13

THE DYNAMIC NATURE OF IDENTITY

From the Brain to Behavior

Dominic Packer and Jay J. Van Bavel

"Do I contradict myself?
Very well then I contradict myself.
(I am large, I contain multitudes.)"

—Walt Whitman, *Songs of Myself*, 1855

In the 1960s, Henri Tajfel and his colleagues were investigating how intergroup bias emerges from aspects of the intergroup context such as competition for resources, status differentials, and negative stereotypes. Recognizing that many of these variables tend to co-occur in real intergroup situations, they created a minimal version of an intergroup context involving two groups without any interaction with ingroup or outgroup members or any of the other aspects typically associated with intergroup conflict. They assigned participants to these so-called "minimal" groups on the basis of arbitrary and rather meaningless criteria, such as their ability to estimate the number of dots on a screen or preferences for abstract artwork,. Participants were then asked to allocate resources between the members of their ingroup and the outgroup. Importantly, the researchers ensured that there was no competition for resources and that participants' decisions had no direct bearing on their own individual outcomes (see Billig, 1976 for a review).

Rather than eliminating intergroup bias—as one might have expected—participants randomly assigned to arbitrary social groups discriminated in favor of their ingroup (Billig & Tajfel, 1970, 1973; Tajfel & Turner, 1979). Specifically, they allocated more money to members of their ingroup than to members of the outgroup (Brewer, 1979). It appeared that social categorization itself—simply belonging to one group rather than another—was sufficient to induce intergroup bias. Tajfel and his colleagues had identified one of the most influential baseline conditions in the history of psychological science.

These minimal group studies helped introduce two seminal concepts in psychology: namely, that people rapidly and flexibly categorize themselves in terms of their group memberships (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987) and that social identities are a central motivational force in human social life (Tajfel, 1982). In other words, when people are part of a group, they can adopt that category as a *social identity*, such that the group and fellow group members are treated as part of the self. Having taken on the group as part of their social identity, people are motivated to enhance the success, reputation, and vitality of the group in the same way that they typically seek to enhance their individual selves.

Overview

In this chapter, we review major theories and empirical findings that highlight the dynamic nature of self-representation with a focus on research that shows that when people perceive themselves as part of a group or coalition, the self-concept shifts from an individual to a collective level (i.e., I is redefined as We; Brewer & Gardner, 1996; Turner et al., 1987). There are several good reasons group-level concerns shape human cognition and behavior. By building and maintaining coalitions with others, people can fulfill a variety of motives more successfully than they can on their own (Brewer, 2004; Caporael, 1997; Wilson & Sober, 1994). In addition to fulfilling basic human psychological needs, such as belonging and distinctiveness (Baumeister & Leary, 1995; Brewer, 1991), working in groups allows people to obtain resources that are only accessible through coordination or division of labor (Allport, 1954; Correll & Park, 2005).

Drawing on insights from large literatures on group behavior and social identity, we argue that self-representation is highly sensitive to social cues, such that people dynamically construct and adopt social identities within group contexts (e.g., when groups represent opportunities for agency; Brewer, 2008). One important functional consequence of social identification is that it shifts motivational concerns from the individual to the collective level. For instance, highly identified group members adopt the goals of the groups with which they identify— pursuing collective interests as actively as they might ordinarily pursue individual self-interest. The types of goals (e.g., for status, stability) often remain constant, but the "agent" or entity for which they are pursued shifts from the individual self to the group. This helps explain why people not only seek to increase their own status by striving for individual success but, in other situations, try to increase their ingroup's status by enhancing collective achievement (Ellemers, Wilke, & Van Knippenberg, 1993).

The dynamic self-concept is both powerful and parsimonious because it suggests the same set of cognitive processes that are used to make decisions with regard to individual-level interests are also used to make decisions with regard to collective, group-level interests. As a result, we posit that shifts toward collective self-representation can account for a great deal of prosocial and collectively

oriented cognition and behavior, from anonymous donations to altruistic self-sacrifice, that otherwise appears puzzling from the perspective of standard theories of personality and decision making.

In the pages that follow, we briefly review classic perspectives on the stability of identity and social preferences before outlining an alternative approach premised on the dynamic nature of identity. We then review research illustrating how cognition shifts from individual to collective concerns as a function of group-level contingencies and discuss the motivational shifts that follow dynamic changes in self-representation and their implications for decision making. We outline the role of social norms and institutions in collectively oriented decisions and conclude by describing the contours of a potential computational framework for modeling the role of dynamic self-representation in decision making.

A Classic View: Stable Identities and Individualistic Preferences

Models of human behavior have often made two assumptions: (1) that people possess reasonably stable preferences and goals and (2) that these preferences are generally individualistic in nature (Tversky & Kahneman, 1986). Much existing literature in psychology, economics, and biology endorses the notion that people—and by extension self-representations—are largely stable. For instance, extensive research on personality has sought to identify traits that capture enduring and essential differences between people, assuming that although people differ meaningfully from each other, there is less meaningful variation within individuals. Similarly, rational choice models in economics have traditionally assumed that preferences possess the property of transitivity, such that preferences for different outcomes are ordered in a stable and internally coherent manner, ensuring consistency in decision making (Elster, 2007). Theorists in personality psychology (Mischel & Shoda, 1995) and economics (Gigerenzer, 2002; Simon, 1982) have long recognized that stability is an assumption that is often not met. But stability is often treated as a normative baseline from which interesting deviations can be catalogued and explained.

Rational-choice perspectives have often also assumed that people think and act on the basis of what they perceive as their individual self-interest (Blau, 1964; Hardin, 1968; Homans, 1961; Olson, 1965; Von Neuman & Morgenstern, 1947). Individualistic models can explain a great deal of human cognition and behavior, but they are often unable to account for common patterns of prosocial behavior in which people make decisions that better promote others' or their groups' interests rather than their own (Akerlof & Kranton, 2010; E. Ostrom, 2005). The difficulty of modeling actual human behavior in entirely individualistic terms is exemplified by social dilemma situations, in which individual-level interests are pitted against collective interests (Kollock, 1998; Messick & Brewer, 1983). Standard rational-choice models predict that decisions that favor collective interests in these

dilemmas should rapidly approach zero (Hardin, 1968). In actuality, however, rates of cooperation tend to be significantly higher than rational–choice theories would generally predict (Camerer, 2003; E. Ostrom, 2005). Theorists have suggested (and data confirm) that cooperative behavior can be maintained in social dilemmas when third parties punish noncooperative behavior (Fehr & Fischbacher, 2004; Fehr & Gächter, 2000). The threat of punishment alters the incentive structure such that individuals' interests are better served by cooperation because there are now costs associated with noncooperation. However, this is, at best, a partial solution because punishing noncooperators is itself an individually costly behavior, theoretically maintainable only when failures to punish are themselves punished by fourth–party punishers, who presumably must be punished for failing to punish nonpunishers, and so on *ad infinitum* (Lewisch, Ottone, & Ponzano, 2011).

Other models instead attempt to account for cooperative and prosocial behavior in terms of stable preferences and goals by relaxing the assumption that people are uniformly motivated by individual self-interest. In these models, interest in other people's outcomes or collective outcomes is integrated into a stable utility function. People may vary, for example, in their social value orientations (SVO), such that there are stable individual differences in the nature of the outcomes to which decision makers are attuned (Kuhlman & Marshello, 1975; Messick & McClintock, 1968; Van Lange, 2000).[1] SVO and related individual difference approaches account for interindividual variation in cooperative and prosocial behavior by positing that people possess different sorts of goals. These models are less attentive, however, to intra–individual variation, such that the same people may behave quite differently from one situation to another (Akerlof & Kranton, 2010).

An Alternative Perspective: The Dynamic Self

Although human behavior is often modeled in terms of stable personal preferences, actual behavior is typically more variable and less individualistic than would expected given these assumptions. To account for these characteristics of decision making, researchers (particularly in social psychology) have posited that self-representations change rapidly to accommodate shifting goals and contextual influences and that dynamic changes in self-representation mediate a great deal of variation in human social behavior. Major theories posit that the self is multifaceted and context dependent (Aron, Aron, Tudor, & Nelson, 1991; Markus & Nurius, 1986; Simmel & Wolf, 1950; Tajfel, 1982; Turner et al., 1994). Individuals can categorize themselves according to multiple dimensions (e.g., age, gender, race, occupation, nationality), and the psychological salience of any of these identities can shift relatively quickly, leading to the online and *ad hoc* construction of self.

Here, we focus on cognitive, motivational, and behavioral consequences as the self-concept shifts from the *personal* level (i.e., defining oneself as unique from others—the "I" or "me") to the *collective* level (i.e., defining oneself in terms of characteristics of one's social group—the "we" or "us"; Turner et al., 1987).

Personal identity tends to be more salient in intragroup contexts, whereas collective or social identity tends to be more salient in intergroup contexts (Hogg & Turner, 1987). As such, self-representation becomes more inclusive as the categorization context is broadened, leading to the inclusion of others who were formerly deemed distinct from the self (Gaertner, Mann, Murrell, & Dovidio, 1989; Turner et al., 1994). This form of "self-expansion" occurs when individuals join and identify with groups, which can range in size from dyadic interpersonal relationships to all of humanity (Brewer & Gardner, 1996; Buchan et al., 2011).

In recent years, researchers have made major strides toward understanding the dynamic, multifaceted nature of self-representation. For instance, studies have shown that relationship partners (Aron et al., 1991) and groups (e.g., fraternities or sororities) become semantically associated with the self (Smith & Henry, 1996)—especially when people identify with the ingroup (Brewer & Pickett, 1999). Moreover, connectionist models suggest that representations of the self and others are not independent structures stored separately in memory but rather reflect online constructions derived from contextually influenced patterns of activation in neural networks (Smith, Coats, & Walling, 1999). These findings reflect the idea that close relationships and group memberships both involve dynamic construction of collective self-representation (Aron et al., 1991; Turner et al., 1987; Van Bavel & Cunningham, 2011).

Effects on Perceptual, Evaluative, and Neural Processes

Our primary contention in this chapter is that changes in self-representation involve not only cognitive changes (e.g., in explicit knowledge about identity), but also a fundamental reconfiguration of the entire motivational and decision-making system. One major implication of a dynamic-self approach is that social identities are likely to shape group members' perception and evaluation of the social world and perhaps even physical reality (Caruso, Mead, & Balcetis, 2009; Hastorf & Cantril, 1954). These processes are not necessarily deliberative or conscious, and we suggest that dynamic aspects of self-representation can shape rapid and ostensibly automatic reactions. When a specific category membership becomes salient, people are more likely to see the world through the lens of this particular identity (Xiao & Van Bavel, 2012). In this way, different social identities may change how people reflexively attend to (Brosch & Van Bavel, 2012) and evaluate (Ashburn-Nardo, Voils, & Monteith, 2001; Otten & Wentura, 1999; Van Bavel & Cunningham, 2009) stimuli in their environment.

Based on these assumptions, we have hypothesized that constructing a novel collective identity would alter rapid evaluations, and possibly override implicit racial biases based on years of experience and exposure to racial stereotypes. To test this prediction, we conducted a series of experiments examining the influence of a minimal group identity on ostensibly automatic racial biases in evaluations (Van Bavel & Cunningham, 2009) and amygdala activity (Van Bavel, Packer, &

Cunningham, 2008). Participants in these experiments were assigned to one of two mixed-race groups (e.g., the Lions or Tigers) with an equal number of Black and White males in each group. Participants were then given a few minutes to memorize the group membership of these faces before we assessed their attitudes (Van Bavel & Cunningham, 2009), memory (Hehman, Maniab, & Gaertner, 2010; Shriver, Young, Hugenberg, Bernstein, & Lanter, 2008; Van Bavel & Cunningham, 2012; Van Bavel, Swencionis, O'Connor, & Cunningham, 2012), and brain activity (Van Bavel et al., 2008; Van Bavel, Packer, & Cunningham, 2011). Assigning people to mixed-race groups allowed us to examine whether a currently shared group membership could override ostensibly automatic racial biases in categorization and evaluation (Brewer, 1988; Devine, 1989; Ito & Urland, 2005)

In a pair of initial experiments, we measured ostensibly automatic evaluations of the faces described using a response-window priming task (Van Bavel & Cunningham, 2009), which allowed us to assess very rapid evaluations of ingroup versus outgroup members (Cunningham, Preacher, & Banaji, 2001; Draine & Greenwald, 1998).[2] As predicted, participants who were assigned to a mixed-race group exhibited positive evaluations of ingroup members, regardless of their race. Specifically, we found that group membership increased positivity toward Black ingroup members relative to Black outgroup members, eliminating the standard pattern of automatic racial bias (Devine, 1989; Fazio, Jackson, Dunton, & Williams, 1995). Thus, participants' evaluations reflected their current self-categorization with a minimal group, even when the ingroup and outgroup had no history of contact or conflict, and when there was an orthogonal, visually salient social category cue (i.e., race) with strong existing evaluative connotations.

This dynamic shift in preferences suggests that social categorization—even with a novel group—can tune seemingly automatic preferences. These promising results lead us to examine the neural mechanisms underlying this pattern of ingroup bias (Van Bavel et al., 2008). Several earlier studies in the domain of social neuroscience had identified a relationship between activity in the amygdala—a small structure in the temporal lobe—and implicit racial bias (Amodio, Harmon-Jones, & Devine, 2003; Cunningham et al., 2004; Phelps et al., 2000). Work in behavioral and cognitive neuroscience, in which the amygdala had been implicated in fear conditioning (LeDoux, 2000) and processing negative stimuli (Cunningham, Johnson, Gatenby, Gore, & Banaji, 2003; Hariri, Tessitore, Mattay, Fera, & Weinberger, 2002) led to the inference that this particular region might reflect fear toward racial outgroup members. However, other research has shown that the amygdala is sensitive to any motivationally relevant input—even positive stimuli (Anderson et al., 2003; Cunningham, Van Bavel, & Johnsen, 2008; Hamann, Ely, Hoffman, & Kilts, 2002). As such, we predicted that amygdala activity would be greater for ingroup members because in many situations, ingroup members—as part of an expanded collective self-representation—are more motivationally important than outgroup members. As predicted, participants had greater amygdala activity to ingroup members than outgroup members (see also Chiao et al.,

2008). Further, this pattern of ingroup bias was not moderated by target race or categorization task (i.e., categorizing by team or by skin color), suggesting that it occurred relatively reflexively. This research suggests that racial biases in amygdala activity are *not* inevitable or hardwired (see also Wheeler & Fiske, 2005). Rather, activity in this region appears to be sensitive to dynamic changes in self-representation as a function of the social context.

We have also examined the effects of social identity on the own-race bias (ORB)—a phenomenon whereby people are better at remembering people from their own race than people from other races. Although the ORB may appear to be relatively innocuous relative to other forms of intergroup bias, it can have severe consequences in the legal domain. For instance, racial biases in face memory may cause eyewitnesses to misidentify suspects from another race, which could lead to the conviction of innocent persons (Brigham & Ready, 2005). In fact, the majority of false convictions of criminals on death row in the United States are based on a cross-race eyewitness misidentification in which a White eyewitness falsely accused a Black defendant (Scheck, Neufeld, & Dwyer, 2000). Traditional explanations for the ORB suggest that people are better at within-race than between-race recognition because they tend to have greater *perceptual expertise* with own-race faces (Malpass & Kravitz, 1969). Recently, however, researchers have proposed that these effects may be more motivational in nature, such that people are motivated to encode ingroup members at a subordinate level (i.e., as individuals) and outgroup members at a superordinate level (i.e., as exemplars of their category; Bernstein, Young, & Hugenberg, 2007; Hugenberg, Young, Bernstein, & Sacco, 2010). We hypothesized that self-categorization with a minimal group might have a similar effect—leading people to encode ingroup members at a subordinate level and outgroup members at a superordinate level. Consistent with this prediction, we found greater activation within the bilateral fusiform gyri—a region of the occipito-temporal lobe involved in face perception (Kanwisher, McDermott, & Chun, 1997) and perceptual expertise (Gauthier, Skudlarski, Gore, & Anderson, 2000)—when members of arbitrary groups viewed ingroup versus outgroup faces (Van Bavel et al., 2008). In a subsequent study, we not only replicated this pattern of ingroup bias in a subregion of the fusiform gyri sensitive to faces called the fusiform face area (FFA; Kanwisher et al., 1997), but we also found that greater ingroup bias in FFA activity mediated the effect of group membership on recognition memory—a behavioral index of individuation (Van Bavel et al., 2011). Specifically, people who had greater activation in the FFA to ingroup versus outgroup faces also had better recognition memory for ingroup over outgroup members. Importantly, there was no main effect of race, and the effects of group membership were not moderated by target race (see also Hehman et al., 2010; Shriver et al., 2008). These results provide evidence that brain regions involved in visual perception, such as the fusiform gyri, are sensitive to shifts in self-categorization, responding selectively to face stimuli imbued with psychological significance by virtue of their group membership.

A dynamic-self approach suggests that group identification affords individuals the opportunity to pursue individual motives at the collective level. Therefore, we also examined the role of collective identification and social motives in this pattern of ingroup bias. Specifically, superior memory for ingroup compared to outgroup members was found only among people who were highly identified with their minimal ingroup (Van Bavel & Cunningham, 2012) or who had a heightened need to belong (Van Bavel et al., 2012). In other words, people for whom the ingroup was important or motivationally relevant were the most likely to selectively encode and recognize ingroup relative to outgroup members. It is also important to note that this type of ingroup bias is not inevitable and that visual and memory processes can rapidly shift to reflect changes in the intergroup context, as well as one's role within the group. For example, we found that enhanced memory for ingroup members was reduced when people were assigned to a role (i.e., spy) that motivated them to attend to outgroup members (Van Bavel & Cunningham, 2012). Thus, our research provides evidence that self-categorization with a group—even a minimal group—can shape the *motivational relevance* of categories in a flexible and dynamic fashion even in the absence of long-term experience with the categories in question.

Motivational Shifts

Changes in self-representation not only entail alterations in self-knowledge but also involve changes in motivation and decision making. When categorization of the self shifts from an individual to a collective level, events and outcomes that were perceived and evaluated in terms of consequences for the individual are more likely to be evaluated in terms of consequences for the collective. These shifts enable people to detect and exploit environmental contingencies that operate at levels higher than the individual (Brewer, 2004; Correll & Park, 2005). These sorts of situations are exemplified by social dilemmas in which collective contingencies differ from those incurred by individuals (Dawes, 1980; Kollock, 1998; Messick & Brewer, 1983). Collective contingencies can reflect positive opportunities, as in the case of public goods dilemmas when a community as a whole can benefit from resources that can only be created or harvested by cooperative collective effort. Examples include agriculture, public railways, roads, fire stations, and broadcasting. Collective contingencies can also be negative, as when a community works together to avoid a negative outcome. In commons dilemmas, for example, overuse of a resource by separate individuals may deplete it to the point of collapse (e.g., fisheries), a calamity they might be able to avoid by coordinating behavior as a group. People are more attuned to group contingencies and more willing to respond in collectively oriented ways when they self-categorize as members of a group.

Dynamic self-representation is often functional because it provides a means by which the human cognitive system can take advantage of positive collective

contingencies and avoid negative ones. When self-representation shifts from a personal to a collective level, contingencies and events that would ordinarily be evaluated in reference to the individual and his/her goals are now evaluated in reference to the currently relevant social identity—promoting group success (and, often, individual success). This has a number of implications. First, it suggests that many of the same motives that animate cognition and behavior at the individual level are likely to operate at the collective level as well. This appears to be case, such that with respect to both individual selves and group memberships, people are motivated to self-enhance and gain status (Tajfel & Turner, 1979; Valdesolo & DeSteno, 2007), self-verify or maintain stability (Chen, Chen, & Shaw, 2004), define boundaries (Brewer, 1991), and improve (Packer, 2008).

More provocatively, a dynamic-self approach predicts that when people identify with a group, they are more likely to make decisions that benefit the group, even if doing so involves personal costs. In contrast to individual-difference models of preference (e.g., SVO), the dynamic perspective posits that acting in the collective interest should be contextually variable. A great deal of prosocial behavior occurs when the self shifts from an individual to a collective level such that *self-interest* is extended to the collective self (De Cremer & Stouten, 2003; De Cremer & Van Vugt, 1999). Importantly, however, this means that cooperative and prosocial behavior is also often parochial in nature, such that it is extended to ingroup members but withheld from outgroup members (Bernhard, Fischbacher, & Fehr, 2006; Choi & Bowles, 2007; Tajfel & Turner, 1979).

Across a variety of experiments, individuals within group contexts have been observed to favor ingroup over personal outcomes (Brewer & Kramer, 1986; Kramer & Brewer, 1984; Van Vugt & Hart, 2004; Zdaniuk & Levine, 2001). As a result, when shared group memberships are salient, the frequency of cooperative and collectively oriented decisions (i.e., decisions that put individual self-interest at risk) are increased in a variety of social dilemmas (Dion, 1973; Miller, Downs, & Prentice, 1998; Tanis & Postmes, 2005; Wit & Kerr, 2002; see also Brewer, 2004; Brewer & Kramer, 1986; De Cremer & Van Vugt, 1999). Evidence further suggests that dynamic-self processes interact with individual preferences and predispositions to predict prosocial attitudes and cooperative behavior. For example, De Cremer and Van Vugt (1999) examined contributions in public goods dilemmas as a joint function of individuals' SVO and level of group identification. They found that participants classified as "prosocials" behaved cooperatively regardless of how connected they felt to the group. Critically, however, identification (measured or manipulated) had a dramatic effect on the behavior of "proselfs." As is typically observed, these individuals were relatively uncooperative when they did not identify with the group. But they were as cooperative as "prosocials" when they identified with the group. These data suggest that both preference-based and dynamic-self approaches may be relevant for understanding prosocial behavior and show that ordinarily selfish people can undergo a transformation of goals from the individual to the collective level when they identify with a group. Moreover,

shifts in self-representation and prosocial behavior are dynamically determined by aspects of the social context, not merely stable individual differences in preferences or collective identification.

There is also substantial evidence that the prosociality extended to ingroups is parochial, such that it stops at the group boundary (Abbink, Brandts, Herrmann, & Orzen, 2012); Bernhard et al., 2006; Choi & Bowles, 2007; Tajfel & Turner, 1979; Valdesolo & DeSteno, 2007). Again, this is illustrated dramatically by the original minimal group studies, which showed that people would often discriminate against outgroups even when ingroup and outgroup outcomes were completely independent of one another (i.e., when discrimination offered no material benefit for their own group). Consistent with a dynamic-self approach, however, when intergroup boundaries are redrawn so that people come to perceive ingroup and outgroup members as sharing a common superordinate identity, attitudes toward outgroup members improve and intergroup discrimination is reduced (e.g., Gaertner et al., 1989; Van Bavel & Cunningham, 2009; Wohl & Branscombe, 2005).

Institutions and Social Norms

As shifts in self-representation orient people toward collective outcomes in decision making, groups need to align their members' goals, such that there is reasonably broad agreement regarding the collective interest. If every group member had their own unique evaluation of group interests and there were no other mechanisms available to facilitate coordination, group activity could descend into chaos. Human groups adopt a variety of mechanisms to coordinate activity among their members, ranging from formal institutions to public rituals to tacit and often unspoken social norms (e.g., Abrams, Wetherell, Cochrane, Hogg & Turner, 1990; Chwe, 2001; Fehr & Fischbacher, 2004; Ostrom, 2005; Terry & Hogg, 1996). Large groups, including companies, cities, and nation states, use institutions to formalize collective decision making and regulate the behavior of group members (Ostrom, 2005; Shariff & Norenzayan, 2007). Less formally, groups regulate their members' behavior by developing social norms—widely held conceptions of how we do and should behave (Abrams et al., 1990; Campbell, 1990; Cialdini, Reno, & Kallgren, 1990). Conforming to norms, as well as enforcing them, appears to be a deeply held human impulse. People look to group norms to help determine the nature of reality and efficacious courses of action, as well as socially appropriate behavior (see Asch, 1955; Deutsch & Gerard, 1955; Hodges & Geyer, 2006; Jacobs & Campbell, 1961; Levine, 1999). When people deviate from norms, they are often punished, and this punishment is typically parochial, such that people are more punitive toward ingroup than outgroup members who engage in deviant behaviors (e.g., Marques & Yzerbyt, 1988). This is particularly the case when deviance puts the ingroup or its reputation at risk (Garcia, Horstman, Amo, Redersdorff, & Branscombe, 2005) and when people highly identify with the ingroup (Branscombe, Wann, Noel, & Coleman, 1993). Group norms can help members

converge on shared conceptions of collective interests, ensuring that they generally act in coordinated and thus collectively efficacious ways.

Critically, however, a dynamic-self approach posits that operating at a collective level of self does not entail automatic or passive conformity to group norms. Rather, the fact that identified group members engage in their own computations about collective utility suggests that they may dissent from normative patterns of behavior if they perceive another course of action as being in the collective interest (Packer, 2008; Packer & Miners, in press). In other words, because individual group members exert their own agency with regard to collective interests, they may sometimes deviate from and act to change normative patterns of group behavior.

We tested this prediction in a series of studies (Packer, 2009; Packer & Chasteen, 2010; Packer, Fujita, & Chasteen, in press). Generally, members who report being strongly identified with a group are more likely than weakly identified members to conform to group norms (Packer, 2009; Terry & Hogg, 1996). We hypothesized, however, that this would not be the case when strongly identified members were aware that a particular social norm was harmful to the collective interests of their group. By measuring or manipulating perceptions of collective harm, we found that strong identifiers are less conforming when they believe a norm is counter to the interests of their group and are more likely to engage in acts of dissent intended to change the norm to benefit the long-term success of the group (Packer, 2009; Packer & Chasteen, 2010). Consistent with our claim that shifts to a collective level of self entail pursuit of collective rather than individual interests, when identified group members are made aware that a group norm harms their personal self-interest, dissent against the norm does not increase (Packer, 2010; Packer & Chasteen, 2010).

This line of research has also allowed us to examine how analogous the dynamics of the motivational processes are at the collective versus individual level. Individual-level self-interest is multifaceted. Even when people are focused entirely on what is good for *me*, what is optimal differs across situations and, in particular, as a function of time. Many behaviors that maximize individual utility in the short term (e.g., eating a delicious dessert) reduce utility in the long run (e.g., odds of a healthy old age; Fujita, 2011; Thaler & Shefrin, 1981). These sorts of competing contingencies create self-control conflicts at the individual level. Recent research has shown that people are more likely to make decisions that favor their long-term interests when they focus on the future or when they approach decisions in a relatively abstract mindset (Fujita, Trope, Liberman, & Levin-Sagi, 2006; Malkoc, Zauberman, & Bettman, 2010). We hypothesized that collective interests are also multifaceted and that similar dynamics would occur among individuals making decisions at a collective level of self (Packer, Fujita, & Chasteen, in press; Packer, Fujita, & Herman, in press). In the short term, a group's interests may be best served by decisions that maintain stability, cohesion, and immediate effectiveness. In the long term, however, a group may be better served by decisions that disrupt stability in favor of change and improvement. We predicted that the same factors that orient people toward longer-term individual interests—a future temporal

AuQ18

AuQ19

AuQ20

focus and abstract mindset (Fujita et al., 2006)—would cause them to make decisions that favor their groups' long term interests and thus increase dissent against problematic group norms. This is exactly what we found: Strongly identified group members were more likely than weak identifiers to challenge group norms in a future-oriented or abstract mindset. In contrast, they were more likely than weak identifiers to conform to group norms when thinking concretely. These data indicate that individuals engage in their *own* computations about collective utility and that those computations are shaped by the same types of motivational dynamics that influence decision making with regard to individual-level concerns.

Contours of a Computational Approach to the Dynamic Self

In the final section, we consider how the dynamic-self approach might be integrated with current models of decision making. It is certainly the case that goal-driven behavior changes to incorporate group-level interests as a function of shifts in self-representation. Precisely how this is accomplished, however, remains largely unknown. We believe that the next major frontier in this field of research will be to investigate the underlying computations involved in these changes. As a starting point, we have adapted a neurobiological framework of (individual-level) goal-directed decision making, which posits that decision making consists of five types of processes (Rangel, Camerer, & Montague, 2008). In an initial *representation* phase, individuals represent the field of possible actions, internal (need or goal) states, and external states of the environment (e.g., affordances).[3] In a second, *valuation* phase, the value of possible actions is computed. This involves computing associations between each action and outcomes, as well as associations between outcomes and value to the individual. The former will be affected, in particular, by environmental affordances and the latter affected by the individual's current goal states. Third, the individual enters an *action selection* phase based on comparison of the values assigned to possible actions. At this point, the individual decides to enact one action (or inaction) from among the set of represented and evaluated possibilities. After which, the individual engages in *outcome evaluation* and *learning*, during which representation, valuation, and action-selection processes are updated in order to improve subsequent decisions. In this framework, as in most standard decision-making models, goals, actions, outcomes, and values are computed in reference to the individual organism. Sophisticated goal-directed cognition allows individuals to learn and respond to complex and diverse contingencies, perhaps approximating utility maximization given current goals and environmental affordances. These phases may occur rapidly and be implemented below conscious awareness.

Integrating a dynamic self into this type of model dramatically increases its computational complexity because the organism must compute goals, affordances, actions, outcomes, and values and engage in action selection in reference to multiple and contextually variable points of reference. A major challenge and

an opportunity for models of decision making is to identify how collective self-representation might affect the decisional processes outlined earlier. At the representational stage, an individual may represent the physical world according to the goals and norms of their group. They may also represent an expanded set of possible actions, incorporating actions that can be made individually with respect to group interests, as well as behaviors that could only be enacted successfully in collaboration with others. Representations presumably also expand to include computation of group-level need states and goals, as well as affordances that affect possibilities for effective group-oriented or collective action. Given representation of these goals and affordances, the valuation phase presumably involves computing group-level contingencies: associations between actions and group outcomes and between outcomes and value to the group. During action selection, a highly valued action (with respect to group goals) is presumably selected from among the set of represented possibilities. Following an action (or inaction), the individual will likely engage in *outcome evaluation* and *learning*, during which representation, valuation, and action-selection processes are updated with respect to collective oriented actions and goals.

Major questions arise from this type of framework, including how dynamic self-representation interfaces with the evaluation and decision-making system. One possibility is that group-level self-representations are modeled by a separate set of cognitive processes that then modulate evaluative and decisional processes in a top-down fashion. It is probably fair to say that this is the dominant (if tacit) assumption among group/self theorists. Social identity and self-categorization theory, for example, suggest that collective self-representation emerges from basic categorization processes (Turner et al., 1987). Having established a group-level sense of self, it can then influence goal-directed decision making. Alternately, collective self-representation may arise when valuation and outcome evaluation processes detect that certain coordinated actions result in shared outcomes and that certain outcomes have shared value. This might have emerged initially as part of a learning process when humans began to engage in ritualistic activity (Sosis & Ruffle, 2003; Wiltermuth & Heath, 2009); now, individuals may also use heuristic cues of group relevance (e.g., being assigned to a team) to trigger computation of shared value. Collective-level value computations might then feed back to influence representation—directing attention to affordances that affect the relevant group and triggering computation of collective-level need states and goals. As such, dynamic self-representation may arise as an emergent property from detection of shared value. We suspect that both of these processes may be at play during collective-level decision making.

Conclusion

The concept of collective self-representation can be understood in different ways. It might be considered a metaphor—that there can be something self-like about group memberships, such that people exhibit some similar proclivities and

behaviors when responding on behalf of their groups as they do when responding in their individual self-interest. Most self researchers, however, believe that social identities are more than a metaphor. From this perspective, the cognitive apparatus that gives rise to an individual sense of self are also involved in generating, maintaining, and acting in the service of social identities. On this basis, the self is constructed ad hoc, by contextually shaped representations of self. When a group membership is relevant, working representations of the self rapidly shift from the individual level (who am I, and what do I want?) to the collective level (who are we, and what do we want?).

The dynamic nature of the self has important implications for understanding change—both within persons and within groups and social systems at large. In contrast to models of personhood that assume largely stable preferences and dispositions, the research we have reviewed on shifts in self-representation indicates that preferences and behavioral patterns are often highly situationally variable. Change the context, change the self. Similarly, intergroup biases are—often surprisingly—malleable. People orient to the contextually relevant social categories and even very rapid ("implicit") evaluative reactions follow suit (Van Bavel & Cunningham, 2009). As such, rather than reflecting stable, slow-learned, and inevitable responses, the evaluations indexed by popular implicit measures appear to capture contextually flexible online computations of value that are sensitive to dynamic self-representations and their associated goals. Finally, we argue that the dynamic-self approach can account for collective change, including when and why group members are motivated to dissent from and challenge stable normative patterns of behavior. Because adopting a collective level of self enables individuals to engage in their own assessments of group outcomes, they can critically evaluate the behaviors of their own groups and may—when sufficiently motivated—attempt to change them.

It may be no accident that the species that displays by far the greatest behavioral flexibility is also the species with the most developed sense of self. The dynamic self allows for the embrace of shifting coalitions and group memberships. It also allows people to take an alternate perspective on reality—established individual habits or collective normative patterns—and envision alternate possibilities: This is not who I am, this is not who we are. The self is, as Charles Taylor (1989) put it, "the horizon within which I am capable of taking a stand."

Notes

1. The SVO construct divides people into four categories: the most common are *individualists*, who are concerned with maximizing their own outcomes, and *prosocials*, who would prefer, when possible, to maximize joint outcomes. Less common categories are composed of *competitors*, who seek to maximize differences between their outcomes and others', and *altruists*, who seek to maximize others' outcomes.
2. In this task, participants are presented in quick succession with a face prime (150 milliseconds) followed by a target word, which they have to classify as positive or negative

within a restricted time window (525 milliseconds). By examining error rates to positive versus negative words following different classes of primes (e.g., ingroup versus outgroup faces), we are able to assess very rapid evaluations of social categories (Cunningham, Preacher, & Banaji, 2001; Draine & Greenwald, 1998).

3. By affordances, we refer to **opportunities** for and **constraints** on action provided by the organism's current environment.

References

Abbink, K., Brandts, J., Herrmann, B., & Orzen, H. (2012). Parochial altruism in inter-group conflicts. *Economics Letters, 117*, 45–48.

Abrams, D., Wetherell, M., Cochrane, S., Hogg, M. A., & Turner, J. C. (1990). Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization. *British Journal of Social Psychology, 29*, 97–119.

Akerlof, G. A., & Kranton, R. E. (2010). *Identity economics: How our identities shape our work, wages, and well-being*. Princeton, NJ: Princeton University Press.

Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison Wesley.

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

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., & Glover, G. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience, 6*, 196–202.

Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Close relationships as including other in the self. *Journal of Personality and Social Psychology, 60*, 241–253.

Asch, S. E. (1955). Opinions and social pressure. *Scientific American, 193*, 31–35.

Ashburn-Nardo, L., Voils, C. I., & Monteith, M. J. (2001). Implicit associations as the seeds of intergroup bias: How easily do they take root? *Journal of Personality and Social Psychology, 81*, 789–799.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*, 497–529.

Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature, 442*, 912–915.

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The Cross-Category Effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science, 18*, 709–712.

Billig, M. (1976). *Social psychology and intergroup relations*. London: Academic Press.

Billig, M., & Tajfel, H. (1973). Social categorization and similarity in intergroup behaviour. *European Journal of Social Psychology, 3*, 27–52.

Blau, P. (1964). *Exchange and power in social life*. New York: Wiley.

Branscombe, N. R., Wann, D. L., Noel, J. G., & Coleman, J. (1993). Ingroup or outgroup extremity: Importance of the threatened identity. *Personality and Social Psychology Bulletin, 19*, 381–388.

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86*, 307–324.

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Lawrence Erlbaum.

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality & Social Psychology Bulletin, 19*, 475–482.

Brewer, M. B. (2004). Taking the social origins of human nature seriously: Toward a more imperialist social psychology. *Personality and Social Psychology Review, 8*, 107–113.

Brewer, M. B. (2008). Depersonalized trust and ingroup cooperation. In J. I. Krueger (Ed.), *Rationality and social responsibility* (pp. 215–232). New York: Psychology Press.

Brewer, M. B., & Gardner, W. L. (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.

Brewer, M. B., & Pickett, C. L. (1999). Distinctiveness motives as a source of the social self. In T. Tyler, R. Kramer, & O. John (Eds.), *The psychology of the social self* (pp. 71–87). Mahwah, NJ: Erlbaum.

Brigham, J. C., & Ready, D. J. (2005). Own-race bias in lineup construction. *Law and Human Behavior, 9*, 415–424.

Brosch, T., & Van Bavel, J. J. (2012). The flexibility of emotional attention: Accessible social identities guide rapid attentional orienting. *Cognition, 125*, 309–316.

Buchan, N. R., Brewer, M. B., Grimalda, G., Wilson, R. K., Fatas, E., & Foddy, M. (2011). Global social identity and global cooperation. *Psychological Science, 22*, 821–828.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction.* Princeton, NJ: Princeton University Press.

Caporael, L. R. (1997). The evolution of truly social cognition: The core configurations model. *Personality and Social Psychology Review, 1*, 276–298.

Caruso, E., Mead, N., & Balcetis, E. (2009). Political partisanship influences perception of biracial candidates' skin tone. *Proceedings of the National Academy of Sciences, 106*, 20168–20173.

Chen, S., Chen, K. Y., & Shaw, L. (2004). Self-verification motives at the collective level of self-definition. *Journal of Personality and Social Psychology, 86*, 77–94.

Chiao, J. Y., Iidaka, T., Gordon, H. L., Nogawa, J., Bar, M., & Aminoff, E. (2008). Cultural specificity in amygdala response to fear faces. *Journal of Cognitive Neuroscience, 20*, 2167–2174.

Choi, J. K., & Bowles, S. (2007). The coevolution of parochial altruism and war. *Science, 318*, 636–640.

Chwe, M. S-Y. (2001). *Rational ritual: Culture, coordination, and common knowledge.* Princeton, NY: Princeton University Press.

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology, 58*, 1015–1026.

Correll, J., & Park, B. (2005). A model of the ingroup as a social resource. *Personality and Social Psychology Review, 9*, 341–359.

Cunningham, W. A., Johnson, M. K., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2003). Neural components of social evaluation. *Journal of Personality and Social Psychology, 85*, 639–649.

Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science, 15*, 806–813.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12*, 163–170.

Cunningham, W. A., Van Bavel, J. J., & Johnsen, I. R. (2008). Affective flexibility: Evaluative processing goals shape amygdala activity. *Psychological Science, 19*, 152–160.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology, 31*, 169–193.

De Cremer, D., & Stouten, J. (2003). When do people find cooperation most justified? The effect of trust and self–other merging in social dilemmas. *Social Justice Research, 16*, 41–52.

De Cremer, D., & Van Vugt, M. (1999). Social identification effects in social dilemmas: A transformation of motives. *European Journal of Social Psychology, 29*, 871–893.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology, 51*, 629–636.

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

Dion, K. L. (1973). Cohesiveness as a determinant of ingroup-outgroup bias. *Journal of Personality and Social Psychology, 28*, 163–171.

Draine, S. C., & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General, 127*, 286–303.

Ellemers, N., Wilke, H., & Van Knippenberg, A. (1993). Effects of the legitimacy of low group or individual status on individual and collective status-enhancement strategies. *Journal of Personality and Social Psychology, 64*, 766–778.

Elster, J. (2007). *Explaining social behavior: More nuts and bolts for the social sciences*: Cambridge, UK: Cambridge University Press.

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

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior, 25*, 63–87.

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review, 90*, 980–994.

Fujita, K. (2011). On conceptualizing self-control as more than the effortful inhibition of impulses. *Personality and Social Psychology Review, 15*, 352–366.

Fujita, K., Trope, Y., Liberman, N., & Levin-Sagi, M. (2006). Construal levels and self-control. *Journal of Personality and Social Psychology, 90*, 351–367.

Gaertner, S. L., Mann, J., Murrell, A., & Dovidio, J. F. (1989). Reducing intergroup bias: The benefits of recategorization. *Journal of Personality and Social Psychology, 57*, 239.

Garcia, D. M., Horstman Reser, A., Amo, R. B., Redersdorff, S., & Branscombe, N. R. (2005). Perceivers' responses to in-group and out-group members who blame a negative outcome on discrimination. *Personality and Social Psychology Bulletin, 31*, 769–780.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*, 191–197.

Gigerenzer, G. (2002). *Bounded rationality: The adaptive toolbox*. Cambridge, MA: MIT Press.

Hamann, S. B., Ely, T. D., Hoffman, J. M., & Kilts, C. D. (2002). Ecstasy and agony: Activation of the human amygdala in positive and negative emotion. *Psychological Science, 13*, 135–141.

Hardin, G. (1968). The tragedy of the commons. *Science, 162*, 1243–1248.

Hariri, A. R., Tessitore, A., Mattay, V. S., Fera, F., & Weinberger, D. R. (2002). The amygdala response to emotional stimuli: A comparison of faces and scenes. *Neuroimage, 17*, 317–323.

Hastorf, A. H., & Cantril, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology, 49*, 129–134.

Hehman, E., Maniab, E. W., & Gaertner, S. L. (2010). Where the division lies: Common ingroup identity moderates the cross-race facial-recognition effect. *Journal of Experimental Social Psychology, 46*, 445–448.

Hodges, B. H., & Geyer, A. L. (2006). A nonconformist account of the Asch experiments: Values, pragmatics, and moral dilemmas. *Personality and Social Psychology Review, 10*, 2–19.

Hogg, M. A., & Turner, J. C. (1987). Intergroup behaviour, self-stereotyping and the salience of social categories. *British Journal of Social Psychology, 26*, 325–340.

Homans, G. C. (1961). *Social behavior: Its elementary forms*: London: Taylor & Francis.

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The Categorization-Individuation Model: An integrative account of the other-race recognition deficit. *Psychological Review, 117*, 1168–1187.

Ito, T. A., & Urland, G. R. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective, and Behavioral Neuroscience, 5*, 21–36.

Jacobs, R. C., & Campbell, D. T. (1961). The perpetuation of an arbitrary tradition through several generations of a laboratory microculture. *Journal of Abnormal and Social Psychology, 62*, 649–658.

Kanwisher, N., McDermott, J., & Chun, M. (1997). The Fusiform Face Area: A module in human extrastriate cortex specialized for the perception of faces. *Journal of Neuroscience, 17*, 4302–4311.

Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology, 24*, 183–214.

Kramer, R. M., & Brewer, M. B. (1984). Effects of group identity on resource use in a simulated commons dilemma. *Journal of Personality and Social Psychology, 46*, 1044–1057.

Kuhlman, D. M., & Marshello, A. (1975). Individual differences in the game motives of own, relative, and joint gain. *Journal of Research in Personality, 9*, 240–251.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience, 23*, 155–184.

Levine, J. M. (1999). Solomon Asch's legacy for group research. *Personality and Social Psychology Review, 3*, 358–364.

Lewisch, P., Ottone, S., & Ponzano, F. (2011). Free-riding on altruistic punishment? An experimental comparison of third-party. *Review of Law and Economics, 7*, 161–190.

Malkoc, S. A., Zauberman, G., & Bettman, J. R. (2010). Unstuck from the concrete: Carryover effects of abstract mindsets in intertemporal preferences. *Organizational Behavior and Human Decision Processes, 113*, 112–126.

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other "race." *Journal of Personality and Social Psychology, 13*, 330–334.

Markus, H. R., & Nurius, P. (1986). Possible selves. *American Psychologist, 41*, 954–969.

Marques, J. M., & Yzerbyt, V. Y. (1988). The black sheep effect: Judgmental extremity towards ingroup members in inter- and intra-group situations. *European Journal of Social Psychology, 18*, 287–292.

Messick, D. M., & Brewer, M. B. (1983). Solving social dilemmas: A review. In L. Wheeler & P. Shaver (Eds.), *Review of Personality and Social Psychology* (Vol. 4, pp. 11–44). Beverly Hills, CA: Sage.

Messick, D. M., & McClintock, C. G. (1968). Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology, 4*, 1–25.

Miller, D. T., Downs, J. S., & Prentice, D. A. (1998). Minimal conditions for the creation of a unit relationship: The social bonds between birthdaymates. *European Journal of Social Psychology, 28*, 475–481.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*, 246–268.

Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups.* Cambridge, MA: Harvard University Press.

Ostrom, E. (2005). *Understanding institutional diversity.* Princeton, NJ: Princeton University Press.

Otten, S., & Wentura, D. (1999). About the impact of automaticity in the Minimal Group Paradigm: Evidence from affective priming tasks. *European Journal of Social Psychology, 29*, 1049–1071.

Packer, D. J. (2008). On being both with us and against us: A normative conflict model of dissent in social groups. *Personality and Social Psychology Review, 12*, 50–72.

Packer, D. J. (2009). Avoiding groupthink: Whereas weakly identified members remain silent, strongly identified members dissent about collective problems. *Psychological Science, 20*, 546–548.

Packer, D. J. (2010). The interactive influence of conscientiousness and openness to experience on dissent. *Social Influence, 5*, 202–219.

Packer, D. J., & Chasteen, A. L. (2010). Loyal deviance: Testing the normative conflict model of dissent in social groups. *Personality and Social Psychology Bulletin, 36*, 5–18.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., & Gore, J. C. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12*, 729–738.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience, 9*, 545–556.

Scheck, B., Neufeld, P., & Dwyer, J. (2000). *Actual innocence: Five days to execution and other dispatches from the wrongly convicted.* New York: Doubleday.

Shariff, A. F., & Norenzayan, A. (2007). God is watching you priming God concepts increases prosocial behavior in an anonymous economic game. *Psychological Science, 18*, 803–809.

Shriver, E. R., Young, S. G., Hugenberg, K., Bernstein, M. J., & Lanter, J. R. (2008). Class, race, and the face: Social context modulates the cross-race effect in face recognition. *Personality and Social Psychology Bulletin, 34*, 260–274.

Simmel, G., & Wolf, K. H. (1950). *The sociology of Georg Simmel.* Glencoe, IL: Free Press.

Simon, H. A. (1982). *Models of bounded rationality: Empirically grounded economic reason* (Vol. 3). Cambridge: MIT Press.

Smith, E. R., Coats, S., & Walling, D. (1999). Overlapping mental representations of self, ingroup, and partner: Further response time evidence and a connectionist model. *Personality and Social Psychology Bulletin, 25*, 873–882.

Smith, E. R., & Henry, S. (1996). An in-group becomes part of the self: Response time evidence. *Personality & Social Psychology Bulletin, 22*, 635–642.

Sosis, R., & Ruffle, B. J. (2003). Religious ritual and cooperation: Testing for a relationship on Israeli religious and secular Kibbutzim. *Current Anthropology, 44*, 713–722.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223*, 96–102.

Tajfel, H. (1982). *Social identity and intergroup behavior.* Cambridge, UK: Cambridge University Press.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 94–109). Monterey, CA: Brooks-Cole.

Tanis, M., & Postmes, T. (2005). A social identity approach to trust: Interpersonal perception, group membership and trusting behaviour. *European Journal of Social Psychology, 35*, 413–424.

Taylor, C. (1989). *Sources of the self: The making of the modern identity*. Cambridge, MA: Harvard University Press.

Terry, D. J., & Hogg, M. A. (1996). Group norms and the attitude-behavior relationship: A role for group identification. *Personality and Social Psychological Bulletin, 22*, 776–795.

Thaler, R. H., & Shefrin, H. M. (1981). An economic theory of self-control. *The Journal of Political Economy, 89*, 392–406.

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: Basil Blackwell.

Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin, 20*, 454–463.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business Ethics*, S251–S278.

Valdesolo, P., & DeSteno, D. (2007). Moral hypocrisy: Social groups and the flexibility of virtue. *Psychological Science, 18*, 689–690.

Van Bavel, J. J., & Cunningham, W. A. (2009). Self-categorization with a novel mixed-race group moderates automatic social and racial biases. *Personality and Social Psychology Bulletin, 35*, 321–335.

Van Bavel, J. J., & Cunningham, W. A. (2011). A social neuroscience approach to self and social categorisation: A new look at an old issue. *European Review of Social Psychology, 21*, 237–284.

Van Bavel, J. J., & Cunningham, W. A. (2012). A social identity approach to person memory: Group membership, collective identification, and social role shape attention and memory. *Personality and Social Psychology Bulletin, 38*, 1566–1578.

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias: A functional magnetic resonance imaging investigation. *Psychological Science, 19*, 1131–1139.

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2011). Modulation of the Fusiform Face Area following minimal exposure to motivationally relevant faces: Evidence of in-group enhancement (not out-group disregard). *Journal of Cognitive Neuroscience, 23*, 3343–3354.

Van Bavel, J. J., Swencionis, J. K., O'Connor, R. C., & Cunningham, W. A. (2012). Motivated social memory: Belonging needs moderate the own-group bias in face recognition. *Journal of Experimental Social Psychology, 48*, 707–713.

Van Lange, P. A. (2000). Beyond self-interest: A set of propositions relevant to interpersonal orientations. *European Review of Social Psychology, 11*, 297–331.

Van Vugt, M., & Hart, C. M. (2004). Social identity as social glue: The origins of group loyalty. *Journal of Personality and Social Psychology, 86*, 585–598.

Von Neuman, J., & Morgenstern, O. (1947). *Theory of games and economic behavior.* Princeton, NJ: Princeton University Press.

Wheeler, M. E., & Fiske, S. T. (2005). Controlling racial prejudice social–cognitive goals affect amygdala and stereotype activation. *Psychological Science, 16*, 56–63.

Wilson, D. S., & Sober, E. (1994). Reintroducing group selection to the human behavioral sciences. *Behavioral and Brain Sciences, 17*, 585–654.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science, 20*, 1–5.

Wit, A. P., & Kerr, N. L. (2002). "Me versus just us versus us all" categorization and cooperation in nested social dilemmas. *Journal of Personality and Social Psychology, 83*, 616–637.

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

Xiao, Y. J., & Van Bavel, J. J. (2012). See your friends close, and your enemies closer: Social identity and identity threat shape the representation of physical distance. *Personality and Social Psychology Bulletin, 38*, 959–972.

Zdaniuk, B., & Levine, J. M. (2001). Group loyalty: Impact of members' identification and contributions. *Journal of Experimental Social Psychology, 37*, 502–509.