

Identity, Belief, and Bias*

Geoffrey L. Cohen

Stanford University

*Chapter to appear in J. Hanson (Ed.), *Ideology, Psychology, and Law*

Abstract

This chapter presents research on the social psychology of identity and belief. Beliefs tied to long-held identities resist change and bias the processing of new information. These phenomena help explain defensive denigration of victims of social problems, resistance to persuasion and intransigence in negotiation, and discrimination in hiring decisions. In each case, intervention strategies based on a social-psychological analysis provide a potential remedy.

Identity, Belief, and Bias

People often persist in long-held beliefs even in the face of evidence that invalidates them. In a classic study, opponents and proponents of capital punishment reviewed the same mixed scientific evidence concerning the ability of the death penalty to deter would-be murderers. Each side saw that evidence as, on the whole, confirming their prior beliefs (Lord, Ross, & Lepper, 1979). They tended to accept the research that supported their prior beliefs and to denigrate the research that contradicted those beliefs. As a consequence, they reported that the evidence made them even more extreme in their beliefs. The tendency to evaluate new information through the prism of preexisting beliefs, known as assimilation bias, is robust and pervasive (Kahan, 2010; Pronin, Gilovich, & Ross, 2004; Tetlock, 2005; cf. Gerber & Green, 1999).

Although people acknowledge assimilation bias in others, they often fail to see it in themselves (Pronin, Linn, & Ross, 2002; Pronin, this volume). In one dramatic study, even so-called political experts displayed severe bias and had no awareness of it (Tetlock, 2005). The subjects, scholars in political science, persisted dogmatically in their economic and political theories, even when confronted with unambiguous evidence that those theories had led them to make erroneous predictions about world affairs. Moreover, knowledge and expertise does not protect against bias and error. Knowledge can simply provide more “informational grist” for the “mill” of people’s prior beliefs (Kruglanski, Webster, & Klein, 1993; see also Vallone, Ross, & Lepper, 1985). For example, among political experts, expertise in a given domain of world affairs predicted greater bias and worse accuracy when it was accompanied by ideological extremity and a thinking style that favored simplification.

People's beliefs and attitudes affect social policy, the law, and government decisions. They affect who will lead the country, whether we go to war, and what economic policies will be implemented. It thus is critical for people to remain both open to strong evidence and skeptical of weak evidence. It can be problematic, as is often the case, when they do not. This chapter reviews research that my collaborators and I have conducted to better understand how people's beliefs, ideologies, and identities impede openness to new information and empathy to victims of social problems. Our research goes beyond the documentation of psychological biases to advance theory-driven intervention strategies. We review strategies, borne out of relevant theory and research findings, that encourage openness to probative evidence, willingness to compromise across partisan divides, and empathy to the struggles of others.

Three areas of research are summarized. The first examines the role of ideologies about race in shaping punitive reactions to minority victims. It focuses on how a colorblind ideology—a common ideology in the U.S. that downplays race-based differences in experiences—can blind people to racial inequities. However, the story will prove complex. While a colorblind ideology can impede empathy, so can a multicultural ideology that values ethnic diversity and downplays people's shared humanity. Racial ideologies, our findings suggest, seldom have simple effects but instead interact with situational factors to affect empathy. Additionally, both colorblind and multicultural ideologies contain assets that, when combined into a hybrid, may promote the most constructive responses to race-based problems in America.

The second area of research focuses on the role of identity in resistance to persuasion and inter-group conflict. This chapter presents evidence that beliefs, when tied to political identities such as “liberal” or “conservative,” are abandoned only with great reluctance. This can make adversaries unwilling to compromise, even when they would benefit materially by doing so. The

chapter also reviews research suggesting that a particular kind of psychological intervention—one in which people affirm *alternative* sources of identity and self-worth—can help overcome closed-mindedness, mistrust, and intransigence in inter-group conflict.

The third area of research examines a subtle psychological bias that may lie at the root of much workplace discrimination. It arises from a tension between people's desire to maintain a personal identity as "fair" and "objective" and their sometimes unconscious prejudices against groups. People resolve that tension by creating *constructed criteria* of merit. This permits workplace discrimination to persist seemingly without challenge to egalitarian principles. Our research, however, offers a promising intervention strategy for reducing discrimination.

In each area of research we ask a question too seldom raised in the social sciences: Given that a problem exists, how might it be overcome? Our approach is informed by Kurt Lewin, one of the founders of experimental social psychology. As Lewin suggested, if one's theory of a social problem is accurate and powerful, one should be able to change it (Lewin, 1952).

Ideologies about race and punitive reactions to minority victims: A blindness in colorblindness?

In a thought-provoking essay, the social psychologist Gustav Ichheiser (1970) suggested that the denial of intergroup difference gave rise to prejudice:

[W]e fail to understand that people whose personalities are shaped by another culture are psychologically different—that they see the (social) world in a different way and react to it as they see it. Instead we tend to resolve our perplexity arising out of the experience that other people see the world differently than we see it ourselves by declaring that those others, in consequence of some basic intellectual and moral defect, are unable to see things "as they really are" and to react to them "in a normal way" The prejudiced are not those who

insist that people are different, in various respects and by various reasons, but those who deny it” (pp. 70-72)

If we start with the “false assumption” that people see the world similarly, and then find that others in fact differ or disagree with us, we will tend to “denounce and persecute” them (Ichheister, 1970, pp. 72-73; see also Ross & Ward, 1995). Drawing on Ichheiser, we thought that a form of this false assumption is particularly pernicious in the U.S.—the assumption that different ethnic groups experience similar objective circumstances and hold similar perceptions of those circumstances. This is referred to as a colorblind ideology. People who endorse it, we expected, would tend to attribute race-based disagreement and difference to the defects of the other. Indeed, people sometimes make sense of disagreements with outgroup members by calling to mind their inherent “otherness” or by marshalling stereotypes that cast them as inherently inferior (Kunda, Davies, Adams, & Spencer, 2002; Miller & Prentice, 1999). The colorblind individual lacks a situational explanation for difference and thus resorts to a dispositional one (see also Hanson, this volume).

In a series of studies, we investigated whether a colorblind ideology impedes interracial empathy. We identified White Americans who espoused colorblindness or did not, and then had them respond to one of two minority victims (Cohen, Scrivener, & Miles, 2009). Both victims experienced the same struggles, but whether those struggles were tied to race or not was experimentally varied. We hypothesized that, when the victim’s struggles were tied to race, people high in colorblindness would defensively resist the information and ascribe the victim’s problems to internal defects.

In studies with sample sizes ranging from 25 to 70, we first measured college students’ endorsement of a colorblind ideology using a survey that included items such as “Members of

different racial and ethnic groups see the world similarly.”¹ Several weeks later we presented those respondents to a minority victim whose problems were described as relevant to race or not. For instance, in one study participants reviewed a case report of a minority woman living in the city and applying for welfare. Her problems, such as lack of education and limited job opportunities, did not vary for participants. However, for a random half of participants, the woman’s problems were explained, by her, as partially related to her race, while for the rest of the participants they were not. For example, in the race-relevant condition, the woman emphasized the “stress from coping with the problems many Black Americans have to face in the U.S.” By contrast, in the race-irrelevant condition, these statements were altered slightly so that instead the woman emphasized the “stress from coping with the pressures of being a single mother in the U.S.” The manipulation was relatively subtle, contained within a couple of sentences in a 3-page report. We assessed how much participants blamed the victim—that is how much they ascribed her problems to personal defects and recommended punitive measures such as denying her benefits.

As expected, ideology shaped empathy. In the race-irrelevant condition, participants high colorblindness were, if anything, *less* victim-blaming and *more* empathetic than participants low in colorblindness (i.e., those who held more multicultural beliefs). When a victim does not contradict their worldview, colorblind individuals may be relatively better able to see common humanity with victims. However, in the race-relevant condition, the pattern reversed. High colorblind subjects’ victim-blaming rose sharply. They proved relatively more likely to judge the victim harshly and recommend that she be forced into a mandatory work program.

¹ All effects reported in this chapter are statistically significant, below the conventional $P = 0.05$ threshold.

One unexpected finding was that individuals low in colorblindness—that is, those who thought that “different races see the world differently”—also based their empathy on how much the victim’s problems fit their worldview. They engaged in marginally less victim-blaming when the woman framed her problems as race-relevant rather than race irrelevant. In this way, both groups proved responsive to whether the victim’s problems were congenial to their worldview. The same pattern replicated in the context of a problem less racially charged than poverty. When presented with a fellow college freshman—an ethnic minority who was having academic troubles—colorblind individuals were less sympathetic and more punitive when that student ascribed some of his difficulties to the struggles of being a minority on campus than when he ascribed them to the struggles generally faced by new students making the transition to college. A final study demonstrated that colorblindness is not colorblind. Rather it is selectively applied to minority outgroups. It was only when the victim was Black, rather than White, that colorblindness mattered. Responses to the White victim were generally positive regardless of whether or not she ascribed her problems to race.

A key implication is that a relatively small manipulation had a large effect on victim-blaming and empathy, in some cases even reversing reactions. Ideology, it seems, can act like a double-edged sword: Colorblindness was associated with greater empathy to victims when this ideology was not under threat, but less empathy when it was.

These processes can lead to ripple effects. People’s judgments of a single victim, biased by their prior beliefs, can cascade to affect their attitudes and policy recommendations with regard to a social problem in general. Later, after reviewing the welfare case report described previously, participants indicated their attitudes on various social issues in the context of a seemingly unrelated study. One series of questions asked participants for their views of poverty, in particular

the extent to which they blamed poverty on the deficits of the poor rather than environmental circumstances. The responses mirrored subjects' previous judgments of the welfare applicants. If they had read about a welfare applicant who ascribed some of her problems to race, colorblind subjects were relatively more likely to ascribe poverty in *general* more to the personal deficits of the poor rather than to their circumstances. They were also relatively more opposed to social policies aimed to redress poverty. However, in the race-irrelevant condition, this pattern reversed: Colorblind participants were relatively more favorable and less victim-blaming in their views of the poor.

This research suggests that responses to victims and social problems are driven, in part, by ideology. It suggests that people are to some extent top-down in their empathy. They use their ideology to assess the legitimacy of an individual's claims rather than evaluate the claims on their own merits. Perhaps, as Ichheister (1970) suggested, "[T]hey restore in this way their peace of mind, for now they can feel that they themselves are right and that 'something is wrong with the other'—to believe which is one of the most essential conditions of happiness in life" (p. 69).

What can be done to forestall victim blaming? One possibility is to affirm people's identities and worldviews through alternative means. Ideologies express identities—who people are, their values, the way they want the world to be. Consequently, people may be more open to information that threatens long-held ideologies when they can affirm alternative identities (Hahn & Cohen, 2009; Cohen et al., 2007). For instance, one study found that having people reflect on an identity unrelated to their political beliefs—for instance, their identity as a "good student" or a "relationship partner"—made them less likely to denigrate victims of inequality (Hahn & Cohen, 2009). It is also possible to reduce defensive denigration by honoring the ideals in the ideology but differentiating those ideals from reality. People who espouse an ideology may assume that

simply because a state of affairs *ought* to be true—such as racial equality—that this state of affairs *is* in fact true (Hahn & Cohen, 2009). We have found that colorblindness ceases to predict empathy when people are given a message that affirms colorblindness as an *ideal* but that portrays the acknowledgment of racial inequality as a necessary *means* to achieving it. Indeed this message lies at the heart of many powerful political speeches, for instance, Martin Luther King Jr.’s “I Have a Dream Speech.” The ideal of racial equality was articulated, but the more subtle message was that attaining this ideal required recognition of our society’s failure to yet reach it. More generally, research has found that people are more open to policies at odds with their ideologies when those policies are framed in ways that affirm rather than threaten their moral allegiances and identities. For instance, individualists are more likely to credit evidence of climate change when they are told that one solution is nuclear power, a commerce that affirms their values, than when told that only increased carbon emission limits, a policy that threatens their values, must be implemented (Kahan, 2010; Kahan, Braman, Cohen, Gastil, & Slovic, in press).

The role of identity in resistance to persuasion and inter-group conflict

People generally resist arguments and evidence that challenge the validity of long-held beliefs. Partisans in negotiation resist agreements that demand compromise, even when the cost of inflexibility is heavy. Liberals and conservatives endlessly debate appropriate social policy, such as health care policy, without either side yielding much ground. In one study, liberals and conservatives rejected welfare policies that were ascribed to the political opposition, even for policies would have received their strong support the absence of any information about their source (Cohen, 2003). Similarly, Israelis and Palestinians rejected even their own side’s peace proposal when led to believe, through an experimental ruse, that it was offered by the other side

(Maoz, Ward, & Ross, 2002). People may resist persuasion attempts and reject pragmatic negotiation compromises, because acquiescing to those attempts, or accepting those compromises, would be costly to their sense of identity. Long-held beliefs and group allegiances are often tied to people's identity. As such they can be like treasured possessions that are difficult to give up (Abelson & Prentice, 1989).

Many economic and political factors underlie the Israeli-Palestinian conflict. However, commentators and researchers suggest that the conflict also involves the defense of identity and identity-relevant symbols (Kelman, 2006). For example, sovereignty of Jerusalem's holy sites is arguably a symbolic issue. "To have 'sovereignty' over the Temple Mount," Margalit (2000) writes, "implies no practical difference. The Palestinians preside over its administration now and will continue to do so. But it makes all the difference in the world, or in the other-world, for the two contesting sides." After Israel acquired Jerusalem, "To lose Jerusalem" came to mean "to lose . . . the belief in things worth sacrificing your life for."

Because of such identity concerns, negotiation benefits from social gestures that affirm each side's cultural narrative, even when they have no material or economic consequence (Kelman, 2006; see also Cohen et al., 2007). For example, when adversaries express empathy for the other side's suffering, or offer an apology for past injustices, this increases the other side's willingness to negotiate and compromise, sometimes dramatically, provided a level of trust has been established (Nadler & Liviatan, 2006; Atran & Ginges, 2009; see also Kelman, 2006).

Intuitive strategies to overcome resistance to persuasion and compromise may prove ineffective or even counter-productive. For instance, telling people to be unbiased (Lord, Lepper, & Preston, 1984) or highlighting their commitment to objectivity (Uhlmann & Cohen, 2007) fails to reduce bias and may even exacerbate it. People may adopt an "I think it, therefore it's true"

mindset, in which they assume that a belief or even hunch, simply by virtue of being theirs, is objective and therefore worthy of being acted on (Uhlmann & Cohen, 2007; see also Kennedy & Pronin, this volume). As a consequence, encouraging people to be objective may lead them to use their prior beliefs more rather than less. For instance, one study found that asking partisans of capital punishment to be objective had no impact on their tendencies to resist belief-confirming evidence and to accept belief-disconfirming evidence (Lord et al., 1984). In another study, having people assert their commitment to objectivity increased their reliance on gender-stereotypical beliefs that they would have otherwise suppressed. They engaged in relatively more discrimination, favoring a male over a female candidate for stereotypically masculine jobs like police chief (Uhlmann & Cohen, 2007).

One effective theory-driven intervention involves affirming alternative sources of identity through self-affirmation. If people focus on important sources of self-worth beyond the persuasion or negotiation topic, it becomes psychologically less painful for them to change their minds or to compromise. By affirming alternative sources of self-worth, people broaden their momentary identity to encompass domains beyond the provoking threat (Schmeichel & Vohs, 2009; Sherman & Cohen, 2006). They can thus yield without its being as costly to their self-worth. An anecdote conveyed by former President Jimmy Carter is suggestive (Carter & Richardson, 1998). In the final hours of the peace talks at Camp David between the Prime Minister of Israel, Menachem Begin, and the President of Egypt, Anwar Sadat, negotiations had reached an impasse. Prime Minister Begin wanted to leave without an agreement. Then Carter gave Begin a series of photographs of him, Carter, and Sadat. Carter had personally addressed each of the photographs with the name of one of Begin's grandchildren:

... I handed it to him. And he started to talk to me about the breakdown of the negotiations and he looked down and saw that I had written all of his grandchildren's names on the individual pictures and signed them, and he started telling me about his favorite grandchild and the characteristics of different ones. And he and I had quite an emotional discussion about the benefits to my two grandchildren and to his if we could reach peace. And I think it broke the tension that existed there, that could have been an obstacle to any sort of resolution at that time.

Carter's actions broadened Begin's perspective beyond the economic and political calculus that can dominate negotiations. More generally, being reminded of higher-order values such as one's family allows people to see disagreement in context. "Being right" or "getting one's way" fades in importance, making room for other motives, such as "being pragmatic" and "finding a compromise" that take into account broader costs and benefits. This can be particularly beneficial in settings where compromise is psychologically painful but pragmatically wise.

Early research predicted that allowing political partisans to affirm an alternative source of identity before reviewing a counter-attitudinal persuasive report would increase their openness (Cohen, Aronson, & Steele, 2000). The participants were proponents and opponents of the death penalty. They reviewed a report concerning capital punishment. For death penalty proponents, the report offered strong arguments and scientific evidence that challenged the deterrent efficacy, economic soundness, and moral justification of the death penalty. For death penalty opponents, the report took the opposite position. (The report was fabricated but appeared to be an authentic article from a social science journal.)

For the self-affirmation manipulation, half of the participants were randomly assigned to complete a self-affirmation exercise before reading the persuasive report. These participants wrote about the importance of a personal value unrelated to politics, such as their relationships with friends, their sense of humor, or their creativity; or in another study they received positive feedback on their social intelligence. Participants thus received an affirmation of an important source of self-worth unrelated to the threatening information. Relative to a no affirmation control condition, the affirmation made participants, regardless of their ideological position, more open to the persuasive evidence. Affirmed proponents and opponents of the death penalty proved significantly more likely to show substantive attitude change in the direction of the report they had read.

Another study addressed why some partisans are sometimes resistant to objectively good news (Bastardi, Ross, & Sherman, 2004). The researchers posited that objectively good news could threaten a partisan's identity. If a person stakes his or her sense of self-worth on an issue, then discovering that the issue is actually a non-issue may be psychologically painful. If true, affirmation should make partisans more open to such identity-threatening "good" news. In one study, a group of roughly thirty environmentalists were found to be significantly more open to evidence that global catastrophe was not imminent when they received a self-affirmation unrelated to their identity as an environmentalist—for instance, when they reflected on an alternative value, such as their relationships with friends and family. By contrast, they were relatively dismissive of such evidence without an affirmation of an unrelated identity—for instance, in a condition where they affirmed their commitment to environmentalism or wrote about a threatening personal experience. Although something of a caricature, it seems that the environmentalists, unless self-affirmed, would rather believe the world could end than be wrong. The results echo earlier

research by Festinger, Riecken, and Schachter (1956), who showed that doomsday groups had a vested interest in seeing their prophecies of world destruction come true and experienced psychological dissonance when they did not. The psychological drive for identity may eclipse even basic human motives to be rational, pragmatic, and optimistic.

Additionally, self-affirmation has been found to increase people's openness to health-risk information. Sherman, Nelson, and Steele (2000) began with the idea that people may feel threatened by information suggesting they engage in behavior that puts their health at risk. Such information can challenge not only their wish to be free of risk but their identity as a rational person. They brought coffee-drinkers to their laboratory and had them read a scientific report linking caffeine intake to cancer (the report was fabricated). Under normal circumstances, coffee-drinkers, in contrast to non-coffee-drinkers, dismissed the evidence. However, coffee-drinkers asked to complete a self-affirmation, in which they reflected on an important personal value, were relatively more open to and accepting of the report's findings. Subsequent studies have found that affirmation-induced attitude change can persist over time and improve health behavior (see Epton & Harris, 2008).

Affirmation-induced openness is not simply a rose-colored glasses effect, in which people become generally more positive toward new ideas and information. Research finds that affirmation increases openness primarily to strong rather than weak counter-attitudinal arguments, while decreasing acceptance of weak rather than strong pro-attitudinal arguments (Correll, Spencer, & Zanna, 2004). Self-affirmation, it seems, gives people the inner security they need to accept valid new ideas that they would otherwise reject and to abandon invalid old ideas to which they would otherwise cling. By allaying self-protective needs, affirmation permits other motives—such as motives for accuracy and objectivity—to predominate (Kunda, 1990).

Another study tested whether affirmation could be used to improve negotiation. Pro-choicers engaged in a negotiation with a pro-lifer over abortion legislation (Cohen et al., 2007). At the outset, they completed an exercise designed to be either self-affirming (writing about a time when they lived up to a personally important value, such as relationships with friends) or self-threatening (writing about a time when they failed to live up to an important value). The study crossed that experimental manipulation with a manipulation of the salience of participants' political identity. In the political identity-salient condition, participants asserted, prior to the negotiation, their true beliefs about appropriate abortion policy. In the political identity non-salient condition, they did not. Participants then took part in a negotiation with their adversary over abortion legislation, and we counted the number of concessions they made to their political adversary. When political identity was not salient—when people were not focused on defending their political identity—there was no effect of affirmation. Because people had relatively less identity or ego at stake in the situation, an ego-buttressing intervention had no impact. But in the condition where people were focused on their political identity, affirmation had a large effect, nearly doubling the number of concessions participants made to their adversary. Perhaps more important, affirmation also increased participants' trust in their negotiation partner. Affirmed negotiators were less likely to ascribe bias to him, and felt more open to negotiating with him in the future. This latter finding is particularly promising in light of the importance of trust in the resolution of inter-group conflict in the real world, where mutual suspicion can lead adversaries to view one another's conciliatory gestures with suspicion, triggering a downward spiral (see Nadler & Liviatan, 2006; Kennedy & Pronin, this volume).

Intuitively, one might think that expressing one's commitment to an issue, one's true beliefs, would make partisans more resistant to change. Indeed, there was a trend in this direction

in the condition among the nonaffirmed. However, when affirmed, commitment made partisans more open to compromise. Bringing to the fore the values associated with a particular identity—and one’s commitment to the cause on which that identity is staked—may facilitate openness, as long as ego-protective motives, such as the desire to be “right” or to “triumph over one’s adversary,” are lifted through the affirmation of alternative sources of self-worth.

Affirmation processes also apply to the effects of racial identity on school achievement. Education is, in many ways, like persuasion. Openness to new information and ideas is critical. But if a minority student is working in a racially diverse school, with much instruction and social interaction and instruction taking place across racial lines, racial identity may be salient. The salience of race could make the school environment more psychologically threatening for minorities (Steele, Spencer, & Aronson, 2002). For example, a minority student may wonder if a teacher’s critical feedback reflects a genuine intent to help or racial bias (Cohen, Steele, & Ross, 1999). Given this, we wondered if affirmation would increase minority students’ openness to instruction in an ethnically diverse school.

To address this question, we conducted a field experiment in an ethnically diverse middle school (Sherman & Cohen, 2006; see also Cohen, Garcia, Apfel, & Master, 2006; Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009). The student body was split roughly equally between Black Americans and White Americans, and virtually all the teachers were White. At the beginning of the school year, in the context of a randomized, double-blind experiment, seventh grade students were assigned to complete either a series of self-affirmation exercises similar to the ones described previously, in which they reflected on an important personal value, or a series of control exercises in which they wrote about an unimportant value or a daily routine. The exercises were presented as a regular classroom assignment. Random assignment occurred within each

participating classroom, such that half the children completed affirmation exercises, while the remainder completed control exercises. Similar exercises were repeated throughout the school year. Careful methodological steps ensured that teachers remained unaware of their students' condition assignment. Additionally, children remained unaware both that they were part of an intervention study and that there were different versions of the exercises given to their peers (see Cohen et al., 2006, 2009). We used a questionnaire to measure students' trust in their teachers—the extent to which they viewed their teachers as fair and “on their side”—both at the beginning of the year, prior to the intervention, and then at the end of the academic year.

In the control condition, a racial divide in perception emerged over the course of middle school (Sherman & Cohen, 2006; see also Cohen et al., 2006). Without affirmation, Black students' trust in their teachers dropped more than did that of White students. Seventh grade is a year in which race and its significance become particularly salient, something that might explain this drop in trust (Simmons, Black, & Zhou, 1991). By contrast, affirmed Black students' trust proved significantly less likely to drop and more likely to maintain itself over time.

If affirmation maintains students' trust of their teachers, this might translate into better learning and grades. Consistent with this notion, affirmed Black students earned significantly better grades than did their peers in the control condition, for instance, reducing the percentage of students earning a D or below in the course from 20% to 9% (Cohen et al., 2006.), with the performance benefits lasting at least two years (Cohen et al., 2009). While many mechanisms probably underlie this performance benefit (Cohen et al., 2006, 2009), trust may be one component.

Those findings provide promising evidence of the applicability of self-affirmation and identity processes to the real world. Affirmation can bridge not only partisan divides but also

racial divides. In real-world settings, where relationships continue over time, these processes can be recursive in nature (Cohen et al., 2009). Mistrust can feed off its consequences, leading people to see bias in the behavior of others, which in turn deepens their mistrust, in a potentially repetitive cycle (Kennedy & Pronin, this volume). Accordingly, if one can intervene early enough to interrupt the recursive process before it accelerates, lasting positive change may follow.

The role of constructed criteria in licensing discrimination

Many Americans experience a tension between their meritocratic values on the one hand and conscious or unconscious prejudicial beliefs on the other (Gaertner & Dovidio, 1986). The mandates of meritocracy demand that they hire people based on their merits rather than their group memberships. However, sexist and racist beliefs push them to discriminate. We have found that people can resolve this tension by changing their very definition of merit (Uhlmann & Cohen, 2005). People spontaneously construct and reconstruct criteria of merit in a way that advantages positively stereotyped groups (see also Hodson, Dovidio, & Gaertner, 2002; Norton, Vandello, & Darley, 2004). Such “constructed criteria” enable people to discriminate while resting assured of their personal fairness and objectivity. Indeed, a recent field experiment found that such shifts in hiring criteria are among three major contributors to real-world discrimination in low-wage labor markets (Pager, Western, & Bonikowski, in press).

In one study, participants reviewed either a male or a female candidate for the job of police chief in a mock hiring scenario (Uhlmann & Cohen, 2005). The candidate was described as either “book smart”—someone with a lot administrative experience, who was well-educated and had a family and kids—or “street-smart”—someone who made a lot of arrests, was tough, worked a lot, and was single. Participants were then asked to rate the importance of each of these traits to the job of police chief.

Whatever credentials the male candidate held were seen by participants as relatively important for the job. For instance, when he had book smarts, book smarts were seen as important to being a police chief. In fact, there was a very strong correlation between ratings of the candidate's standing on a given trait (for example, "administrative skills") and the perceived importance of that trait. Even seemingly peripheral traits, like having children, or stereotypically feminine traits, like being "family oriented," were rated as more important when the male candidate had them than when he did not. Likewise, when the male candidate had street smarts, the book-smart traits were seen as relatively less important and alternative traits like "being tough" and "making a lot of arrests" were rated as relatively more important. In short, participants' response to the male candidate seemed to be, "I can see how this person will bring assets to the position."

By contrast, there was no such effect when the candidate was female. If anything, there was a non-significant reverse tendency for subjects to see whatever credentials the female applicant had as less important to the job. On the whole, the relationship between the female candidate's perceived standing on a given trait and the perceived importance of that trait was nil. Participants' response was one of neutrality. They did not allow the female applicant's credentials to affect their perceived importance. Their response seemed to be, "I'm going to be neutral. On the one hand, she has certain credentials. On the other hand, here's what's important to the job. I'm not going to confuse the two." In this way, participants applied different definitions of fairness when evaluating candidates of different genders. For male candidates, they defined fair as appreciating what the candidate could bring to the position. For female candidates, they defined fair as being neutral. In both cases, participants were fair, but what they failed to recognize was that they had applied different definitions of fairness for men and women. Importantly, men may

face the same bias when they apply for jobs where they are stereotyped as inferior. We found that people constructed hiring criteria that disadvantaged men applying for high-status jobs typically filled by females, such as “women’s studies professor” (Uhlman & Cohen, 2005).

Participants thus defined the criteria of merit in a way congenial to the idiosyncratic credentials of the positively stereotyped applicant. As a consequence, they could discriminate while resting assured that their hiring decisions followed from “objective” standards of merit. Indeed, those participants who constructed more discrimination-justifying criteria proved relatively more likely to see their hiring decisions as objective. That is, those who engaged in the most bias claimed the least bias. Participants, it seems, believed they had picked exactly the right man for the job when in reality they had picked exactly the right criteria for the man.

These results suggest a theory-driven intervention: Have people commit to hiring criteria before reviewing applicants. In a follow up study, we had a random subset of participants rate the importance of various criteria for the job of police chief, such as traits related to book smarts and street smarts, *before* they saw the applicants’ credentials. Those participants who did not do this exercise engaged in gender discrimination. They preferred the male over the female candidate. By contrast, participants who pre-committed to hiring criteria showed no gender discrimination. When people no longer had the ability to construct biased hiring criteria, discrimination disappeared. More research is needed to assess whether the same de-biasing strategy would work in actual work settings with real stake-holders, but these results provide a promising foundation for effective intervention.

Conclusion

The research discussed here illustrates the role of ideology in punitive reactions to victims; the role of identity in bias, closed-mindedness and social conflict; and the role of constructed

criteria in the maintenance of discrimination. Beliefs, ideologies, and identities subtly and unconsciously bias thoughts, feelings, and behaviors, and thus exacerbate social problems such as inequality and inter-group conflict. However, theory-driven interventions, attuned to important psychological processes, can reduce bias and change outcomes for the better.

References

- Abelson, R. P., & Prentice, D. A. (1989). Beliefs as possessions: A functional perspective. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 361-381). Hillsdale, NJ: Erlbaum.
- Atran, S., & Ginges, J. (2009, January 25). How words could end a war. *New York Times* (New York), pp. WK12. .
- Bastardi, A., Ross, L., & Sherman, D. (2004). Maintaining self-integrity: Effects of threat to personal identity on evidence evaluation and concession making. Unpublished manuscript.
- Carter, J., & Richardson, D. (1998). *Conversations with Carter*. Boulder, CO: Lynee Rienner Publishers.
- Cohen, G. L. (2003). Party over policy: The dominating impact of group influence on political beliefs. *Journal of Personality and Social Psychology*, *85*, 808-822.
- Cohen, G. L., Aronson, J., & Steele, C. M. (2000). When beliefs yield to evidence: Reducing biased evaluation by affirming the self. *Personality and Social Psychology Bulletin*, *26*, 1151-1164.
- Cohen, G. L., Sherman, D. K., Bastardi, A., Hsu, L., McGoey, M., & Ross, L. (2007). Bridging the partisan divide: Self-affirmation reduces ideological closed-mindedness and inflexibility in negotiation. *Journal of Personality and Social Psychology*, *93*, 415-430.
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*, 1307-1310.
- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science*, *324*, 400-403.

Cohen, G. L., Scrivener, R., & Miles, J. (2009). The blindness in colorblindness: The effect of racial ideologies on empathy toward outgroup members. Manuscript in preparation.

Cohen, G. L., Steele, C. M., & Ross, L. D. (1999). The mentor's dilemma: Providing critical feedback across the racial divide. *Personality and Social Psychology Bulletin*, *25*, 1302-1318.

Correll, J., Spencer, S. J., & Zanna, M. P. (2004). An affirmed self and an open mind: Self-affirmation and sensitivity to argument strength. *Journal of Experimental Social Psychology*, *40*, 350-356.

Epton, T., & Harris, P. R. (2008). Self-affirmation promotes health behavior. *Health Psychology*, *27*, 746-752.

Festinger, L., Riecken, H. W., Schachter, S. (1956). *When prophecy fails: A social and psychological study of a modern group that predicted the destruction of the world*. Minneapolis, MN: University of Minnesota Press.

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

Gerber, A., & Green, D. (1999). Misperceptions about perceptual bias. *Annual Review of Political Science*, *2*, 189–210.

Hahn, A., & Cohen, G. L. (2009). From an “ought” to an “is”: Defensive maintenance of social values. Manuscript in preparation.

Hodson, G., Dovidio, J.F., & Gaertner, S.L. (2002). Processes in racial discrimination: Differential weighting of conflicting information. *Personality and Social Psychology Bulletin*, *28*, 460-471.

Ichheiser, G. (1970). *Appearances and reality*. San Francisco: Jossey-Bass.

Kahan, D. (2010). Fixing the communications failure. *Nature*, *463*, 296-297.

Kahan, D. M., Braman, D., Cohen, G. L., Gastil, J., & Slovic, P. (in press). Who Fears the HPV Vaccine, Who Doesn't, and Why? An Experimental Study of the Mechanisms of Cultural Cognition. *Law and Human Behavior*.

Kelman, H. (2006). Interests, relationships, identities: Three central issues for individuals and groups in negotiating their social environment. *Annual Review of Psychology, 57*, 1-26.

Kruglanski, A.W., Webster, D.M., & Klein, A. (1993). Motivated resistance and openness to persuasion in the presence or absence of prior information. *Journal of Personality and Social Psychology, 65*, 861-876.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*, 480-498.

Kunda, Z., Davies, P. G., Adams, B. D., & Spencer, S. J. (2002). The dynamic time course of stereotype activation: Activation, dissipation, and resurrection. *Journal of Personality and Social Psychology, 82*, 283-299.

Lewin, K. (1952). Group decision and social change. In G. E. Swanson, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in social psychology*. (pp. 459-473). New York: Henry Holt.

Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology, 47*, 1231-1243.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*, 2098-2109.

Miller, D., & Prentice, D. (1999). Some consequences of a belief in group essence: The category divide hypothesis. In D. Prentice, D. Miller (Eds.), *Cultural divides: Understanding and overcoming group conflict* (pp. 213-238). New York: Russell Sage Foundation.

- Maoz, I., Ward, A., & Ross, L. (2002). Reactive devaluation of an “Israeli” vs. “Palestinian” peace proposal. *Journal of Conflict Resolution*, 46, 515-546.
- Margalit, A. (2000, September 21). The odds against Barak. *The New York Review of Books*. Retrieved February 10, 2010, from <http://www.nybooks.com/articles/13812>
- Nadler, A., & Liviatan, I. (2006). Intergroup reconciliation: Effects of adversary’s expressions of empathy, responsibility, and recipient’s trust. *Personality and Social Psychology Bulletin*, 32, 459-470.
- Norton, M. I., Vandello, J. A., & Darley, J. (2004). Casuistry and social category bias. *Journal of Personality and Social Psychology*, 87, 817-831.
- Pager, D., Western, B., & Bonikowski, B. (in press). Race at work: A field experiment of discrimination in low-wage labor markets. *American Sociological Review*.
- Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review*, 111, 781-799.
- Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28, 369-381.
- Ross, L., & Ward, A. (1995). Psychological barriers to dispute resolution. In M. Zanna (Ed.), *Advances in Experimental Social Psychology*, Volume 27 (pp. 255-304). San Diego: Academic Press.
- Schmeichel, B. J., & Vohs, K. D. (2009). Self-affirmation and self-control: Affirming core values counteracts ego depletion. *Journal of Personality and Social Psychology*, 96, 770-782.
- Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 38, pp. 183-242). San Diego, CA: Academic Press.

Sherman, D. A. K., Nelson, L. D., & Steele, C. M. (2000). Do messages about health risks threaten the self? Increasing the acceptance of threatening health messages via self affirmation.

Personality and Social Psychology Bulletin, 26, 1046-1058.

Simmons, R. G., Black, A., & Zhou, Y. (1991). African-Americans versus White children and the transition into junior high school. *American Journal of Education*, 99, 481-520.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379-440). New York, NY: Academic Press.

Tetlock, P. E. (2005). *Expert Political Judgment: How Good is It? How Can We Know*. Princeton, NJ: Princeton University Press.

Uhlmann, E., & Cohen, G. L. (2005). Constructed criteria: Redefining merit to justify discrimination. *Psychological Science*, 16, 474-480.

Uhlmann, E., & Cohen, G. L. (2007). "I think it, therefore it's true": Effects of self-perceived objectivity on hiring discrimination. *Organizational Behavior and Human Decision Processes*, 104, 207-223.

Vallone, R. P., Ross, L., & Lepper, M. R. (1985). The hostile media phenomenon: Biased perception and perceptions of media bias in coverage of the Beirut Masacre. *Journal of Personality and Social Psychology*, 49, 577-585.

*