Chapter 6

An Examination of Categorization Processes in Organizations: The Root of Intergroup Bias and a Route to Prejudice Reduction

Melissa J. Ferguson and Shanette Porter

Abstract

To understand how people experience diversity, researchers have focused on category-based perceptions of others who belong to different social groups and the subsequent impact on intergroup dynamics. Specifically, scholars have focused on the automaticity of such categorizations, including stereotyping and implicit bias. This chapter will discuss these automatic, nonconscious processes, including their antecedents and their association with and effects on individual, group, and organizational outcomes. Given the automaticity of such processes, measurement strategies and suggestions for their use in future research will also be discussed.

Key Words: social identity, categorization processes, intergroup bias, diversity

The study of intergroup bias has a long history in psychology. Our understanding of this phenomenon continues to evolve as the social cognitive, intergroup, and organizational literatures grow. In an increasingly global society, where diversity of gender, ethnicity, nationality, and culture is valued but not maximally taken advantage of, understanding group processes is essential. The United States is rapidly becoming more diverse; according to the U.S. Census Bureau, by about 2040, Latinos, Blacks, and Asians will constitute more than 50% of the population. The candidate pool for jobs in America is similarly becoming more diverse (e.g., Reskin & Bielby, 2005). Recent years have also
brought increasing numbers of women qualified for competitive jobs in higher education, the corporate world, and the sciences (e.g., Fiske, 2010; Heneman, Judge, & Heneman, 1999; Offerman & Gowing, 1990). These changes in demographics offer increased opportunities for diverse work environments, which boast several positive outcomes (e.g., Boyett & Conn, 1991; Mannix & Neale, 2005; Reskin & Bielby, 2005).

In the literature, diversity can refer to a variety of types of heterogeneity, such as surface-level heterogeneity, in which visible characteristics differ within a group, or deep-level heterogeneity, in which invisible characteristics differ within the group, such as attitudes, opinions, or skills (e.g., Barrick, Stewart, Neubert, & Mount, 1998; Lawrence, 1997). Unless otherwise noted, here diversity is used to refer to surface-level, or social category, differences, such as gender, ethnicity, nationality, and so on.

Understanding and supporting diversity in the workplace, on the one hand, and inclusion practices on the other, is important for a number of reasons. For one, understanding diversity within both organizations and teams is important for combating issues of historical injustice. Segregation within organizations is associated with unequal job outcomes for ethnic and gender groups, for example (e.g., Hellerstein & Neumark, 2008). Two recent lines of research highlight other important advantages of diversity in organizations and teams. First, there is evidence that diverse environments promote prosocial outcomes for both minority and majority members, such as reduced intergroup bias, and increased helping and support for out-group members (Allport, 1954; Hurtado, Dey, Gurin, & Gurin, 2003). Second, diversity can help an organization’s bottom line: even with only one nonmajority group member, diversity is associated with the productive debate of unique ideas, which is particularly beneficial for creative problem
solving (for a review, see Mannix & Neale, 2005). And although research on the productivity of diverse groups more generally is mixed, the take-home message from this pool of research seems to be that diverse groups can produce more unique ideas, innovative responses, and creative solutions than homogenous groups—but the extent to which the benefits of diverse groups are realized depends largely upon how diversity within groups is managed by both workgroup members and superordinates (e.g., Phillips, Kim-Jun, & Shim, 2010; Phillips, Liljenquist, & Neale, 2009; Mannix & Neale, 2005).

If diversity is so beneficial and potentially effective, then why are organizations still largely homogenous at both the organizational and team level? There is no single answer to this question, as evidenced by the sheer density of research on this topic. Among the explanations is that there are still large education gaps, divided along race and gender lines. Cognitive and behavioral psychologists tend to focus instead on the intrapersonal, interpersonal, and intergroup processes that occur within organizations that might result in a lack of diversity, unequal intergroup outcomes (e.g., the tendency for traditionally lower-status ethnic group members and women to achieve lower pay and job ranks than traditionally higher-status groups), and a lack of inclusiveness in the workplace. In particular, there is much research to suggest that a preference for one’s in-group, or homophily, and anti–out-group stereotyping and bias play large roles in these outcomes (Bielby, 2000; Brown & Turner, 1981; Byrne, 1971). Many important theoretical and practical questions have been raised regarding these phenomena in this volume and elsewhere. In the following, we primarily examine just one of those questions: What are the antecedents and processes that produce intergroup stereotyping and bias in the workplace? In our examination of that question we also touch upon how
understanding the antecedents of intergroup stereotyping and bias can be informative for understanding (1) the consequences of bias and the extent to which intergroup bias is inevitable, (2) what can be done to prevent, or ameliorate the consequences of, intergroup bias and prejudice in organizations, and (3) how organizations can capitalize on diversity. Note that although it has long been noted in a variety of fields that in-group preference, stereotyping, and prejudice can be explicit, overt, and intentional, it is far more common in recent decades in America for bias to manifest in subtler and perhaps unintentional, manners, and thus this chapter will largely focus on this latter type of stereotyping and bias (e.g., Greenwald & Banaji, 1995).

**Categorical thinking**

Prejudice is one of the most insidious issues still facing society. Although this phenomenon can be complex and multiply determined, research suggests that the root cause—that is, the most distal necessary antecedent from a social psychological standpoint—is social categorization (Allport, 1954; Bodenhausen & Macrae, 1998; Brewer, 1988; Tajfel, 1978). Categorization is the grouping of similar things (in the case of social categorization, types of people) together with other similar things, and apart from distinct things, in a way that makes sense to an individual either over time or in that moment (e.g., Bruner, 1957; Tajfel, 1970). Thus, groups are defined both in relation to one another and in the context of one’s current situation. For example, depending on whether one is abroad, discussing the weather, or attending a baseball game, Americans versus Italians, Northerners versus Southerners, or New Yorkers versus Chicagoans might be deemed the appropriate categories. Given the complexity of our social environments, social categorization provides invaluable functions, such as organizing
one’s social environment, allowing one to engage in sense making, and preparing one to interact efficiently and cost-effectively with his or her environment (e.g., Bodenhausen, 1988; Tajfel, 1970). Social categorization also, however, lays the groundwork for prejudice.

As noted, social categorization is the momentary or stable grouping of subjectively similar people into groups, which are also differentiated from each other. Importantly, during social categorization, people also categorize the self (Hogg & Terry, 2000; Oakes, Haslam, & Turner, 1994)—that is, individuals orient themselves in their mapping of similarity and differentiation, thus creating a personal social identity and engendering a “we” (or “us”) and “they” (or “them”) distinction between relevant groups. For example, we are Americans versus they are Italians; we are Northerners versus they are Southerners, and so on. Note that these self and other categorizations denote a shift from the interpersonal to the intergroup. Wilder (1986) suggested that this mere categorization of the self and others into groups leads to a different set of outcomes than one might expect if individuals were not considered in relation to one another (also see Tajfel, Billig, Bundy, & Flament, 1971). Although this claim has in many ways stood the test of time, experiments conducted by these researchers and others have identified other important psychological factors related to the basic process of social categorization. First, a “we” versus “they” distinction is largely innocuous without a crucial finding in evolutionary and social psychology: namely, individuals tend to engage in social comparisons between self (in-group) and other (out-group) automatically—that is, without the effort or intention of the individual (Barkow, Cosmides, & Tooby, 1992; Tajfel, Billig, & Bundy, 1971; Tajfel & Turner, 1979). Second, these two processes of
social categorization and social comparison produce intergroup bias is due to individuals’ preferences for people who are perceived to be similar to them—a preference found to be largely based in the fundamental motivational need for self-esteem or self-enhancement, and specifically in this case, a positive social identity (Barkow, Cosmides, & Tooby, 1992; Tajfel & Turner, 1979). A positive social identity refers to the subjective favorability of one’s in-group when compared to one’s out-group, and thus when achieved through social categorization processes, fulfills one’s basic need to feel good about oneself, given that the group offers both a general sense of belongingness and, importantly, a sense of belongingness to a group that is both distinctive and comparatively better than other groups.

Abrams, Hogg, and colleagues (e.g., 1988; 1990; 1999) have found that the preference for one’s own in-group is also motivated by the need to resolve self-concept uncertainty (e.g., uncertainty about how to act, what to believe, and how this relates to others’ behavior and beliefs). These researchers suggest that individuals resolve such uncertainty by subscribing to responses, beliefs, and so on that are subjectively prototypical of their group in a given situation (also see Schimel et al., 1999, and Greenberg et al., 1990). There are several factors influencing the extent to which uncertainty reduction is a motive in self-categorization. Of course, uncertainty reduction through group identification is most useful when one is experiencing a time of high uncertainty and/or is high in dispositional uncertainty and is most successful when one is highly identified with the group (e.g., Hogg, 2000; Mullin & Hogg, 1998). Thus, individuals are, for example, more likely to show evidence of prejudice when they face uncertainty about an important issue or when their group membership is otherwise not
salient, thus conceivably inducing feelings of uncertainty (e.g., Grieve & Hogg, 1999; Hogg, 2000). Moreover, features of the group influence the extent to which uncertainty reduction is a motivational pull. For example, individuals are more likely to reduce uncertainty by identifying with groups that are high, versus low, in group entitativity—that is, the extent to which a group has clear boundaries, is homogenous, and generally holds together cohesively as a group. Finally, individuals low in status are more likely to seek uncertainty reduction through group identification than those with high status (Reid & Hogg, 2005).

Brewer (1991) offered a third motivational theory for intergroup bias: individuals contend with two opposing motivations, assimilation with an in-group and differentiation. Optimal distinctiveness is achieved when the needs are both perfectly satisfied. One way in which this might occur is when a group is large enough to promote belongingness but small enough to allow for distinctiveness. Intergroup bias, then, tends to occur when these competing needs are not met at optimal levels. As with the other motives described above, the extent to which optimal distinctiveness needs play a role in bias depends upon other factors, such as the status of the in-group and the individual’s level of in-group identification (e.g., Leonardelli & Brewer, 2001; Leonardelli, Pickett, & Brewer, 2010). Recent work, for example, suggests that this motive might account for in-group bias evidenced by minority groups more than majority groups, except when group identification is high (e.g., Leonardelli & Brewer, 2001).

Taken together, these theories of intergroup bias suggest that individuals mentally divide others and the self into groups and, on comparison between one’s own group and other groups, show a preference for—that is, bias toward—their in-group and for the
individuals that make up the in-group. These categorizations and resultant bias have been shown to be in the service of self-esteem striving, self-esteem maintenance, uncertainty reduction, assimilation, and differentiation.

Consequences of social categorization processes

The social categorization and social comparison processes described above produce a host of intergroup outcomes, documented by years of research in this area. This chapter will focus primarily on just two of those consequences: stereotyping and prejudice. The relationship between social categorization and stereotyping has been examined in two ways.

The first type of this sort of research addresses the extent to which placing individuals into groups leads to the generation and application of stereotypes about out-groups. For example, early research found that dividing individuals into groups led to depersonalization wherein out-group members are seen as prototypical representatives of their respective groups rather than individuals with unique characteristics (e.g., Tajfel, 1969). Depersonalization is driven by accentuation, or the perception of increased similarity among individuals in any given out-group, as well as differences (exaggerated both in size and consistency) between groups (e.g., Tajfel, 1969). In many of the experiments demonstrating the phenomenon of accentuation, participants are first divided into groups based upon arbitrary shared characteristics, such as preference for paintings (i.e., a minimal groups paradigm). Next, all participants are asked to rate the degree to which out-group members share beliefs and attitudes on a variety of topics. Participants tend to believe that members of an out-group with a known similarity (e.g., preference for paintings) share opinions and attitudes with one another on both related (e.g., attitudes
toward art) and unrelated (e.g., politics) topics, but not with members of other out-groups or the in-group (e.g., Allen & Wilder, 1979). These sorts of effects are pervasive in this literature (e.g., Hamilton, Sherman, & Rodgers, 2004; Hogg & Abrams, 1988; Jetten, Hogg, & Mullin, 2000). These similarity and dissimilarity accentuation effects occur primarily in characterizations of out-groups (e.g., Park & Rothbart, 1982). An exception is minority or low-status group members, who tend to view their own in-group as largely homogeneous, perhaps due to the strategic value of collective similarity in interacting with majority members (Brown & Smith, 1989; Simon & Brown, 1987).

The second manner of examining how social categorization influences stereotyping is to measure the extent to which existing cultural, or collective, stereotypes are endorsed or applied to out-group members when one is reminded of his or her in-group identity. Here, too, there is evidence for accentuation and depersonalization. For example, when White men are reminded of their own in-group, out-group members (e.g., women, African-Americans) are perceived as more similar to the prototypical out-group member than they are otherwise (Hogg & Turner, 1987). There is similarly extensive evidence that individuals are more likely to recall, endorse, and apply negative stereotypes to out-group members when in-group and out-group social identities are salient (e.g., Oakes, 1987; Oakes, Turner, & Haslam, 1991). For example, one study found that simply making normative groups salient (e.g., medical students with anti-alternative medicine vs. pro-alternative medicine attitudes) led to great activation of concepts associated with intergroup categorizations, such as “subject of study” (Blanz & Aufderheide, 1999). Moreover, Oakes and colleagues (1991) found that not only were group concepts activated, but people behaved in a way that suggested that the concept of
group (rather than person) was activated. In their study, individuals who were presented with group information about normative academic groups (i.e., science students who endorsed hard work vs. a social life) were more likely to categorize people based on gender than were those who were presented with information about nonnormative groups. Finally, if in-group and out-group identities are apparent, in-group members will interpret out-group members’ behaviors (e.g., kicking a tree) in line with dispositional attributes (e.g., X is violent) that are negative and/or known stereotypes associated with the group (e.g., Maass, Silvi, Arcuri, & Semin, 1989; Pettigrew, 1979; Taylor & Jaggi, 1974; also see Hewstone, Jaspars, & Lalljee, 1982).

The findings for the impact of social categorization on prejudice mirror those for stereotyping. Here again, minimal group paradigms have been used to demonstrate the impact of mere categorization on intergroup bias (Brewer, 1979; Hamilton & Trolier, 1986; Messick & Mackie, 1989). The primary finding is that when categorized into groups, individuals report more positivity toward their in-group (e.g., Otten & Moskowitz, 2000; Tajfel, 1969). Individuals also favor in-group members over out-group members on a variety of other measures of bias, such as resource or reward allocations, intergroup trust, prosocial helping, and evaluations of performance (e.g., Dovidio, Gaertner, Kawakami, & Hodson, 2002; Dovidio, Gaertner, Validzic, Matoka, Johnson, & Frazier, 1997; Mullen, Brown, & Smith, 1992; Tajfel, Billig, Bundy, & Flament, 1971). In-group members cooperate more with one another and work hardest when paired with other in-group members (Kramer & Brewer, 1984; Worchel, Rothgerber, Day, Hart, & Butemeyer, 1998).
Perhaps surprisingly, Brewer and colleagues (1981, 1999) have found that often the sort of outcomes described above do not reflect hostility toward the out-group. Instead, these researchers find that manifestations of intergroup bias commonly result from feelings of positivity toward the in-group (i.e., *absolute* liking), and comparatively less positivity (although not necessarily negativity) toward the relevant out-groups (i.e., *relative* disliking). In a practical sense, this means that although in-group bias is relatively common, hostility toward out-groups, especially when avoidable, is not. This has been illustrated empirically, for example, in Tajfel’s rewards allocation studies, in which participants are asked to divvy up rewards between an in-group member and an out-group member. In the rewards allocation task, people tend to choose a strategy that favors their in-group member, regardless of whether this bias is at the expense of the out-group member or not (for a full description of the task and results, see Tajfel et al., 1971; Bigler, Jones, & Lobliner, 1997; Brewer, 1979). If one were truly concerned with penalizing the out-group, then the opposite strategy would be best—that is, to allocate the least amount of rewards to the out-group, regardless of the outcome for the in-group (i.e., even if it meant foregoing a net profit to the in-group). Moreover, the fact that researchers do not find evidence for either in-group favoritism or out-group negativity when one is asked to divvy up negative outcomes rather than rewards in this sort of task further suggests that intergroup bias is often the result of favoring the in-group rather than a desire to disparage the out-group (Mummendey, 1995).

Unfortunately, negativity or prejudice against out-groups that cannot be explained by a preference for one’s own in-group does sometimes emerge, in certain contexts and for certain people. For example, some research suggests that individuals who are strongly
identified with their in-group are more likely to exhibit bias against out-groups (e.g., Struch & Schwartz, 1989). Likewise, when one’s in-group status is threatened, bias is evidenced on both the rewards and the punishment allocation tasks described above (Otten, Mummendey, & Blanz, 1996). In general, a threat to one’s in-group’s beliefs or values, fear of harm to an in-group member, recognizing that there is a lack of available resources, or otherwise engaging in a conflict-laden or competitive situation, increases the likelihood that out-group derogation and bias will be evidenced (Branscombe, Ellemers, Spears, & Doosje, 1999; Castano, Yzerbyt, Paladino, & Sacchi, 2002; Haslam, Turner, Oakes, McGarty, & Hayes, 1992; Sherif, 1967; Tajfel & Turner, 1979).

Alternatively, reducing or eliminating the competitive or threatening nature of a situation reduces out-group hostility (e.g., Mummendey & Schreiber, 1984; Rabbie & Horowitz, 1988).

Early work on social identity responded to the focus on the individual in bias research—work that explicitly or implicitly relied on the assumption that prejudice is driven by one’s motivation or affect in a given situation. At least partially in response to this work, Tajfel (1969), in his now-renowned précis on prejudice, *Cognitive Aspects of Prejudice*, underscored the banality of intergroup bias by highlighting its cognitive basis. In essence, Tajfel, and later others, suggested that across situations and people, and regardless of one’s current motivation, social categorization processes will occur when there are multiple social groups. Was he correct, and if so, does this mean that intergroup bias is unavoidable? As noted above, the research conducted by Mummendey and colleagues (1995; 1999), for example, suggests that at least derogation of the out-group and extreme, blatant prejudice depend on the context and the person and thus are not
inevitable. But whether social categorization processes themselves are automatic, even innate—and what the answer to this question means for the inevitability of bias that is the byproduct of in-group preference—are also important questions, to which this chapter turns next.

**Is intergroup bias inevitable?**

Whether social categorization and social comparison processes are evolved, innate cognitive functions, or instead socially learned is not entirely clear; however, some have argued that evidence for the former has accumulated. For example, children as young as seven months old can distinguish between gender, and preschool-aged children distinguish among ethnicities and age groups (e.g., see Aboud, 1988; Baron & Banaji, 2006; Dunham, Baron, & Banaji, 2006, 2007; Miller, 1983). Also, research has found that a preference for similar others (i.e., one’s subjective in-group) develops concomitantly with social categorization, and thus gender preferences develop earliest, followed by ethnic and age-group preferences (e.g., Aboud, 1988; Katz & Zalk, 1974).

The intergroup bias research conducted using infants as subjects is also informative for clarifying the processes that lead to in-group bias. Namely, this research suggests that while familiarity breeds liking, in-group preferences are not always explained by greater exposure to, or familiarity with, one group (Zajonc, 1968; also see Dasgupta, McGhee, Greenwald, & Banaji, 2000). That is, children under the age of one year develop in-group gender preferences, even though this cannot be due solely to familiarity with one or the other group. This contention has also been supported by research using minimal groups paradigms, wherein groups are formed by an experimenter based upon an arbitrary and/or randomly assigned distinction between groups, such as the
color of one’s shirt (e.g., Brewer & Silver, 1978; Tajfel, 1970). Evidence is increasing that suggests that social categorization processes appear to occur spontaneously and automatically, and that in-group preferences follow directly from these processes, and thus develop from an early age (see also Hardin & Banaji, in press).

These findings suggest that categorization processes occur quite easily, and there is plenty of evidence of implicit bias toward out-groups that would follow from such categorization processes, which we will review in the next section. Is there any way to combat these automatic categorization processes? A recent, intriguing line of work suggests that it is possible to undermine the prejudicial responses that follow from such early group-based processes. In a series of experiments, Moskowitz and colleagues (2000, 2005, 2011) found that for those for whom being egalitarian is a chronic goal, intergroup salience actually led to decreased implicit stereotyping of the out-group. In other words, categorization processes occurred, but an intergroup egalitarian response was enacted automatically rather than an intergroup bias one. Indeed, this appears to be an important lesson for bias reduction, more generally: if one wants to reduce intergroup bias, exploring and exploiting categorization processes (not stopping them) is one viable route. The ways in which this has been done, and potentially could be done, will be discussed in the context of organizations and workgroups next.

**Intergroup bias in organizations**

Intergroup bias has been measured both implicitly and explicitly. The meaning of implicit (i.e., automatic) and explicit (i.e., controlled) attitude measurement, and the various theoretical matters concerning this dichotomy, have been given much attention in recent years (e.g., Bargh, 1994; Devine, 1989; Fazio, Jackson, Dunton, & Williams, 1995;
Greenwald, McGhee, & Schwartz, 1998; Jacoby & Witherspoon, 1982; Langer, 1975; McConnell & Leibold, 2001; Meyer & Schvaneveldt, 1971; Neely, 1977). The terms *implicit* (automatic) attitudes and *explicit* (controlled) attitudes have variously been used to refer to types of attitudes, the types of processes that produce attitudes and behaviors, and the ways in which attitudes are measured (e.g., De Houwer et al., 2009; Fazio et al., 1995; Ferguson & Fukukura, 2012; Greenwald & Banaji, 1995; Jacoby & Witherspoon, 1982; Petty, Fazio, & Brinol, 2007; Schacter, 1987; Squire, 1986; Wittenbrink & Schwarz, 2007). Keeping with recent literature, we use the terms *implicit* and *explicit* to refer to the processes that produce attitudes and behaviors. Explicit attitudes, then, are those produced by controlled, intentional regulation processes. Thus, individuals are aware of, and are able to self-report, these attitudes. Explicit attitudes tend to correlate with many overt behaviors, such as name calling and the use of slurs. Implicit attitudes tend to correlate with nonverbal behaviors, such as smiling, and the distance that one sits or stands from an out-group member. These attitudes are produced automatically—that is, without conscious intention, and sometimes without effort or awareness. As such, implicit attitudes are not directly self-reported but rather are measured with various tools, typically, but not always on the computer (for reviews see Fazio, 1995; Fazio & Olson, 2003; Greenwald & Banaji, 1995; Petty et al., 2007; Wittenbrink & Schwarz 2007). In many cases, these tasks assess attitudes toward an attitude object (e.g., studying, African-Americans, one’s mother, animals, coworkers) by measuring how quickly people are able to categorize a positive or negative word (as positive or negative, respectively) after first seeing an attitude object. The more quickly one categorizes a positive word, and the more slowly one categorizes a negative word, after first seeing the word “mom,” for example,
is an indication of how positively one feels toward one’s mother (for a description of other tasks, see Wittenbrink & Schwarz, 2007).

The Implicit Association Test, perhaps the most widely used tool for measuring implicit attitudes, assesses attitudes using a similar principle, but the task involves a pairing of attitude objects with attributes. For example, the categories of “women” and “men” (the attitude objects) are paired with the categories of “positive” and “negative” (the attributes), respectively, and then people are asked to quickly classify words (e.g., “Betty,” “sunshine”) associated with each attitude object and attribute. To the extent that it is easier to complete the categorization task when women and positive, and men and negative, are paired, versus the reverse pairing, is taken as evidence that one has a positive implicit bias toward women, and vice versa (e.g., see Greenwald et al., 2009; Greenwald, McGhee, & Schwarz, 1998).

Note that these categories of implicit and explicit are distinct from the in-group preference versus out-group hostility bias distinction discussed earlier—that is, implicit and explicit can reflect either in-group preference or out-group hostility. Nevertheless, one might expect that negative out-group attitudes are more likely to be spontaneously, explicitly expressed under the conditions that produce out-group hostility (i.e., threat or competition).

An added complexity to the implicit versus explicit attitudes is that there can be within-person dissociation in the two types of attitudes. Modern theories of prejudice, in fact, are largely based upon the phenomenon that people espouse egalitarianism or positive out-group attitudes explicitly, but display prejudice in their implicit evaluations, nonverbal behavior, and so on. Further, negative out-group behaviors seem to “leak out”
when individuals can attribute their behavior to a nonprejudicial source, or otherwise rationalize or justify their behavior. For example, individuals may be equally likely to help in-group and out-group members during an emergency, but more likely to help an in-group than an out-group member when there is ambiguity in the degree to which the situation is a true emergency. Gaertner and Dovidio (1986) have termed this *ambivalence aversive racism* and note that it stands in contrast to the more overt, old-fashioned racism that was far more common in the United States prior to the 1970s or 1980s (see Duckitt, 1992). Other theories of contemporary prejudice have similarly suggested that under most conditions, prejudicial attitudes since 1980 tend to be expressed in unintentional, subtle, and indirect manners, rather than in explicit statements or overt behaviors (e.g., Crandall & Eshleman, 2003; Dovidio & Fazio, 1992; Dovidio & Gaertner, 2004; Greenwald & Banaji, 1995; Hardin & Banaji, in press; McConahay, 1983).

That attitudes have increasingly become more explicitly positive toward out-groups is in large part due to changes in official and unofficial legal and social policies, which in turn have shaped the norms for the intentional (explicit) expression of prejudice. Although a reduction in the explicit expression of prejudicial attitudes and behavior marks progress in efforts to reduce prejudice, as alluded to above, it is still the case that ethnic and gender out-groups face discrimination and prejudice at the individual, organizational, institutional, and societal level (e.g., see Bobo, 2001; Dovidio, 2001; Sniderman & Carmines, 1997). Moreover, it also has not been met with a concomitant decrease in implicit intergroup bias (see Fazio & Olson, 2003; Hardin & Banaji, in press). Plenty of recent studies have demonstrated implicit prejudice toward groups based on ethnicity and race (e.g., see Dovidio & Gaertner, 2004; Fazio et al., 1995; Nosek, Banaji,
& Greenwald, 2002), gender (e.g., Banaji & Hardin, 1996), sexual orientation (e.g., Dasgupta & Rivera, 2009), and according to many other stigmatized characteristics (for a review see Olson & Fazio, 2004; Greenwald et al., 2009).

The importance of reducing implicit bias in the workplace cannot be overstated. Implicit intergroup bias has far-reaching negative effects in many organizational domains, including, but not limited to, selection, retention (including compensation and promotion issues), teams-related issues, general work environment, and worker self-esteem and well-being (Beckman & Phillips, 2005; Forman, 2003; Zatzick, Elvira, & Cohen, 2003. For example, processes important for working in teams, such as information exchange, suffer in diverse groups that are divided along racial lines (e.g., Lau & Murnighan, 2005; Li & Hambrick, 2005; Sawyer, Houlette, & Yealey, 2006). In other words, fostering harmonious intergroup interactions is at the crux of producing the best possible outcomes in organizational productivity, organizational climate, and social justice.

The following will focus primarily on groups with perceptually detectable differences such as ethnicity and gender rather than “invisible” differences (e.g., in political opinion or skills). Although much of the work discussed next will apply across types of diversity, the visibility of perceptually detectable differences, juxtaposed against the historical backdrop of tense relations in the U.S. workforce for these groups, provides a context that is perhaps uniquely suited for considering the robust effects of social categorization. Even relatively recent research finds evidence of interethnic and gender-based tension in the workplace. For example, compared with their White counterparts, African-Americans report feeling undervalued, as well as feeling that the workplace is
less positive, with less access to opportunities (e.g., Reskin, 1998; for a review, see Smith, Brief, & Colella, 2010). Likewise, White Americans report lower job satisfaction and sociability in diverse work environments (Mannix & Neale, 2005; Riordan & Shore, 1997).

Fortunately, much progress has been made in the area of intergroup relations, and specifically on how modulating or exploiting the process of social categorization can reduce bias. From a social identity or self-categorization theory perspective, there are three potential ways in which intergroup bias has been reduced and cooperation has been increased: eliminating or deemphasizing distinctions between social categories (decategorization), creating or emphasizing an already existing higher-order category (e.g., factory worker or team member identity; recategorization), or retaining the social categories but revaluing the boundaries between the groups (mutual differentiation). As many scholars have noted, intergroup bias is multiply determined by cognitive, affective, and motivational causes. As such, it should be acknowledged that there are many other means to stereotyping and prejudice reduction; however, interventions aimed at social categorization processes and intergroup boundaries represent viable, often straightforward, and practical responses to a complex problem.

**Decategorization**

As prejudice reduction research began to take hold in the 1950s and 1960s, decategorization was quickly recognized as a method for achieving the goal of that research. Recall that social categorization processes engender a move from the individual to the group level of thinking (i.e., one categorizes self and other into groups). Although thinking at the group level itself is not necessarily problematic, one issue that arises is
that group-level negativity towards out-groups can be, and often is, applied to
individuals. Decategorization is anything that creates circumstances in which
interpersonal, rather than intergroup, thinking is encouraged (Brewer & Miller, 1984;
1988). Decategorization has two possible benefits. First, otherwise negative interactions
with an out-group member may be more positive if one is considering one’s contact with
an individual rather than one’s cognitions or affect toward the group more generally.
Second, just as stereotypes about, and affect toward, an out-group can influence
experiences with an individual out-group member, experiences with an out-group
member can generalize to group-level affect and beliefs. Of course, these positive
outcomes are predicated on positive interpersonal experiences with out-group members.

Perhaps the best-known, and most elaborate, method of decategorization was
proposed by Gordon Allport (1954) in his Contact Hypothesis, which details the optimal
conditions needed to encourage positive interpersonal contact between members or out-
groups. Allport, and later others, began to unpack the mechanisms through which the
military might have provided a special situation in which intergroup conflict is reduced
(see also Dovidio, Kawakami, & Gaertner, 2002; Hamberger & Hewstone, 1997;
Pettigrew, 1997; Pettigrew & Troop, 2006; Richeson & Shelton, 2003; Rothbart & John,
1985; Shelton & Richeson, 2005). In short, the Contact Hypothesis suggests that
prejudice can be reduced by repersonalizing individual out-group members in light of a
shared, positive experience in which information about differentiated members is
attended to (especially expectancy-inconsistent information) and subsequently used as a
basis for guiding judgments about those individuals (e.g., Blair, 2002; Brewer & Miller,
1984; Erber & Fiske, 1984; Miller, Brewer, & Edwards, 1985). For attitude change to be
generalized from attitudes toward the individual to attitudes toward the group, a higher threshold must be met. Rothbart and John (1995) suggested that the behavior of the out-group member must be inconsistent with a known stereotype, but that the out-group member should otherwise seem to be a typical member of the out-group rather than an exception. Contact must also happen frequently and across a variety of contexts.

Experimentally, a host of studies support the effectiveness of the Contact Hypothesis for changing the attitudes of both majority- and minority-group members (for a review, see Pettigrew & Tropp, 2006). The controlled circumstances under which the Contact Hypothesis creates attitude change, coupled with the constraints of the workplace, make the Contact Hypothesis somewhat challenging to implement in organizations; however, it is more promising than it might seem at first blush. For example, as mentioned, the military has institutionalized policies that by design encourage situations in which bias should be reduced. Moreover, it is not necessary to meet all of the optimal criteria outlined by the hypothesis in order to prompt attitude change. It a meta-analysis of studies, Pettigrew and Tropp (2006) found that although optimal conditions led to the greatest attitude change toward the individuals in the interaction, merely increasing contact between members of out-groups sometimes decreased prejudice. Finally, even indirect contact seems to confer benefits for intergroup relations. Learning that a member of one’s in-group is friends with out-group members, observing positive cross-group interactions, and even imagining contact with out-group members has been shown to reduce prejudicial beliefs (Dovidio, Eller, & Hewstone, 2011; Mazziotta, Mummendey, & Wright, 2011; Wright, Aron, McLaughlin-Volpe, & Ropp, 1997). If mere contact with out-groups (or its observation) seems to improve
intergroup relations, why do negative intergroup attitudes in organizations still persist? There are many answers to this question. One is that organizations can be contexts rife with conflict and competitiveness, both of which are inversely related to intergroup communication, as well as risk factors for exaggerated social categorization effects (e.g., Nelson, 1989; Richter, West, & van Dick, & Dawson, 2006). In other words, sometimes even in situations in which diversity exists, individuals can observe and experience a homogenous situation, and this is why the opportunity for intergroup contact (as opposed to actual intergroup contact) has a much smaller relationship with intergroup attitudes (e.g., Pettigrew, 2008. The critical element, then, seems to be encouraging intergroup contact, whether it is first-person or third-person, real or imagined. This intergroup contact also ought to be experientially positive (or neutral) in order to encourage personalization, and in turn for attitudes to improve, because aversive experiences can simply reinforce prejudicial attitudes.

Recategorization

Like decategorization, recategorization is a method of reducing the emphasis on boundaries between groups. Recategorization, however, encourages group-level thinking rather than personalization. In recategorization, old boundaries between groups are deconstructed, and a new, inclusive group is formed or emphasized (Brown & Turner, 1981; Doise, 1978; Sherif et al., 1961). In most cases, individuals are asked to form or attend to a higher-order identity that they share with their out-group, such as a work team identity or a shared organizational identity (e.g., Gaertner, Dovidio, Anastasio, Bachman, & Rust, 1993). A special case of recategorization occurs when individuals are given, or asked to attend to, information about a characteristic or attribute that their in-group and
the out-group share (e.g., Urban & Miller, 1998). In these ways, recategorization exploits social categorization processes. That is, rather than attending to one’s in-group or out-group status, one attends to an inclusive grouping of the in-group and out-group. As one might expect, the typical social categorization processes occur; however, in this case, positivity increases for the more inclusive group and group members (e.g., Mottola et al., 1997). One of the features of Allport’s (1954) Contact Hypothesis was sharing an interdependent goal, and this point highlights the bidirectional relationship between cooperation and liking. That is, liking leads to increased cooperation, but cooperation can also lead to liking; these effects reinforce one another and thus have benefits that last far beyond the length of any intervention (Allport, 1954; Brown & Turner, 1981; Sherif et al., 1961).

Research on recategorization within the organizational field largely finds support for its effectiveness. Some studies, done in both educational and organizational settings, find that creating an inclusive “we” group is sufficient for reducing stereotyping and intergroup bias (Hogg & Terry, 2000). For example, Haslam and Ellemers (2005) summarize evidence suggesting that when individuals from different groups share an in-group identity, their expectations and motivations regarding the interaction differ from a situation in which no such identity is made salient. Notably, when individuals expect to share beliefs, cooperate, and reach an agreement, their behaviors are guided by these expectations. Thus, constructive disagreement, information sharing, use of adaptive persuasion techniques, and amount of communication are all greater in intergroup contexts in which a common in-group identity is made salient (Haslam & Ellemers, 2005). The expectations for interactions with a shared group, in essence, catalyze a self-
fulfilling prophecy. Finally, individuals with a salient common in-group identity self-disclose more, build intimacy and trust, and are engaged in more cross-group helping (e.g., Archer & Berg, 1978; Dovidio et al., 1997).

Although recategorization is causally related to reducing discrimination within the context of the situation in which a common identity is formed, these effects may not generalize to attitudes and beliefs about the out-group. Forming a common group identity reduces bias through the de-emphasis on between-group divisions; thus, individuals may indeed become fond of out-group members with whom they share an in-group identity, but that affective positivity may be detached from the out-group members’ out-group identity. In other words, a man may respect and fully cooperate with a woman who is on the same work team, but the attitudinal shift may be specific to the woman in the work team and may not translate to women with whom he does not share an in-group status. As a related point, these effects may be even further specified by the context. That is, even the positive attitudes toward targets achieved through a shared in-group status could be limited to the time that one shares an in-group status, particularly when the collective in-group salience is itself short-lived (e.g., Brewer, 1991). Finally, efforts at recategorization can sometimes be met with resistance, given that individuals may feel strongly tied to their already established social identity.

One final method for reducing bias at the workplace addresses these latter concerns by leaving intergroup boundaries intact but revaluing and recontextualizing those boundaries.

**Mutual differentiation**
Mutual differentiation differs from the other two types of interventions in that rather than deconstructing groups, boundaries within groups are maintained. The groups are recontextualized, however, to reduce the threat that might otherwise be present. Both self-categorization and social identity theory predict that individuals are threatened by ambiguous or weak boundaries between groups, as well as the fear that these boundaries might be reduced or eliminated. Hewstone and Brown (1986) suggested that one way to encourage intergroup positivity would be to affirm the boundaries between groups, or what they termed mutual differentiation. Similar to the other two types of interventions, the mutual differentiation model retains the use of a cooperative context in order to reduce the threat associated with interacting with out-group members. Because mere intergroup contact is sufficient for making category memberships salient in many cases, this theory lends itself to study in naturalistic settings (e.g., see also Brown & Hewstone, 2005).

The extent to which mutual differentiation reduces intergroup bias depends on the amount of intergroup contact. At high levels of (positive or friendly) contact, if social identities are salient, and if an out-group member is perceived to be typical of other out-group members, the amount of anxiety experienced by members of both groups should be lower, and the amount of empathy and positivity higher, than under any other conditions. Mutual differentiation seems to be especially effective for generalizing attitudes, and indeed is superior to the other two types of intervention under the conditions just outlined. For example, in one study, groups of four individuals, two from one social group and two from another social group, worked on a project. The results suggested that dividing the members’ roles along the lines of their social groups (i.e., both members
from one social group were responsible for one task, and both members from the other
social group were responsible for a different task) created mutual differentiation and
resulted in reduced negativity toward the out-group relative to groups in which roles were
not divided along group lines (e.g., Brown & Wade, 1987; also see Deschamps & Brown,
1983). In addition, several studies have shown that counter-stereotypic, yet typical,
members are essential for the generalization of positive attitudes toward the out-group,
measured by both attitudes and behavior (e.g., Brown, Eller, Leeds, & Stace, 2007;
Hewstone, 1994; Hewstone, Hassebrauk, Wirth, & Waenke, 2000). Like
decategorization, cooperation is an antecedent for mutual differentiation. The major
difference is that with mutual differentiation, cooperation is operationalized in a manner
that highlights the distinctions between groups rather than minimizes them (e.g.,
Kenworthy, Turner, Hewstone, & Voci, 2005).

Capitalizing on diversity: multiculturalism versus colorblindness

There is no single best method for reducing bias in organizations. Moreover, the various
methods for reducing bias are not entirely mutually exclusive. That is, to the extent that
these types of interventions each offer maximum benefits under different conditions, they
can each be used at different times to achieve optimal outcomes. To that end, researchers
have offered integrative models for prejudice reduction (e.g., Brewer & Gaertner, 2001).
Of course, what counts as an optimal outcome might also vary. For example, one might
be concerned with promoting creative performance on a task, reducing majority-group
bias, reducing minority-group bias, improving attitudes toward an individual,
generalizing positive attitudes from a member of a group to the group more generally, or
some combination of these outcomes. Thus, on the one hand, it is important to identify
the relevant conditions and desired outcomes when developing a strategy for reducing bias. On the other hand, there is developing research to suggest that in cases of surface-level, demographic diversity, there is an intergroup ideology that seems to be superior for meeting most intergroup-related goals.

Regardless of the particular prejudice reduction strategy—decategorization, recategorization, or mutual differentiation—one could choose to either ignore between-group differences or recognize and acknowledge them. Multiculturalism is an ideology that promotes the recognition and acceptance of group differences (e.g., Lott, 2009). Color- (or gender-) blindness is an ideology that stresses ignoring group differences, and points to a lack of belief that racism still exists or is influential (Lott, 2009). Although multiculturalism is perhaps more obviously compatible with decategorization or mutual differentiation than with recategorization, the ideology could be employed for any of the strategies. For example, while emphasizing a common in-group identity, one could still recognize, value, and accept that differences exist between the groups. Alternatively, a strategy of colorblindness seems to be compatible only with either decategorization or recategorization.

Research on these two types of ideologies is in some ways still in its infancy, but several important findings have already emerged. First, for majority-group members, belief in a multicultural ideology leads to less bias than belief in a colorblind ideology (Neville, Lilly, Duran, Lee, & Browne, 2000; Richeson & Nussbaum, 2004). The opposite is also true: increasing adherence to colorblindness is associated with increasing levels of intergroup bias (Neville et al., 2000; Richeson & Nussbaum, 2004; Verkuyten, 2005). Second, minority-group members have a greater tendency to endorse a
multicultural ideology, and among minority-group members, greater endorsement of multiculturalism is associated with stronger in-group identification as well as greater positivity toward the in-group (Verkuyten, 2005). Although an individual’s endorsement of one or the other ideology predicts in-group identification and intergroup bias, experimental manipulations meant to encourage a multicultural or colorblind ideology also affect intergroup attitudes, stereotyping, and bias, suggesting that these ideologies are malleable (e.g., Richeson & Nussbaum, 2004). This research has been done in both experimental settings and naturalistic field settings (e.g., Plaut, Thomas, & Goren, 2009; van Knippenberg, Haslama, Alexander, & Platow, 2007).

Third, multiculturalism is more likely to facilitate organizations’ desires to capitalize on diversity. Research on the influence of diversity in organizations on productivity and creativity is mixed, but the bottom line appears to be that well-managed and well-functioning diverse groups and environments can outperform homogenous groups and environments (e.g., Mannix & Neale, 2005; Pelled, Eisenhardt, & Xin, 1999). Multiculturalism influences many outcomes that in turn predict both the functionality of diverse groups and their ultimate productivity and performance. For example, Plaut and colleagues (2009) found that majority-group members’ multicultural attitudes were negatively associated with minority-group members’ perceptions of bias in the workplace and positively associated with minority-group members’ psychological engagement at their organizations. Psychological engagement at the workplace is predictive of productivity, among other positive work-related outcomes. Taken together, these findings suggest that a multicultural ideology is more beneficial than a colorblind ideology for both majority groups and minority groups, and for multiple intergroup and work-related
outcomes. Moreover, the linear relationships suggest that the more strongly one adheres to multiculturalism, the more adaptive the outcomes.

Further research is needed to fully understand the multiple mechanisms through which multiculturalism influences intergroup outcomes, as well as how multicultural ideologies interact with the strategies for prejudice reduction outlined earlier in the chapter to produce intergroup outcomes. So far, however, the existing research seems to underline the themes of this chapter: that social identity and social categorization processes are integral to understanding bias; that prejudice is rooted in basic cognitive processes but is not a foregone conclusion; and that the reduction of intergroup bias in organizations is important for promoting both social justice and work performance outcomes.

**Future directions for research**

We have reviewed a great deal of research that speaks to the ways in which prejudice and stereotyping can emerge, in the workplace or elsewhere. Today, intergroup prejudice and stereotyping are likely to emerge on an implicit rather than an explicit level, and thus may be somewhat more difficult to measure casually, even while their effects are significant, pervasive, and destructive at the individual and societal level. What are some of the directions that research might go in the future to understand how implicit bias exists and can be overturned in the workplace?

One important initial note is the need for continued examination, in the basic sciences, of the ways in which intergroup attitudes and relations unfold. As many scholars have argued (see Hardin & Banaji, in press), the area of implicit social cognition is a relatively new field, and there are still very fundamental questions to be addressed,
including issues surrounding construct and predictive validity of implicit constructs. To be able to apply findings and research from social psychology to the workplace, and beyond, there needs to be continued special attention to the pressing issues and developments surrounding this topic in basic science.

For example, in the realm of predictive validity, there is a solid, and still burgeoning, amount of evidence showing the existence of implicit prejudice among “known groups” in the world (especially as measured by the IAT; see Greenwald et al., 2009). This means that groups that have a history of prejudice or intergroup rivalry or strife show correspondent preferences on the IAT, illustrating its sound construct validity. However, work is ongoing to fully understand the depth and boundaries of the predictive validity of implicit measures of prejudice and stereotyping. Although recent reviews and meta-analyses (Fazio & Olson, 2003; Greenwald et al., 2009) suggest that such implicit constructs can meaningfully predict both subtle as well as overt behavior—and even out-predict more traditional, explicit measures—there still needs to be careful work on when and how such implicit constructs shape and guide behavior.

Early research on implicit prejudice and stereotyping suggested that there is likely no way to avoid implicit bias (e.g., Bargh, 1999; Devine, 1989; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997). However, as we noted earlier, more recent work finds that implicit bias and implicit stereotyping are subject to various situational constraints (e.g., Lowery, Hardin, & Sinclair, 2001; Rudman, Ashmore, & Gary, 2001). Moreover, various researchers have shown that implicit bias can be experimentally changed through “retraining” procedures in which participants learn new automatic (and positively valenced) associations with stigmatized groups (e.g., Kawakami, Dovidio,
Moll, Hermsen, & Russin, 2000; Kawakami, Dovidio, & van Kamp, 2005). This work shows numerous, interesting ways in which implicit bias and stereotyping can be altered, minimized, or altogether erased (at least temporarily), and future work will undoubtedly shed more light on the circumstances in which such modification is possible and durable.

One particularly interesting line of work comes from the laboratory of Moskowitz and colleagues and concerns the display of implicit egalitarianism (Moskowitz, 2000, 2005, 2011). In this research, participants who have shown a strong and chronic goal of avoiding prejudice and striving toward egalitarianism show no evidence of implicit stereotyping. Importantly, this work shows that such participants possess the knowledge about the specific stereotypes—it is not as though such participants simply somehow do not have the same knowledge as almost everyone else; it is just that their experience and motivation have allowed them to develop new kinds of implicit associations with stigmatized groups. It appears that they in fact show automatic inhibition of stereotypes associated with a group. This work essentially highlights a group of people according to a personality difference (i.e., those with a strong egalitarian motive) who have been able to “retrain” themselves. It is a nice parallel to the research showing that participants without such a motive can be retrained in the laboratory. Future research could continue to examine the ways in which implicit egalitarianism develops, generalizes, and shapes intergroup decisions and behavior.

A final note about future research concerns the increased interest in thinking and speculating about how implicit bias has consequences for the economic, labor, legal, and public policy realms (e.g., Ayres, 2001, Banaji & Bhaskar, 2000; Banaji & Dasgupta, 1998; Chugh, 2004; Greenwald & Krieger, 2006; Jost, Rudman, Bair, Carney, Dasgupta,
Glaser, & Hardin, 2009; Kang & Banaji, 2006; Tetlock & Mitchell, 2009). Surely the growing and solid evidence for the existence of implicit prejudice means that thinking about its real-world applications is imperative. This line of theory and empirical research will have implications for the workplace, both in a direct manner as well as in terms of policies that influence the workplace. We strongly argue here for the importance of a constant dialogue between basic scientific research findings on implicit bias and their possible applications and implications for the world beyond academia.

References


