incorporate people from the same community. At the extreme, the scope of justice can include all of humanity and even other species.

We argue in this chapter that the expansion or contraction of the scope of justice is one psychological mechanism that determines which of two attitudes—a negative implicit or a positive explicit—will be retrieved when people form fairness judgments. Moreover, it is the activation of

specific kinds of identities that partly determines the expansiveness of the scope of justice. There are many identification processes that might conceivably influence the likelihood of attitude retrieval through contraction or expansion of a person's scope of justice. SRIT defines two general concepts as two higher order self-regulatory mechanisms that influence this process: in-group identity, which in our model refers only to PDG members, and counter-identity.

Identity of PDG Members

SRIT proposes that when people have stronger (rather than weaker) identities as members of a PDG group, their scope of justice will be more restrictive. For example, research shows that in-group identification leads people to selectively search for intergroup differences that favor the in-group and to dismiss information that favors the out-group (e.g., Brewer & Kramer, 1985; Devine, 1989; Mikulincer & Shaver, 2001). This can lead in-group members to view out-group members as inferior, to formulate self-serving causal explanations for positive in-group outcomes (e.g., Hamilton & Trolier, 1986), and to use linguistic labels to differentiate in-groups and out-groups (e.g., Dovidio & Gaertner, 1993). SRIT suggests that when in-group identity occupies a high level of self-importance within the working self-concept relative to other identities, it can reinforce the link between the negative implicit attitude and negative fairness judgments about a policy that benefits an out-group (i.e., a group that is outside of one's scope of justice), while disadvantaging the in-group (i.e., a group that is inside the scope of justice). The accessibility of in-group identity, coupled with the threat posed by the organizational policy to the welfare of the in-group, will make it more likely that an in-group member's fairness judgments will be driven primarily by their concern for their own group's outcomes.

The above reasoning follows Brickson's (2000) argument that when a collective or social identity is salient, people are motivated to enhance the welfare of their own group relative to other groups. As a result, we expect them to engage in less effortful cognitive processing when forming fairness judgments. In turn, this makes the retrieval of positive attitudes toward an out-group that potentially benefits from a fairness enhancing policy less likely. According to dual attitude theory (Wilson et al., 2000), fairness judgments will be influenced primarily by preconscious and negative implicit attitudes. The process just described is illustrated by Path C in Figure 6.1.

However, an identity that is based on one's membership in a particular social group is not the only kind of identity that can be used to organize the working self-concept. According to SRIT, if another type of identity,

referred to here as a counter-identity, is activated, it can increase the motivation to retrieve the positive explicit attitude.

Counter-Identity

By counter-identity, we refer to any self-representation available in the working self-concept that is likely to expand the scope of justice. The idea of a counter-identity as used in SRIT refers specifically to those identities that are rooted in people's internalized notions of *ought* and *should* (e.g., Bandura, 1986; Higgins, 1987; Steele, 1999), and personal values that are terminal goals on their own (e.g., individualized commitment to values like equality, freedom, or the sanctity of life). These are identities that might make people sensitive to broader issues of social justice and to social injustice (Skitka, 2003). As Wu (2002) has argued, peoples' behavior toward others is not driven primarily by color blindness or meritocracy, but by community and membership. According to Wu (2002), membership confers merit. Consequently, the more a given identity promotes the inclusion of others within a person's scope of justice, the more accepting a person should be of policies that benefit those whom they have judged worthy of moral concern.

We theorize that adopting a counter-identity that expands the scope of justice can motivate explicit attitude retrieval by facilitating *de-categorization*, whereby bias is reduced by moving (former) in-group members away from the self and towards out-group members (Brewer, 1999; Hewstone, Rubin, & Willis, 2002). Similarly, *recategorization* tends to alter which social categorizations are used and to replace subordinate (us and them) with super-ordinate (we) categorizations (Gaertner & Dovidio, 2000). Research suggests that bias is reduced by improving attitudes toward former out-group members, owing to their recategorization from out-group to in-group (Hewstone et al., 2002). This process is depicted by the concentric circles in the center of Figure 6.1. Thus, if a highly salient or self-important counter-identity weakens the psychological boundaries that separate *us* from *them*, it may also motivate a person to engage in more complex and effortful cognitive processing about an out-group member because that member is more likely to be included within the perceiver's scope of justice. Although it is possible that exposure to an out-group member may activate preconscious negative stereotypes which lead to negative implicit attitudes, SRIT theorizes that the activation of a specific counter identity—if sufficiently strong—can neutralize the association between the negative implicit attitude and negative social judgments by motivating the retrieval of existing positive explicit attitudes toward that out-group. This effect is illustrated by Path D in Figure 6.1.

Dueling Identities and Attitude Retrieval

SRIT refers to in-group identity and counter-identity as *dueling* because they have potentially competing effects on the cognitive boundaries that define the self's relation to others and, by extension, the scope of justice. The question of which identity may be most salient to a person in a given situation is relevant for understanding how that person might interpret and evaluate social information. Based on this argument, we consider possible consequences of a highly self-important in-group identity and the implications of having a highly self-important counter-identity. We do so by using specific identities as examples and discussing their substantive implications.

TRANSLATING SRIT INTO TESTABLE HYPOTHESES

Operationalizing In-Group and Counter-Identities

There are many social identities a person may potentially adopt as a PDG identity (cf. Deaux et al., 1995). In our model, the assumption is that the relevant identity will be one that distinguishes those who benefit from a fairness enhancing organizational policy (PAG) from those who do not (PDG). Given this limiting assumption, the model can be applied to many different social identities such as those based on gender, race/ethnicity, nationality, physical condition, and so forth. Applying SRIT to the entire universe of social identities is beyond the scope of this chapter. Instead, we seek to illustrate the concrete application of the model by specifying a relevant in-group and counter-identity in a situation that is likely to activate negative stereotypes. We do not claim that these are the only PDG identities that matter; there are obviously other identities that could be substituted. Describing these is an appropriate undertaking for future theory and research. In this chapter, the PDG social identity we emphasize for illustrative purposes is *White racial identity*. We offer the example of *moral identity* as a counter-identity that can expand the scope of justice, thereby facilitating positive explicit attitude retrieval.

We chose these particular identities for two reasons. First, studies have documented important gains in racial equality as a direct result of policy driven initiatives that can pose real or perceived threats to the interests of historically dominant racial groups (Bowen & Bok, 1998). For example, a U.S. Labor Department report detailed that affirmative action, which some people regard as a fairness enhancing organizational policy, has helped five million minority members and six million White and minority women move up in the workforce ("Reverse Discrimination," 1995). Since one assumption of our theory is that conditions that pose threats to goal

attainment by nonbeneficiaries of such policies can activate negative stereotyping, examining how Whites react to policies that benefit racial minorities should illustrate SRIT's explanatory usefulness. Second, with regard to moral identity, it has been suggested (e.g., Younis & Yates, 1999), and shown empirically (Reed & Aquino, 2003), that this identity influences the expansiveness of one's scope of justice. Consequently, there is reason to believe that this identity may affect how motivated a person may be to override negative attitudes toward an out-group member. We present moral identity as *one possible exemplar* of a counter-identity in the context of our theoretical model, recognizing that in reality there could be other identities that serve the same function.

White Identity

Racial identity is an individual's general perception of self with respect to race. Some research has suggested a link between White racial identity and biased judgments related to out-groups. For example, Levin, Sidanius, Rabinowitz, and Federico (1998) reported racial identity among a sample of White undergraduates was positively associated with making *internal attributions* for poverty among African Americans and the belief that the Los Angeles (California) riots of 1992 were caused by criminal elements rather than a protest against injustice. A recent study by Negy, Shreve, Jensen, and Uddin (2003) also found a positive association between White racial identity and ethnocentrism among a sample of undergraduate and graduate students. Based on these and similar empirical studies (e.g., Crocker, Luhtanen, Blaine, & Broadnax, 1994; Larkey & Hecht, 1995), we expect White racial identity to be negatively associated with attitudes towards beneficiaries of social policies designed to redistribute valuable resources to racial out-group members.

Undoubtedly, White Americans differ in the degree to which their racial identity is a central or important aspect of their overall self-concept. This means their social identities as Whites are not equally accessible or cognitively salient relative to other identities that they may use to define their working self-concept. However, in consideration of the discussion above, we suggest that Whites who do identify themselves strongly in terms of racial group membership are more likely to show preferential moral concern for the welfare of members of their White in-group than for members of other (nonWhite) out-groups, as per intergroup and categorization theories (Turner et al., 1987). In other words, their scope of justice is more likely to be restricted to racial in-group members under certain conditions.

According to both SRIT and AIM, the greater accessibility of their in-group social identity should increase their sensitivity to *group-based* norms of fairness. Thus, when an organizational policy is implemented to

achieve macrojustice, White employees will be more likely to evaluate the policy in terms of how it affects the welfare of the different groups to which they may or may not closely identify. Assuming that the policy is seen as potentially disadvantaging Whites as a racial group, while benefiting certain minority groups (e.g., African Americans, Hispanics), we expect someone whose White racial identity is highly salient *and* who also holds dual attitudes towards minority groups who benefit from the policy, to rely primarily on their negative implicit attitude to form fairness judgments. This also means they will be less motivated to engage in the effortful cognitive processing required to retrieve the positive explicit attitude because they will be focused primarily on the adverse impact that the policy has on their racial in-group. Our argument is consistent with past research suggesting that people have different perceptions of fairness depending on whether they take a more individual versus social identity perspective (e.g., Davidson & Friedman, 1998; Wenzel, 2001). What we add to these findings is one explanation for these differences; namely, that identity accessibility facilitates the retrieval of different attitudes.

Having described how a highly salient in-group identity can reinforce the association between a negative implicit attitude and fairness judgments, we now turn to an identity whose effect is hypothesized to have the opposite effect.

Moral Identity

Moral identity refers to the concept of a person's character held internally and projected to others. According to Lapsley and Lasky (2001, p. 347), a person who has a moral identity is "one for whom moral schemas are chronically available, readily primed, and easily activated for information processing." Extending this definition, Hart, Atkins, and Ford (1998, p. 515) defined moral identity as "a commitment to one's sense of self to lines of action that promote or protect the welfare of others." The self-other relationship implied by Hart et al.'s definition is fundamental to a person's moral identity (Reed & Aquino, 2003; Younis & Yates, 1999). Some writers have theorized that a person whose moral identity has high self-importance should show greater concern for a larger segment of humanity than someone whose moral identity is less important (Hart et al., 1998; Younis & Yates, 1999). Reed and Aquino (2003) provided one of the only empirical tests of this argument in four studies that examined peoples' willingness to demonstrate concern for the welfare of out-groups as a function of the self-importance of their moral identities. Reed and Aquino (2003) theorized that people whose moral identity has high self-importance would show decreased negativity towards out-groups and would be more willing to behave in ways that benefit members of these groups. They reasoned that one consequence of having a self-important

moral identity is that it expands people's circle of moral concern. Reed and Aquino (2003) showed empirically that this circle of moral concern appears to be larger (i.e., moves outward toward encompassing all of humanity) as the self-importance, or salience, of moral identity increases. Their findings therefore support the notion that the self-importance of moral identity can influence the scope of justice.

Based on past theorizing about how moral identity defines the self-other relationship, and empirical research suggesting that moral identity can motivate an expansion of the scope of justice, we present this identity as an exemplar of a counter-identity to White identity because it may facilitate the effortful retrieval of a positive explicit attitude when dual attitudes exist. Returning to our previous example, assuming that an organizational policy is perceived as disadvantaging Whites (the PDG) and benefiting minorities (the PAG), and that a particular White employee holds dual attitudes toward minorities, SRIT suggests that if this PDG member's moral identity is highly accessible relative to other identities, it can motivate them to retrieve their positive explicit attitude when forming fairness judgments about that policy. In other words, these attitudes will override their negative implicit ones. The above analysis implies that moral identity can potentially compete with White racial identity because it expands the scope of justice. This effect is illustrated by Path D in the SRIT model.

THEORETICAL AND PRACTICAL IMPLICATIONS

Having provided concrete examples of the identity driven processes in our model, we now turn to some of its theoretical and practical implications. An implication of SRIT is that if an organizational policy designed to achieve macrojustice is to be widely accepted, judgments of its social fairness should not be predominantly driven by negative implicit attitudes towards those who are perceived to benefit from the policy. SRIT suggests that *observable and openly expressed* attitudes may not fully explain fairness judgments. Instead, some individuals may hold different evaluative reactions toward certain groups that they do not publicly reveal. They may harbor no *observable* animosity towards out-groups but they may still be resentful of and actively oppose policies that benefit these out-groups if these policies pose a threat to their goals. Such reactions can potentially undermine initiatives designed to promote diversity, cooperation, and mutual respect among members of diverse groups. For this reason, the theoretical mechanisms outlined by SRIT have important implications for practitioners and scholars interested in promoting macrojustice in organizations. Our discussion of these implications is organized around the four

key mechanisms of SRIT (Paths A, B, C and D), with emphasis on the higher-order regulatory mechanisms (Paths C and D).

Minimizing the Availability of Negative Implicit Attitudes toward Out-Groups (Path A)

No matter how credible and persuasive are interventions meant to create favorable reactions toward out-group members, the SRIT model suggests that any management intervention whose aim is to create support for a fairness enhancing policy will be inadequate if such interventions do not consider how negative implicit attitudes affect fairness judgments. Although such attitudes are sometimes undone by heightened exposure to members of the stereotyped out-group, SRIT implies that even this exposure falls short unless societal members are made aware of their implicit biases and are somehow motivated to try and override them. For example, Gaertner and Dovidio's (1986) research on *aversive racism* shows that people who consciously endorse egalitarian values can also hold negative feelings about minority groups that are excluded from awareness. On the surface, the implicit attitudes that result from negative stereotype activation processes (Wheeler, Jarvis, & Petty, 2001) can likely be changed or even modified by calling them to a person's conscious awareness. For example, one approach that has been found to be somewhat effective at mitigating bias is encouraging people to bring their prejudices to mind and to then invoke self-directed guilt derived from the discrepancy between personal values and actual behavior (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002).

Maximizing the Availability of Positive Explicit Attitudes toward Out-Groups (Path B)

A variety of interventions have been proposed for generating positive explicit attitudes towards an object, person, or idea. Among these are educating others about why social diversity is desirable (Crosby & Clayton, 2001), publicizing the unique contributions that out-group members make to the larger social collectives such as organizations (Pratkanis & Turner, 1996), or highlighting the accomplishments of out-group members (e.g., the achievements of work groups that are composed of these out-groups; Konrad & Linnehan, 1995). However, a key insight from SRIT is that only focusing on the formation of positive attitudes is not sufficient because some people may simultaneously hold negative implicit attitudes. The interventions cited above seem ill equipped to *replace* nega-

tive implicit attitudes, particularly if dual evaluations are never brought into conflict. Simple exposure to out-group members may raise awareness that negative stereotypes are inaccurate and this realization could *possibly* weaken the automatic negative associations that are the basis for negative implicit attitudes. But if this does not occur, then some persons may continue to harbor resentment and hold negative evaluations of out-group targets. Indeed, for some people these reactions may be particularly strong as a result of having highly salient in-group identities.

The Salience of PDG In-Group Identity (Path C)

The immediate social context in which individuals go about their day-to-day activities can increase the accessibility of certain social identities. For example, imagine an African American executive in a meeting room full of either male Caucasians or of female African Americans. Due to differential distinctiveness, it is likely that this person's ethnic social identity would be highly salient in the first situation, and that his gender social identity would be highly salient in the second (McGuire, McGuire, & Winton, 1979). SRIT posits that in-group identity is a key determinant of attitude retrieval. Since the maintenance of positive self-esteem is among the strongest and most persistent human goals (Hales, 1985), future research should focus on understanding how in-group identities are formed and maintained in a way that reinforces the retrieval of negative implicit attitudes. Building on the notion that increasing the salience of a PDG identity can restrict a person's scope of justice, one implication of SRIT is that it is important to minimize the use of language or situations that make in-group identities highly salient. It may be more prudent instead to focus on heightening the salience of a counter-identity.

The Salience of a Counter Identity (Path D)

SRIT suggests that organizations may instead try to make alternative, counter-identities such as moral identity more salient to encourage the retrieval of positive explicit attitudes. A key intervention that can influence the accessibility of moral identity is the presence of contextual cues in the environment. For example, fairness enhancing policies could be worded to emphasize larger self-identity representations. The important factors are the skills and experiences the individual is bringing to the situation that will benefit his or her broader constituents, and organizations' efforts to select members utilizing fair and socially responsible proce-

dures. Exposure to these kinds of contextual cues may be an important way of activating Path D in the SRIT model.

With respect to moral identity, environmental cues can activate moral identity related concepts in memory and thereby increase the likelihood that organizational members will use their moral identity in the evaluation of relevant stimuli (cf. Forehand & Deshpandé, 2001; Forehand, Deshpandé, & Reed II, 2002; Hong, Morris, Chiu, & Benet-Martinez, 2000; Wyer & Srull, 1986). Perhaps by invoking concepts that focus on terminal values like equality, solidarity, or compassion, managers may be able to prime moral identities, making it more likely that the fairness judgments of nonbeneficiaries of these policies who also hold positive explicit attitudes will be driven primarily by these attitudes and not implicit negative ones. These values and traits are also often those that organizations adopt and assert as their own in mission statements, statements of values, and so forth. Emphasis on the congruence of these values with those of the organizations', and the role that social policies such as affirmative action or family-friendly policies play in manifesting those values, may help to access or reinforce employees' moral identities, or other relevant counter-identities associated with the desired values such as individuals' professional and/or organizational identities.

Dueling Counter and In-Group Identities and the Congruency of the Self

One way to influence the activation of a specific counter identity that can neutralize the impact of a negative implicit attitude is to make the individual aware of identity driven conflict between his/her desired and actual self. On seeing the conflict between explicit and implicit attitudes, the individual may experience *tension* from this *self-discrepancy*, which might elicit a higher order cognitive process. With this higher-order executive control mechanism active, the individual can take positive steps towards neutralizing the influence of a negative implicit attitude. Future research might examine the possibility that when a person strives to maintain consistency between in-group identity, moral identity, and his/her actions in the world, they may be motivated to internalize and or recruit a more positive attitude towards out-groups. However, as a caveat, future research should also examine the extent to which making people aware of their implicit biases will cause them to fortify in-group identity driven boundaries as a defense mechanism (a boomerang effect). An important question that our model does not fully address is how strong do in-group and counter identities have to be to experience conflict between them? If someone can be pressured to change their attitudes with the tides and

currents of the social groups with which they engage, how does this reinforce their other social identities?

CONCLUSIONS

A key matter in organizational settings is the need to effectively manage diversity in the workplace while also promoting some measure of macrojustice. At times, the psychological experiences driven by perceptions of macrojustice policies can adversely affect work performance and motivation. To create and manage macrojustice instruments of social fairness, it is therefore prudent and necessary to closely examine the psychological mechanisms that may underlie how people evaluate such instruments and their beneficiaries. Toward this end, there are several directions for future study based on applying SRIT to situations in which fairness judgments about macrojustice policies are likely to occur. Briefly, we suggest that researchers should: (1) examine the different kinds of identities that are likely to expand the scope of justice, (2) determine the extent to which these identities can be developed and reinforced within organizations, (3) explore how unfavorable identities might be suppressed or their influence diminished, (4) investigate what happens when two competing identities are equally powerful and the psychological mechanisms that determine which identity *prevails*.

We hope that the model developed in this chapter can help spur research to answer questions about how organizational leaders can most effectively implement justice enhancing policies even when, at first glance, such policies may seem to be making *beneficiaries* of some but not others.

REFERENCES

Au: Aquino entry not found in text.

Aquino, K. F., & Reed, A., II. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology, 83,* 1423-1440.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Bowen, W. G., & Bok, D. (1998). *The shape of the river: Long-term consequences of considering race in college and university admissions.* Princeton, NJ: Princeton University Press.

Brewer, M. B. (1999). The psychology of prejudice: In-group love or out-group hate? *Journal of Social Issues, 55,* 429-444.

Brewer, M. B., & Gardner, W. (1996). Who is this "we?" Levels of collective identity and self representations. *Journal of Personality and Social Psychology, 71,* 83-93.

Brewer, M. B., & Kramer, R. M. (1986). Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing. *Journal of Personality and Social Psychology, 50,* 543-549.

Brickson, S. (2000). The impact of identity orientation on individual and organizational outcomes in demographically diverse settings. *Academy of Management Review, 25,* 82-101.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior.* New York: Cambridge University Press.

Crocker, J., Luhtanen, R., Blaine, B., & Broadnax, S. (1994). Collective self-esteem and psychological well-being among White, Black, and Asian college students. *Personality and Social Psychology Bulletin, 20,* 503-513.

Crosby, F. J., & Clayton, S. (2001). Affirmative action: Psychological contributions to policy. *Annals of Social Issues and Public Policy, 1,* 71-87.

Davidson, M., & Friedman, R. (1998). When excuses don't work: The persistent injustice effect among Black managers. *Administrative Science Quarterly, 43,* 154-183.

Deaux, K., Reid, A., Mizrahi, K., & Ethier, K. A. (1995). Parameters of social identity. *Journal of Personality and Social Psychology, 68,* 280-291.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5-18.

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

Dovidio, J., & Gaertner, S. (1993). Stereotypes and evaluative inter-group bias. In D. Mackie & D. Hamilton (Eds.), *Affect, cognition, and stereotyping: Interactive processes in intergroup perception* (pp. 167-193). Orlando, FL: Academic Press.

Erikson, E. (1964). *Insight and responsiblity: Lectures on the ethical implications of psychoanalytic insight.* New York: W. W. Norton.

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

Forehand, M. R., & Deshpandé, R. (2001). What we see makes us who we are: Priming ethnic self-awareness and advertising response. *Journal of Marketing Research, 38,* 336-348.

Forehand, M. R., Deshpandé, R., & Reed, A., II. (2002). Identity salience and the influence of differential activation of the social self-schema on advertising response. *Journal of Applied Psychology, 87,* 1086-1099.

Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61-89). Orlando, FL: Academic Press.

Gaertner, S. L., & Dovidio, J. F. (2000). *Reducing intergroup bias: The common ingroup identity.* New York: Psychology Press.

Gilens, M. (1999). *Why Americans hate welfare.* Chicago: University of Chicago Press.

Goffman, E. (1959). *The presentation of self in everyday life.* Harmondsworth, United Kingdom: Penguin Books.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4-27.

Hales, S. (1985). The inadvertent rediscovery of self in social psychology. *Journal of the Theory of Social Behavior, 15,* 237-282.

Hamilton, D. L., & Trolier, T. K. (1986). Variables influencing judgments of correlational relations. *Journal of Personality and Social Psychology, 50,* 879-888.

Hart, D., Atkins, R., & Ford, D. (1998). Urban America as a context for the development of moral identity in adolescence. *Journal of Social Issues, 54,* 513-530.

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53,* 575-604.

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

Hong, Y., Morris, M. W., Chiu, C., & Benet-Martínez, V. (2000). Multicultural minds: A dynamic constructivist approach to culture and cognition. *American Psychologist, 55,* 709-720.

Konrad, A. M., & Linnehan, F. (1995). Race and sex differences in line managers' reactions to equal employment opportunity and affirmative action interventions. *Group and Organization Management, 20,* 400-439.

Lapsley, D. K., & Lasky, B. (2001). Prototypic moral character. *Identity, 1,* 45-363.

Larkey, L. K., & Hecht, M. L. (1995). A comparative study of African American and European American ethnic identity. *International Journal of Intercultural Relations, 19,* 483-504.

LeBoeuf, R. A., & Shafir, E. (2003). *Alternating selves and conflicting choices: Identity salience and preference inconsistency.* Manuscript submitted for publication.

Levin, S., Sidanius, J., Rabinowitz, J. L., & Federico, C. (1998). Ethnic identity, legitimizing ideologies, and social status: A matter of ideological asymmetry. *Political Psychology, 19,* 373-404.

Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology, 35,* 63-78.

Markus, H., & Kunda, Z. (1986). Stability and malleability of the self-concept. *Journal of Personality and Social Psychology, 51,* 858-868.

McGuire W. J., McGuire, C. V., & Winton, W. (1979). Effects of household sex composition on the salience of one's gender in the spontaneous self-concept. *Journal of Experimental Social Psychology, 15,* 77-90.

Mead, G. H. (1934). *Mind, self and society from the standpoint of a social behaviorist.* Chicago: University of Chicago Press.

Mikulincer, M., & Shaver, P. R. (2001). Attachment theory and intergroup bias: Evidence that priming the secure base scheme attenuates negative reactions to out-groups. *Journal of Personality and Social Psychology, 81,* 97-115.

Negy, C., Shreve, T. L., Jensen, B. J., & Uddin, N. (2003). Ethnic identity, self-esteem, and ethnocentrism: A study of social identity versus multicultural theory of development. *Cultural Diversity and Ethnic Minority Psychology, 9,* 333-344.

Opotow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues, 4,* 1-20.

Opotow, S. (1996). Affirmative action, fairness, and the scope of justice. The affirmative action debate: What's fair in policy and programs? *Journal of Social Issues, 52,* 19-24.

Pratkanis, A. R., & Turner, M. E. (1996). The proactive removal of discriminatory barriers: Affirmative action as effective help. *Journal of Social Issues, 52,* 111-132.

Reed, A., II, & Aquino, K. (2003). Moral identity and the circle of moral regard towards out-groups. *Journal of Personality and Social Psychology, 84,* 1270-1286.

Reid, P. T. & Clayton, S. (1992). Racism and sexism at work. *Social Justice Research, 5,* 249-268.

Reverse discrimination of whites is rare, labor study reports. (1995, March 31). *New York Times*, p. A23.

Rosenberg, S., & Gara, M. A. (1985). The multiplicity of personal identity. In P. Shaver (Ed.), *Self, situations, and social behavior: Review of personality and social psychology* (Vol. 6, pp. 87-113). Beverly Hills, CA: Sage.

Schlenker, B. R. (1980). *Impression management.* Belmont, CA: Brooks/Cole.

Skitka, L. J. (2003). Of different minds: An accessible identity model of justice reasoning, *Personality and Social Psychology Review, 7,* 286-297.

Staub, E. (1990). Moral exclusion, personal goal theory, and extreme destructiveness. *Journal of Social Issues, 46,* 47-64.

Steele, C. M. (1999, August). Thin ice: "Stereotype threat" and black college students. *The Atlantic Monthly,* 44-54.

Stryker, S. (1980). *Symbolic interactionism: A social structural version.* Menlo Park, CA: Benjamin/Cummings.

Tajfel, H. (1959). Quantitative judgement in social perception. *British Journal of Psychology, 50,* 16-59.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory.* Oxford, England: Blackwell.

Wenzel, M. (2001). A social categorization approach to distributive justice: Social identity as the link between relevance of inputs and need for justice. *The British Journal of Social Psychology, 40,* 315-336.

Wheeler, S. C., Jarvis, W. B. G., & Petty, R. E. (2001). Think unto others: The self-destructive impact of negative racial stereotypes. *Journal of Experimental Social Psychology, 37,* 173-180.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107,* 101-126.

Wu, S. The Psychology of Color. Retrieved September 17, 2002, from http://psychologyabout.com/library/weekly/aa031501a.htm

Wyer, R. S., & Srull, T. K. (1986). Human cognition in its social context. *Psychological Review, 93,* 322-359.

Younis, J., & Yates, M. (1999). Youth service and moral-civic identity: A case of everyday morality. *Educational Psychology Review, 11,* 361-376.

Au: Volume number?

Au: Online publication year?