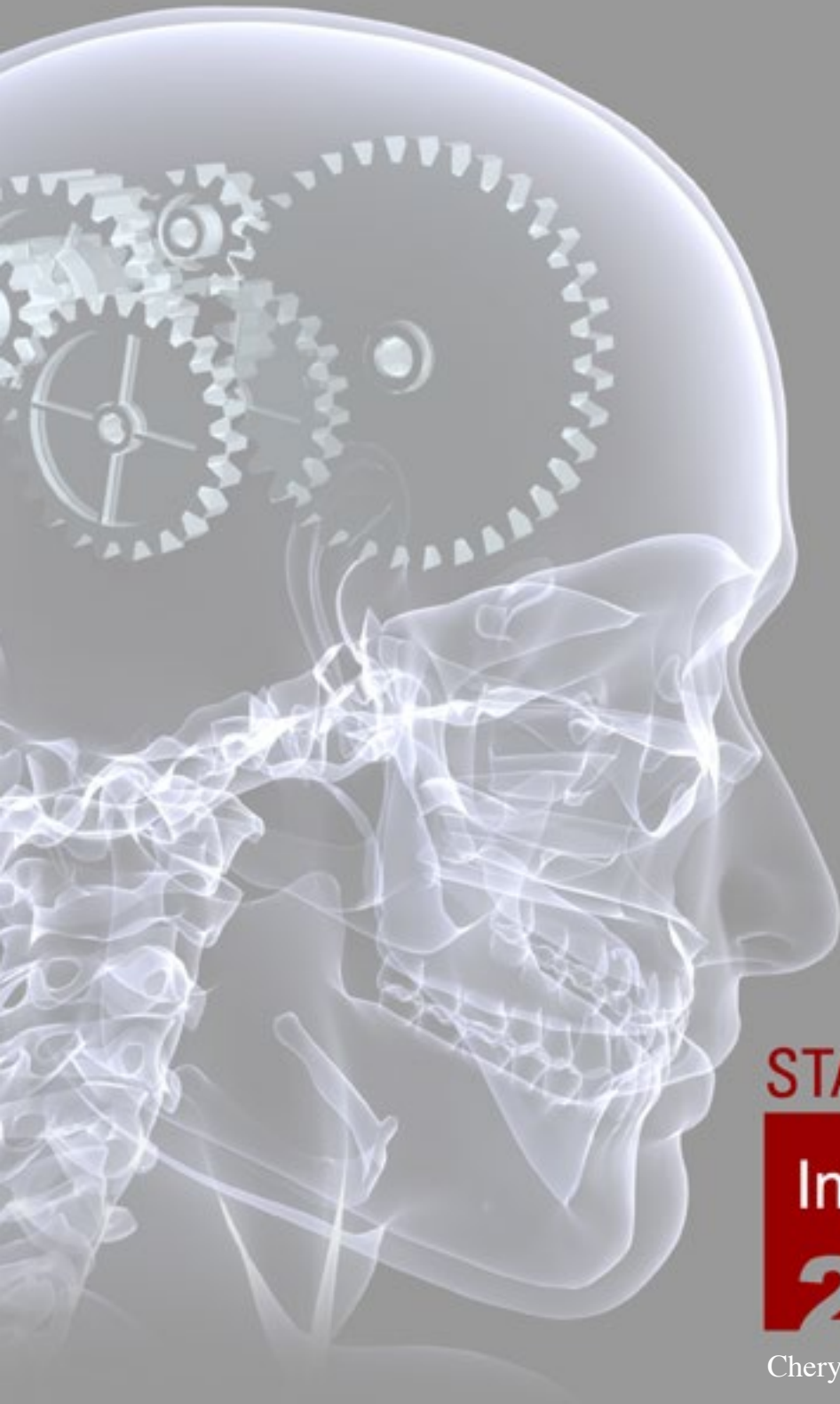THE OHIO STATE UNIVERSITY

KIRWAN INSTITUTE
for the Study of Race and Ethnicity

STATE OF THE SCIENCE

Implicit Bias Review

2013

Cheryl Staats, Research Associate

with contributions from
Charles Patton, Graduate Research Associate

kirwaninstitute.osu.edu

KIRWAN INSTITUTE
for the Study of Race and Ethnicity

THE OHIO STATE UNIVERSITY

The Kirwan Institute works to **create** a just and inclusive society where all people and communities have **opportunity** to succeed.

33 West 11th Avenue
Columbus, Ohio 43210
(614) 688-5429

Sharon Davies, Executive Director

Jason Reece, Research Director

Christy Rogers, Outreach Director

Thomas Rudd, Education and Emerging Research Director

STATE OF THE SCIENCE

Implicit Bias Review
2013

The Purpose and Goal of The Implicit Bias Review

The Kirwan Institute for the Study of Race and Ethnicity has become increasingly mindful of how implicit racial biases shape not only individuals' cognition and attitudes, but also their behaviors. Indeed, a large body of compelling research has demonstrated how these unconscious, automatically activated, and pervasive mental processes can be manifested across a variety of contexts yielding significant impacts. Consider these striking examples:

In a video game that simulates what police officers experience, research subjects were instructed to "shoot" when an armed individual appeared on the screen and refrain from doing so when the target was instead holding an innocuous object such as a camera or wallet. Time constraints were built into the study so that participants were forced to make nearly instantaneous decisions, much like police officers often must do in real life. Findings indicated that participants tended to "shoot" armed targets more quickly when they were African American as opposed to White. When participants refrained from "shooting" an armed target, these characters in the simulation tended to be White rather than African American. Moreover, in circumstances where the target was "shot" in error (i.e., was "shot" even though they were wielding a harmless object), those targets were more likely to be African American than White. Research such as this highlights how implicit racial biases can influence decisions that have life or death consequences.

A 2012 study used identical case vignettes to examine how pediatricians' implicit racial attitudes affect treatment recommendations for four common pediatric conditions. Results indicated that as pediatricians' pro-White implicit biases increased, they were more likely to prescribe painkillers for vignette subjects who were White as opposed to Black patients. This is just one example of how understanding implicit racial biases may help explain differential health care treatment, even for youths.

Examining implicit bias research is important for all who work for racial justice because of the rich insights into human behavior that this work generates. Moreover, as convincing research evidence accumulates, it becomes difficult to understate the importance of considering the role of implicit racial biases when analyzing societal inequities. Implicit biases, explicit biases, and structural forces are often mutually reinforcing, thus multiple levels of analysis are necessary to untangle the nuances of these complex dynamics.

With this in mind, and as a testament to the Kirwan Institute's belief in the importance of understanding implicit bias, we present to you this inaugural edition of our "State of the Science Review of Implicit Bias Learning." As an annual publication, subsequent editions of this Review will highlight the latest research findings and underscore trends in the field. Our goal is to provide a comprehensive resource that communicates this research in a concise and accessible manner while stimulating further dialogue on implicit bias and its implications for the pursuit of social justice.

Introduction

Introduction

> " *At the nexus of social psychology, cognitive psychology, and cognitive neuroscience has emerged a new body of science called "implicit social cognition" (ISC). This field focuses on mental processes that affect social judgments but operate without conscious awareness or conscious control.* "
>
> *– Kang & Lane, 2010, p. 467 –*

Although not yet a widely-known concept outside of the social science community, knowledge of implicit bias is gradually infiltrating the public domain. Attention from the media and other sources devoted to how implicit biases may have influenced voters' decisions in the 2008 and 2012 presidential elections is permeating public consciousness (see, e.g., Greenwald, Smith, Sriram, Bar-Anan, & Nosek, 2009; McElroy, 2012; NPR, 2012; Payne, et al., 2010). The term was also emphasized in an April 2012 decision by an Iowa district court judge, in a class-action suit brought forth by African Americans who claimed that implicit racial biases influenced employment and promotion decisions for state jobs ("Iowa: Ruling for State in 'Implicit Bias' Suit," 2012). As the body of literature on implicit bias expands and the scholarship gains traction outside of academic circles, one can reasonably anticipate that implicit bias will increasingly enter public discourse.

Implicit bias refers to the attitudes or stereotypes that affect our understanding, actions, and decisions in an unconscious manner. Among the key attributes of implicit biases are the following:

- **Implicit or Unconscious:** Much of the literature suggests that these biases are activated unconsciously, involuntarily, and/or without one's awareness or intentional control (see, e.g., Dovidio, Kawakami, Smoak, & Gaertner, 2009; Greenwald & Krieger, 2006; Kang, et al., 2012; Nier, 2005; Rudman, 2004). The appropriateness of using terms such as 'unconscious' and 'non-conscious' with respect to the activation and application of implicit biases has been questioned by some (Fazio & Olson, 2003; Rudman, 2004). Implicit biases are defined as "attitudes and stereotypes that are not consciously accessible through introspection"(Kang, et al., 2012, p. 1132). This is in direct contrast

with explicit biases, meaning those that are held or endorsed on a conscious level. The distinction between implicit and explicit is further discussed in chapter 2 of this literature review.

- **Bias:** Bias "denotes a displacement of people's responses along a continuum of possible judgments" (Greenwald & Krieger, 2006, p. 950). This bias may skew toward either a favorable or an unfavorable assessment (Greenwald & Krieger, 2006).

- **Automatically Activated / Involuntary:** Implicit biases can activate without intention and/or without being explicitly controlled (i.e., not deliberate) (Blair, 2002; Rudman, 2004).

- **Pervasiveness:** Substantial research has established that implicit attitudes and stereotypes are robust and pervasive (Greenwald, McGhee, & Schwartz, 1998; Kang, et al., 2012; Kang & Lane, 2010; Nosek, Smyth, et al., 2007).

The Formation of Implicit Biases

Regardless of whether implicit associations are positive or negative in nature, everyone is susceptible to implicit biases, including children (Nosek, Smyth, et al., 2007; Rutland, Cameron, Milne, & McGeorge, 2005). Implicit biases vary among individuals (see, e.g., D. M. Amodio, Harmon-Jones, & Devine, 2003).

The classic nature versus nurture debate is one way to approach the question of how our implicit associations are formed. A nature-based argument would assert that biases are hardwired. Research by Mahajan and colleagues broadly supports this argument, as they identified the existence of implicit ingroup preferences even among a nonhuman species, rhesus macaques (Mahajan, et al., 2011).

Renown implicit bias scholar Jerry Kang addresses the nature vs. nurture debate and sides with nurture, writing that "even if nature provides the broad cognitive canvas, nurture paints the detailed pictures – regarding who is inside and outside, what attributes they have, and who counts as friend or foe" (Kang, 2012, p. 134). Supporting

this declaration, he notes how entertainment media perpetuates stereotypes, citing one study that documented the subtle transmission of race biases through nonverbal behaviors seen on television. Exposure to nonverbal race bias on television can influence individuals' race associations and attitudes, as "exposure to pro-white (versus pro-black) nonverbal bias increased viewers' bias even though patterns of nonverbal behavior could not be consciously reported"(Weisbuch, Pauker, & Ambady, 2009, p. 1711). Kang also discusses how news programming, particularly the excessive portrayal of Blacks as criminal (see, e.g., Dixon & Linz, 2000; Oliver, 2003), helps foster the formation of implicit biases (Kang, 2012). He even recognizes how online virtual gaming worlds can contribute to implicit biases.

Rudman (2004) outlines five factors that influence the formation of implicit orientations more so than explicit ones. She asserts that early experiences may lay the foundation for our implicit attitudes while explicit attitudes are more affected by recent events (Rudman, 2004). Also on Rudman's list are affective experiences (vs. self-reports), cultural biases (vs. explicit beliefs), and cognitive balance principles (Rudman, 2004). While these previous four principles were already established in the literature, Rudman adds a fifth for consideration: the self. She writes, "it may be difficult to possess implicit associations that are dissociated from the self, whereas controlled evaluations may allow for more objective responses. If the self proves to be a central cause of implicit orientations, it is likely because we do not view ourselves impartially, and this partisanship then shapes appraisals of other objects that are (or are not) connected to the self" (Rudman, 2004, p. 137).

The Importance of Implicit Bias to the Work of Social Justice Advocates

Understanding the nuances of implicit bias is critical for addressing the inequalities that are byproducts of structural forces. Indeed, Kang notes that implicit biases, explicit biases, and structural forces are not mutually exclusive but instead often reinforce one another (Kang, et al., 2012). Elsewhere he affirms this interrelation by writing that "the deepest understanding of any process such as racialization comes from multiple levels of analysis that can and should be integrated together" (Kang, 2010, p. 1147).

powell and Godsil emphasize that the human behavior insights gleaned from the study of implicit biases are key to achieving social justice goals (powell & Godsil, 2011). They declare that knowledge of how the brain functions, particularly how we understand our positioning with respect to our environment, is key for the creation of "a political space in which it is possible to first have a constructive dialogue about the continuing salience of race, then generate support for the policies necessary to address the role race continues to play, and finally, and as importantly, develop implementation measures that will allow these policies to achieve the sought-after outcomes"(powell & Godsil, 2011, p. 4).

Rudman unequivocally asserts the significance of understanding implicit biases with these cautionary words: "for a deep and lasting equality to evolve, implicit biases must be acknowledged and challenged; to do otherwise is to allow them to haunt our minds, our homes, and our society into the next millennium"(Rudman, 2004, p. 139).

About this Implicit Bias Review

Although a wide variety of characteristics (e.g., gender, age) can activate implicit biases, this literature review primarily focuses on implicit racial and ethnic biases. While this review is intended to be as complete as possible, it should not be regarded as comprehensive given the vast literature devoted to this topic.

STATE OF THE SCIENCE

Background on Implicit Bias

Background on Implicit Bias

Key concepts

A few fundamental definitions related to mental associations are important to understand before unpacking the nuances of implicit bias.

- **Schemas** are "templates of knowledge that help us to organize specific examples into broader categories" (Kang, 2008). These mental shortcuts allow us to quickly assign objects, processes, and people into categories (Kang, 2009). For example, people may be placed into categories based on traits such as age, race, gender, and the like. Once these categories have been assigned, any meanings that we carry associated with that category then become associated with the object, process, or person in question. The chronic accessibility of racial schemas allow them to shape social interactions (Kang, 2005).

- **Stereotypes** are beliefs that are mentally associated with a given category (Blair, 2002; Greenwald & Krieger, 2006). For example, Asians are often stereotyped as being good at math, and the elderly are often stereotyped as being frail. These associations – both positive and negative – are routinized enough that they generally are automatically accessed (Rudman, 2004). Stereotypes are not necessarily accurate and may even reflect associations that we would consciously reject (Reskin, 2005).

- **Attitudes** are evaluative feelings, such as having a positive or negative feeling towards something or someone (Greenwald & Krieger, 2006; Kang, 2009).

- **Ingroups and Outgroups** – As soon as we see someone, we automatically categorize him or her as either 'one of us', that is, a member of our ingroup, or different from ourselves, meaning a member of our outgroup. Making this simple 'us vs. them' distinction is an automatic process that happens within seconds of meeting someone (Reskin, 2005). Deservedly or not, ingroup bias leads to relative favoritism compared to outgroup members (Greenwald & Krieger, 2006). We extrapolate characteristics about ourselves to other ingroup members, assuming that they are like us compared to outgroup members (Reskin, 2005). By favoring ingroup members, we tend to grant them

a measure of our trust and regard them in a positive light (Reskin, 2005). This ingroup favoritism surfaces often on measures of implicit bias (see, e.g., Greenwald, et al., 1998). The strength of ingroup bias has been illustrated in various studies. For example, research has found that people tend to display ingroup bias even when group members are randomly assigned (Tajfel, Billig, Bundy, & Flament, 1971), and, even more incredibly, when groups are completely fictitious (Ashburn-Nardo, Voils, & Monteith, 2001).

Moreover, given the dynamics of race in our society, it is no surprise that extensive research has documented White Americans' implicit ingroup bias and relative bias against African Americans (see, e.g., Dasgupta & Greenwald, 2001; Dasgupta, McGhee, Greenwald, & Banaji, 2000; Devine, 1989; Dovidio, Kawakami, & Gaertner, 2002; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, et al., 1998; McConnell & Liebold, 2001; Nosek, Banaji, & Greenwald, 2002; Richeson & Ambady, 2003).

Select Seminal Works

Laying the foundation for implicit bias research was a 1983 seminal article by Gaertner and McLaughlin. Employing a lexical decision task, they found that research subjects, regardless of their personal prejudices, were reliably faster at pairing positive attitudes with Whites than with Blacks (in experiment 1) or with Negroes (in experiment 2) (Gaertner & McLaughlin, 1983). However, negative attributes were equally associated with Blacks and Whites. This article is regarded as the first piece to demonstrate implicit stereotyping.

In 1989, an article by Patricia G. Devine was the first to argue that "stereotypes and personal beliefs are conceptually distinct cognitive structures" (Devine, 1989, p. 5). Her dissociation model differentiated between automatic processes (i.e., those that "involve the unintentional or spontaneous activation of some well-learned set of associations or responses that have been developed through repeated activation in memory") and controlled processes (i.e., those that are "intentional and require the active attention of the individual") (Devine, 1989, p. 6). Devine tested the implications of this dissociation model with respect to prejudice and found that in-

dividuals can hold a "clear distinction between knowledge of a racial stereotype …
and personal beliefs about the stereotyped group" (Devine, 1989, p. 15). In short,
automatic and controlled processes can be dissociated such that an individual can
rate as low-prejudiced while still holding knowledge of the existence of that given
stereotype in his/her memory system.

Following this dissociation revelation, researchers who studied automatic associa-
tions generally regarded them as automatic and inflexible; that is, these associations
were deemed spontaneously triggered and inescapable (Bargh, 1997, 1999; Devine,
1989; Dovidio & Fazio, 1992). This notion of these associations being automatic
and unavoidable led to the conclusion that biases remained stable over time and, be-
cause they were so deep seated, were not open to manipulation. In a seminal work,
Irene V. Blair upended this notion by establishing that automatic stereotypes and
prejudice are, in fact, malleable (Blair, 2002). She found that automatic stereotypes
and prejudices may be moderated by events such as contextual cues, the perceivers'
focus of attention, and the perceivers' motivation to maintain a positive self-image,
among others (Blair, 2002). Other researchers built on this scholarly foundation,
suggesting that even if automatic, stereotype activation is not necessarily uncontrol-
lable (Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000).

These works opened up the door to a plethora of research that further examined the
nature of associations, how implicit biases operate, and ways in which they may be
countered. This body of work will be further discussed throughout the course of this
literature review.

Understanding the Relationship Between Implicit and Explicit Bias

Implicit and explicit biases are generally regarded as related yet distinct concepts
(Kang, 2009; Wilson, Lindsey, & Schooler, 2000). They are not mutually exclusive
and may even reinforce each other (Kang, et al., 2012). "Neither should be viewed
as the solely 'accurate' or 'authentic' measure of bias" (Kang, 2009, p. 3).

The main distinction between implicit and other types of bias centers on level of
awareness (Petty, Fazio, & Briñol, 2009). Explicit biases "can be consciously detect-

ed and reported" (D. M. Amodio & Mendoza, 2010, p. 355). Processes that are not explicit are implicit, meaning that they occur without introspective awareness (D. M. Amodio & Mendoza, 2010; Greenwald & Banaji, 1995; Wilson, et al., 2000). Explicit attitudes tend to be associated with deliberate responses that individuals can control (Dovidio, et al., 1997; Nier, 2005). They are often measured by instruments such as feeling thermometers and semantic differentials, in addition to other forms of direct questioning.

Given that implicit associations arise outside of conscious awareness, these associations do not necessarily align with individuals' openly-held beliefs or even reflect stances one would explicitly endorse (Graham & Lowery, 2004; Kang, et al., 2012; Reskin, 2005).

Following the 1989 debut of Devine's dissociation model, further research has explored the idea of whether implicit and explicit biases are dissociated (see, e.g., Dasgupta & Greenwald, 2001; Dovidio, et al., 1997; Green, et al., 2007; Greenwald & Banaji, 1995; Nier, 2005). A vast body of empirical literature documents studies in which respondents' implicit and explicit attitudes do not align (see, e.g., Cunningham, Preacher, & Banaji, 2001; Dasgupta, et al., 2000; Devine, 1989; Dunton & Fazio, 1997; Fazio, et al., 1995; Fazio & Olson, 2003; Greenwald, et al., 1998; Phelps, et al., 2000; von Hippel, Sekaquaptewa, & Vargas, 1997). That said, the literature is inconsistent and hardly conclusive, as other studies have found that implicit and explicit attitudes seem to align, thus calling into question this notion of dissociation (see, e.g., McConnell & Liebold, 2001; Wittenbrink, Judd, & Park, 1997).

To help get to the root of this debate about dissociation, Hofmann et al. performed a meta-analysis to examine the correlation between a measure of implicit bias (the Implicit Association Test, discussed in-depth in chapter 3) and explicit self-report measures. Analyzing this across 126 studies, they uncovered a mean effect size of 0.24, which is relatively small yet significant (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005). Thus, their meta-analysis concluded that this implicit measure and explicit self-reports are systemically related rather than dissociated. They also noted that variations in correlations between implicit and explicit measures can be attributed to how spontaneous the self-reports are and the degree of conceptual

correspondence between measures (Hofmann, et al., 2005).

In a search to explain these seemingly contradictory results between implicit and explicit measures, other factors have been identified as moderating variables (Rudman, 2004). These include individual motivation to report explicit attitudes that align with one's implicit attitudes (Dunton & Fazio, 1997; Fazio, et al., 1995; Nier, 2005), the psychometric properties of the specific measurement techniques (Cunningham, et al., 2001; Greenwald, Nosek, & Banaji, 2003) and impression management or social desirability concerns, as discussed in the next section (Dunton & Fazio, 1997; Nier, 2005; Nosek & Banaji, 2002).

Downfalls of self-reports and other explicit measures of bias

Early researchers relied on explicit measurements of prejudice, such as the Bogardus Social Distance Scale (Bogardus, 1933). But as norms discouraging prejudice gained societal traction, straightforward approaches to measuring bias became less useful and increasingly suspect. Researchers were left to wonder whether stereotypes were fading, whether the content of stereotypes had changed, or whether people were simply suppressing their negative views of others (D. Amodio & Devine, 2009).

The downfalls of self-reports have been well-documented since at least the 1960s (Orne, 1962; Weber & Cook, 1972). Impression management can undermine the validity of self-report measures of bias, as the desire to be perceived positively can influence people to distort their self-reported beliefs and attitudes (D. Amodio & Devine, 2009; Dovidio, et al., 1997; Fazio, et al., 1995; Greenwald & Nosek, 2001; Greenwald, Poehlman, Uhlmann, & Banaji, 2009; Nier, 2005; Nosek, Greenwald, & Banaji, 2007). In 1971, Harold Sigall and colleagues famously exposed how social desirability can taint self-reports by employing a "bogus pipeline" machine that the researchers claimed would reveal participants' true inner attitudes (Jones & Sigall, 1971; Sigall & Page, 1971). The portion of participants who were led to believe in the effectiveness of the machine reported attitudes that more closely reflected their true beliefs compared to those who did not believe they were being monitored and thus were free to distort their responses to whatever they deemed was socially appropriate (Sigall & Page, 1971). The underlying principles of the bogus pipeline's mild

deception have become a classic technique that is still employed in current research on an array of topics (see, e.g., Myers & Zeigler-Hill, 2012; Nier, 2005).

This inclination for impression management that distorts the validity of self-reports is particularly likely when individuals are questioned about politically or socially sensitive topics such as interracial or intergroup behaviors (Dovidio, et al., 2009; Greenwald & Nosek, 2001; Greenwald, Poehlman, et al., 2009). As such, self-reports are generally regarded as being inadequate for capturing all aspects of individual prejudice (Tinkler, 2012), although other researchers have indicated that they may still be accurate when used in conjunction with implicit measures (Greenwald, Poehlman, et al., 2009) or when used on people who have low motivation to control their prejudiced reactions (Dunton & Fazio, 1997).

On controlling responses

The notion of what processes are automatic or controlled has received further examination following Devine's 1989 seminal work.

It is important to note that implicit and automatic are not perfect synonyms, nor are explicit and controlled. Amodio and Mendoza conceptualize automatic processes as those that are unintentional while controlled processes are intentional and often goal-oriented (D. M. Amodio & Mendoza, 2010). Defining control, they write, "Control does not relate to content *per se*, such as an explicit belief, but rather to the deliberate adjudication of an endorsed response over a different, undesired response" (D. M. Amodio & Mendoza, 2010, p. 355). They note that while automaticity and implicitness may align, ultimately whether responses are automatic or controlled is distinct from the implicit or explicit nature of a response (D. M. Amodio & Mendoza, 2010).

Providing support for the value of automatic associations, Reskin writes of their "survival value," noting that it is impossible for individuals to consciously process all of the stimuli around us, thus automatic associations release cognitive resources for other uses (Reskin, 2005, p. 34).

Racial attitude work by Fazio and colleagues led to the distinction of three types of individuals that differ due to what processes are automatically activated in them and how they then do or do not counter or control those evaluations (Fazio, et al., 1995). One group is comprised of individuals who are non-prejudiced; in Fazio's work these are the folks who do not experience the activation of negative attitudes toward Black people. A second grouping captures those who are truly prejudiced, meaning those who experience a negative automatic association and do nothing to negate or control expression of that association. Finally, a third grouping involves those who may experience a negative automatic association but, like those in Devine 1989, are motivated to counter that sentiment.

Building on this work, researchers have identified several factors that influence individuals' abilities to control responses or act in a manner that is consistent with one's explicit position. These factors include:

The role of motivation

> Individual motivation is a commonly cited factor in the controlling responses literature. For example, Dunton and Fazio studied the role of motivation on how people differ in the extent to which they seek to control expressions of prejudice. They found that motivation to control prejudiced reactions stemmed from two sources: one being a concern with acting prejudiced, and the other being a desire to avoid dispute or confrontation regarding one's thoughts or positions (Dunton & Fazio, 1997). Dunton and Fazio also concluded that self-reports of racial attitudes can be reasonably accurate for individuals with low motivation to control prejudiced reactions (Dunton & Fazio, 1997).

> Work by Devine et al. found that the implicit and explicit racial biases displayed by participants were a function of their internal and external motivation to respond in a non-prejudiced manner, with explicit biases moderated by internal motivation and implicit biases moderated by the interaction of both internal and external motivation (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002).

Relatedly, researchers have explored how motivation can lead to people "over-correcting" for their biases. Work such as that by Olson and Fazio found that some participants sought to control the racial attitudes they exhibited, and in so doing, over-corrected to display a exaggeratedly positive or negative response that concealed their true attitudes (Olson & Fazio, 2004). Earlier work by Dunton and Fazio concluded that prejudiced people who are highly motivated to control their prejudiced reactions may overcompensate for their automatically activated negativity (Dunton & Fazio, 1997). Regarding the extent of this overcorrection, they found that individuals' "high motivation to control prejudiced reactions led to the expression of judgments that were more positive than the responses of individuals for whom positivity was activated" (Dunton & Fazio, 1997, p. 324). Additional research found this same trend toward motivation moderating bias overcorrection (Towles-Schwen, 2003). This phenomena of overcorrection aligns closely with Wegener and Petty's Flexible Correction Model (see Wegener & Petty, 1995).

Also attesting to the significance of the role of motivation, multiple researchers have developed scales designed to measure the effect of motivation on controlling prejudiced reactions. Dunton and Fazio developed a Motivation to Control Prejudiced Reactions Scale that aimed to measure individual differences in motivation (Dunton & Fazio, 1997). Similarly, Plant and Devine developed two measures for assessing motivations to respond without prejudice, the Internal Motivation to Respond Without Prejudice Scale (IMS) and the External Motivation to Respond Without Prejudice Scale (EMS) (Plant & Devine, 1998).

The role of time

Another widely regarded factor that influences decision-making and whether individuals are able to control their reactions is time (Kruglanski & Freund, 1983; Payne, 2006; Sanbonmatsu & Fazio, 1990). Time pressures have been shown to be a condition in which implicit attitudes may appear (Bertrand,

Chugh, & Mullainathan, 2005), even despite explicit attempts at control.

The role of cognitive "busyness"

In the words of Reskin, "anything that taxes our attention – multiple demands, complex tasks, time pressures – increases the likelihood of our stereotyping" (Reskin, 2005, p. 34). Gilbert and Hixon studied how cognitive busyness can affect the activation and application of stereotypes. They found that cognitive busyness can decrease the likelihood of stereotype activation; however, should the stereotype be activated, cognitive busyness makes it more likely that that stereotype will be applied to the individual(s) in question (Gilbert & Hixon, 1991). Similarly, in an experimental design, Payne found that the group that was cognitively overloaded showed more bias, which he regards as a byproduct of individuals' reduced level of control over their responses (Payne, 2006). Finally, Bertrand et al. cite three conditions that are conducive to the rise of implicit attitudes, including lack of attention being paid to a task, time constraints or cognitive load, and ambiguity (Bertrand, et al., 2005).

Monitoring verbal behaviors but not nonverbals (known as leakages)

While people can monitor their verbal behaviors pretty well, they do not monitor and control their nonverbal behaviors as effectively or as often; the prejudiced attitudes they are trying to conceal can "leak," thereby revealing their true stances (Dovidio, et al., 1997; Fazio, et al., 1995; Olson & Fazio, 2007; Stone & Moskowitz, 2011).

CHAPTER 4

Measuring Implicit Cognition

Measuring Implicit Cognition

While implicit measures are less likely to be tainted by impression management tactics, measuring implicit biases remains a challenging task. Legal scholar Jerry Kang articulates the challenges of learning about people's implicit biases as a two part ("willing and able") problem (Kang, 2009, p. 2). As noted above, some people are unwilling to share their true feelings with researchers in order to maintain a sense of political correctness.

In terms of the "able" part of Kang's assertion, the challenges of assessing implicit biases are compounded by the fact that some people may be unable to share their implicit biases. Broadly speaking, we are weak at introspection and therefore often unaware of our own biases (Greenwald, et al., 2002; Kang, 2005; Nisbett & Wilson, 1977; Nosek, Greenwald, et al., 2007; Nosek & Riskind, 2012; Wilson & Dunn, 2004). Explicit measures such as self-reports can only reflect what we believe to be true about ourselves, which may be an incomplete assessment (Rudman, 2004).

Similarly, Nosek and colleagues provide a three-part assessment of why implicit measures (in this case, the Implicit Association Test, discussed later in this chapter) and self-reports differ, including:

1. "the individual is unaware of the implicitly measured associations and uses introspection to generate a unique explicit response;
2. the individual is aware of the implicitly measured associations, but genuinely rejects them as not conforming to his or her belief system and so reports a distinct explicit response; or
3. the individual is aware of the implicit associations, but chooses to report an alternative explicit response due to social concern about the acceptability of such a response." (Nosek, Greenwald, et al., 2007, p. 282)

As the credibility of bias self-reports were increasingly questioned, calls came for greater use of indirect measures and unobtrusive measures that accurately captured racial attitudes, as implicit measures are regarded as less susceptible to social desirability concerns (D. M. Amodio & Mendoza, 2010; Dunton & Fazio, 1997; Greenwald & Banaji, 1995; Jones & Sigall, 1971; Nosek & Banaji, 2002; Petty, et al., 2009; von Hippel, et al., 1997). Since that time, researchers have developed numer-

ous instruments and techniques designed to measure cognitions implicitly. Because implicit measures do not ask research subjects directly for a verbal report, in many cases the subjects may not even be aware of what constructs are being assessed and measured, thereby minimizing social desirability concerns (Fazio & Olson, 2003). This chapter provides a brief overview of several of these instruments.

Physiological approaches

Physiological instruments assess bodily and neurological reactions to stimuli. These instruments provide insights into implicit biases because they measure reactions that are not easily controlled, and the individuals involved may not even realize that they are reacting in any manner whatsoever.

In one study, Phelps et al. (2000) used functional Magnetic Resonance Imaging (fMRI) to examine unconscious evaluations of Blacks and Whites. The amygdala was of particular interest because of its known role in race-related mental processes (Hart, et al., 2000). It is also the part of the brain that reacts to fear and threat (Davis & Whalen, 2001; Pichon, Gelder, & Grèzes, 2009; Whalen, et al., 2001). Phelps and colleagues found that White subjects generally showed greater amygdala activation when exposed to unfamiliar Black faces compared to unfamiliar White faces; however, the fMRI data lacked any consistent patterns of amygdala activity when the subjects were viewing well-known Black and White faces (Phelps, et al., 2000). The researchers also examined the association between the strength of amygdala activation and other measures of race evaluation. They found that the amygdala activity for unfamiliar faces correlated with two different unconscious measures of race evaluation but not with explicit measures of racial bias. Overall, their research indicated that the amygdala response of White subjects to Black faces versus White faces is a byproduct of cultural evaluations that have been modified by individuals' experiences (Phelps, et al., 2000).

Similarly, Cunningham et al. compared a measure of implicit racial bias to amygdala activity and found a significant correlation between the two (Cunningham, et al., 2004). In short, higher levels of anti-Black implicit bias were associated with greater amygdala activity, as measured by fMRI (Cunningham, et al., 2004).

These conclusions align with that of Lieberman et al. (2005) wherein researchers found that the amygdalas of both African American and Caucasian participants displayed greater levels of activity when viewing African American faces than Caucasian American faces (Lieberman, Hariri, Jarcho, Eisenberger, & Bookheimer, 2005). Lieberman et al. suggest that the amygdala activity recorded in this study reflect "culturally learned negative associations regarding African-American individuals" (Lieberman, et al., 2005, 722). Additional research expanded on Phelps et al. (2000) and Lieberman et al.'s (2005) studies by finding that beyond race, skin tone variations also affect amygdala activation, with darker skin tones provoking more amygdala activity than lighter tones (Ronquillo, et al., 2007).

Further research on the neural basis of implicit biases articulated a three-part model in which the amygdala is only one component. Stanley and colleagues' model asserted that the anterior cingulate is involved in the detection of implicit attitudes, and the dorsolateral prefrontal cortices help regulate implicit biases (Stanley, Phelps, & Banaji, 2008). This conclusion aligns with earlier work by Cunningham et al. that suggested that the controlled processes that originate in the anterior cingulated and dorsolateral prefrontal cortex can override the automatic processing of Black and White faces that occurs in the amygdala (Cunningham, et al., 2004).

Vanman et al. (2004) explored another physiological approach to measuring implicit prejudices. They used facial electromyography (EMG) to examine the micro-movements of muscles used in smiling and frowning while exposing research participants to images of White and Black faces and ultimately concluded that facial EMG can be used as an implicit measure of racial prejudice related to discrimination (Vanman, Saltz, Nathan, & Warren, 2004).

Finally, work by Bascovich and colleagues employed cardiovascular and hemodynamic measures as a means to understanding how study participants responded to stigmatized individuals (in this case, individuals perceived to have port-wine birthmarks on their faces) (Blascovich, Mendes, Hunter, Lickel, & Kowai-Bell, 2001).

Priming methods

In psychological terms, priming is simply the act of being exposed to a stimulus that influences how an individual later responds to a different stimulus. Often used in experimental settings, priming methods typically feature a subliminal initial prime that influences or increases the sensitivity of the respondent's later judgments or behaviors (Tinkler, 2012). Implicit racial bias research often uses a race-related term or image as the initial stimulus, followed by the measurement of a later stimulus that compares responses by participants who were shown a race-related prime versus those who were not. For example, Philip Goff and colleagues conducted a study in which they primed some participants with Black faces, White faces, or a non-facial control image. Participants were then presented degraded images of animals that gradually came into focus, making the animal incrementally easier to identify.

As hypothesized, the research team found that participants who were primed with Black male faces required fewer image frames to identify drawings of apes compared to those primed by White male faces or when not primed at all, yet participants remained unaware that they had ever been primed in any fashion (Goff, Eberhardt, Williams, & Jackson, 2008). These kinds of priming methods are thought to yield insights on associations that, despite their implicit nature, influence individuals' perceptions.

Response latency measures

Response latency measures represent "the most widely used strategies to assess implicit prejudice" (Dovidio, et al., 2009, p. 170). These measures rely on reaction times to specific tasks to uncover individuals' biases (Rudman, 2004). The quick speed of reply assumes that responses are uncontaminated and reflected "true" content of stereotypes (D. Amodio & Devine, 2009). The underlying premise of these reaction-time studies is that individuals are able to complete cognitively simple tasks relatively more quickly than those that are mentally challenging (Kang & Lane, 2010).

Thus, measuring the difference in response latency times can provide insights into how strongly two concepts are associated. While the premise may not sound particularly sophisticated, reaction-time instruments Kang and Lane assert that these are

the most reliable measures of implicit cognitions (Kang & Lane, 2010).

Implicit Association Test

One of the most popular, sophisticated, and recognizable response latency measures is the Implicit Association Test (IAT). Pioneered by Anthony Greenwald and colleagues, the IAT measures the relative strength of associations between pairs of concepts (Greenwald, et al., 1998; Greenwald & Nosek, 2001). These associations addressed by the IAT include "attitudes (concept-valence associations), stereotypes (group-trait associations), self-concepts or identities (self-trait or self-group associations), and self-esteem (self-valence associations)" (Greenwald, Poehlman, et al., 2009, p. 19). The IAT operates "on the assumption that if an attitude object evokes a particular evaluation (positive or negative), it will facilitate responses to other evaluatively congruent and co-occurring stimuli" (Dasgupta & Greenwald, 2001, p. 801). Stated another way, the IAT asks respondents to sort concepts and measures any time differences between schema-consistent pairs and schema-inconsistent pairs (Kang, et al., 2012). As a response latency measure, the IAT operates on the supposition that when the two concepts are highly associated, the sorting task will be easier and thus require less time than when the concepts are not as highly associated (Greenwald & Nosek, 2001; Reskin, 2005). This difference in mean response latency is known as the IAT effect (D. M. Amodio & Mendoza, 2010; Greenwald, et al., 1998). The time differentials of the IAT effect have been found to be statistically significant and not simply due to random chance (Kang, 2009). The IAT effect reveals the role of both automatic and controlled processing: the strength of an automatic association and the challenge associated with sorting a bias-inconsistent pair (D. M. Amodio & Mendoza, 2010). Some studies have found IAT results to be generally stable over time (Cunningham, et al., 2001; Egloff, Schwerdtfeger, & Schmukle, 2005).

One notable benefit to the IAT is that is relatively resistant to social desirability concerns (Greenwald, et al., 1998). Numerous studies have examined whether individuals can "fake out" the IAT by intentionally controlling their responses in such a way as to produce desired rather than authentic results. Outcomes and musings from these studies are varied (see, e.g., Cvencek, Greenwald, Brown, Gray, & Snowden,

2010; Egloff & Schmukle, 2002; Fiedler & Bluemke, 2005; Fiedler, Messner, & Bluemke, 2006; Kim, 2003; Steffens, 2004). That said, studies that compare individuals' explicit responses with their implicit attitudes, as measured by the IAT ,consistently find that the people's implicit biases are actually higher than what they self-report (Nosek, et al., 2002; Sabin, Nosek, Greenwald, & Rivara, 2009). This finding aligns well with the social desirability/impression management literature discussed in the previous chapter.

Since its conception, researchers have subjected the IAT to numerous and rigorous tests of its reliability (see, e.g. Bosson, William B. Swann, & Pennebaker, 2000; Dasgupta & Greenwald, 2001; Greenwald & Farnham, 2000; Greenwald & Nosek, 2001; Kang & Lane, 2010; Nosek, Greenwald, et al., 2007). Similarly, the validity of the IAT has been examined extensively (for overviews and meta-analyses, see Greenwald; Greenwald, Poehlman, et al., 2009; Jost, et al., 2009). Studies designed to probe the IAT's internal validity have been particularly extensive, with researchers examining potential confounds such as the familiarity of the stimuli (Dasgupta, Greenwald, & Banaji, 2003; Dasgupta, et al., 2000), the order of the tasks (Greenwald, et al., 1998; Nosek, Greenwald, & Banaji, 2005), previous experience with the IAT (Dasgupta, et al., 2000; Greenwald, et al., 2003), and various procedural nuances (Greenwald, et al., 1998; Greenwald & Nosek, 2001), among others.

Of particular interest to many researchers is the question of the IAT's predictive validity (see, e.g., Blanton, et al., 2009; Egloff & Schmukle, 2002; Fazio & Olson, 2003; Greenwald & Krieger, 2006; Greenwald, Poehlman, et al., 2009; McConnell & Liebold, 2001). Is the IAT able to accurately predict attitudes and behaviors, and if so, can it do so better than explicit self-reports? Greenwald and colleagues' 2009 meta-analysis addressed the predictive validity of the IAT in 122 research reports. Overall they found a predictive validity of r = 0.274 for the IAT, which is regarded as moderate (Greenwald, Poehlman, et al., 2009). Explicit measures were also effective predictors (r = 0.361); however, the predictive validity of these explicit self-report measures declined dramatically when the topic was socially sensitive. Thus, for topics such as interracial and intergroup behavior, the IAT held greater predictive validity than the self-reports did, which justifies the IAT's use, particularly when in combination with self-report measures (Greenwald, Poehlman, et al., 2009).

As a whole, the IAT has performed extremely well to this scrutiny of reliability and validity, so much so that a 2010 article concluded, "After a decade of research, we believe that the IAT has demonstrated enough reliability and validity that total denial is implausible" (Kang & Lane, 2010, p. 477).

IAT Findings on Race

One of the most well-known versions of the IAT is the Black/White IAT, which examines the speed with which individuals categorize White and Black faces with positive and negative words. Faster reaction times when pairing White faces with positive words and Black faces with negative terms suggests the presence of implicit pro-White/anti-Black bias. Considerable research has indicated that most Americans, regardless of race, display a pro-White/anti-Black bias on this IAT (Dovidio, et al., 2002; Greenwald, et al., 1998; Greenwald, Poehlman, et al., 2009; McConnell & Liebold, 2001; Nosek, et al., 2002), even in children as young as six years old (A. S. Baron & Banaji, 2006).

The documented presence of pro-White bias even among nonwhites has intrigued researchers that study ingroup/outgroup dynamics. Dasgupta sheds light on the internal conflict that may help explain this unusual finding when she writes, "In the case of individuals who belong to disadvantaged or subordinate groups ... the desire to protect self-esteem should lead to ingroup favoritism and outgroup bias, but the desire to maintain current social arrangements leads to predictions of outgroup favoritism" (Dasgupta, 2004, p. 148). This leads Dasgupta to question whether there are two separate sources of implicit attitudes – one that focuses on one's group membership, and another that seeks to maintain current social hierarchies (Dasgupta, 2004). Several studies lean towards the latter explanation, citing the presence of implicit outgroup favoritism (or, in some cases, less ingroup favoritism) for a dominant outgroup over one's own subordinated ingroup (Ashburn-Nardo, Knowles, & Monteith, 2003; Nosek, et al., 2002; Rudman, Feinberg, & Fairchild, 2002).

Implicit Biases and the Effects on Behavior

Regardless of how they are measured, researchers agree that implicit biases have

real-world effects on behavior (Dasgupta, 2004; Kang, et al., 2012; Rooth, 2007). These effects have been shown to manifest themselves in several different forms, including interpersonal interactions. For example, McConnell and Liebold found that as White participants' IAT scores reflected relatively more positive attitudes towards Whites than Blacks; social interactions (measured by focusing on 13 specific behaviors) with a White experimenter were more positive than interactions with a Black experimenter. In this study, larger IAT effect scores "predicted greater speaking time, more smiling, more extemporaneous social comments, fewer speech errors, and few speech hesitations in interactions with the White (vs Black) experimenter" (McConnell & Liebold, 2001, p. 439). Another study by Dovidio et al. found that White individuals with higher levels of racial bias blinked more and maintained less visual contact with Black interviewers than White ones (Dovidio, et al., 1997).

Several studies look at interracial interactions and behaviors with a focus on friendliness to examine how implicit biases can affect behavior (Dasgupta, 2004; Dovidio, et al., 2002; Fazio, et al., 1995; Sekaquaptewa, Espinoza, Thompson, Vargas, & Hippel, 2003). Perceptions of friendliness are often but not necessarily entirely assessed through nonverbal body language, such as having an open vs. closed posture or degree of eye contact maintained. These behaviors are insightful because individuals are often relatively unaware of such actions and thereby unlikely to attempt to control or correct these behaviors (Dasgupta, 2004). In one study, Dovidio, Kawakami, and Gaertner established that the implicit biases of White participants significantly predicted the degree of nonverbal friendliness they displayed towards their Black partners in an experimental setting (Dovidio, et al., 2002). This result echoes earlier work by Fazio et al. that found that White students who possessed more negative attitudes towards Blacks were less friendly and less interested during their interactions with a Black experimenter (Fazio, et al., 1995).

Having established that implicit biases affect individuals' behaviors, the next logical step is to consider the ramifications of those behaviors. Indeed, implicit biases have a tremendous impact on numerous social situations. In the words of Rudman, "biases that we do not acknowledge but that persist, unchallenged, in the recesses of our minds, undoubtedly shape our society" (Rudman, 2004, p. 130). The next three chapters examine how implicit racial biases affect three specific realms: education, criminal justice, and health/health care.

CHAPTER FIVE

Implicit Bias In Education

Implicit bias in Education

Implicit bias can permeate educational settings in several forms, all of which can yield disadvantageous consequences for students of color. Teacher expectations of student achievement, teacher perceptions of student behavior, and students' self-perceptions are three key themes highlighted in the literature.

Teacher Expectations of Student Achievement

Teacher expectations related to student achievement is one area in which implicit biases can have detrimental effects. A 2010 study by van den Bergh et al. sought to determine whether teachers' expectations for students and the ethnic achievement gaps found in their classrooms were related to the teachers' prejudiced attitudes. Conducted in the Netherlands, researchers assessed the prejudiced attitudes of elementary school teachers using self-reports and results from Implicit Association Tests. Results indicated that "teachers generally hold differential expectations of students from different ethnic origins" and that implicit prejudiced attitudes were responsible for these differential expectations as well as the ethnic achievement gap in their classrooms (van den Bergh, Denessen, Hornstra, Voeten, & Holland, 2010). van den Bergh et al. assert that teachers who hold negative prejudiced attitudes "appeared more predisposed to evaluate their ethnic minority students as being less intelligent and having less promising prospects for their school careers" (van den Bergh, et al., 2010, p. 518).

Indeed, many studies have shown that teacher expectations tend to vary based on the demographic characteristics of their students. Tenenbaum and Ruck (2007) performed a meta-analysis to determine whether teachers' expectations, referrals (i.e., recommendations for special education, disciplinary action, or gifted programs), and speech patterns (i.e., positive, neutral, and negative speech) differ toward European American students as opposed to African American, Asian, or Latino/a students. They found statistically significant results that teachers hold lower expectations for African American and Latino/a children compared to European American children, and, per the Pygmalion Effect, these expectations may affect student academic performance (Rosenthal & Jacobson, 1968; Tenenbaum & Ruck, 2007). The results of this study align with previous meta-analyses (R. M. Baron, Tom, & Cooper, 1985; Dusek & Joseph, 1983).

McKown and Weinstein (2002) affirmed the role of teacher expectancy effects on the achievement of African American students. Using a sample of 561 elementary school children, the researchers examined whether students' ethnicity played a role in their susceptibility to teacher expectancy effects. By conceptualizing teacher expectations as the degree to which teachers over- or under-estimated achievement compared to the students' actual academic performance, McKown and Weinstein found that African American children are more likely than Caucasian children "to confirm teacher underestimates of ability and less likely to benefit from teacher overestimates of ability" (McKown & Weinstein, 2002, p. 176). Thus, implicit biases held by teachers, that affect the expectations they hold for students, have real consequences in the classroom for African Americans.

Students' verbal nuances and vernacular dialects can also arouse implicit biases in teachers. Following an assertion by Christian (1997) that "people who hear a vernacular dialect make erroneous assumptions about the speaker's intelligence, motivation, and even morality," Cross et al. (2001) asked prospective teachers to blindly evaluate the personal characteristics of anonymous speakers reading a neutral and minimally difficult passage, judging the speakers' intelligence, personality, social status, and ambition (Christian, 1997, p. 43; Cross, DeVaney, & Jones, 2001). The prospective teachers were found to draw conclusions about these traits based on perceptions of dialect, and ethnicity was one factor that influenced these judgments; White prospective teachers regarded White speakers as most favorable and Black speakers as least favorable (Cross, et al., 2001). With White women comprising the majority of our nation's teachers (84% of public school teachers in 2011 were White), ostensibly trivial issues such as dialects can implicitly affect teachers' preconceptions of students (Feistritzer, 2011). Cross et al. warns that biases stemming from verbal nuances can snowball, as "teachers' preconceptions of students may be reflected in students' grades and impact their self-perception, beginning a cycle of self-fulfilling prophecy that contributes to the eventual academic failure of speakers of nonstandard dialect" (Cross, et al., 2001, p. 223).

Finally, Ferguson notes that the oft-reported Black-White test score gap can influence teachers to develop stereotypical perceptions, expectations, and behaviors that most likely perpetuate this gap in test scores (R. F. Ferguson, 2003). For instance,

teachers may be less likely to help and support Black children than White children because they may underestimate Blacks' intellectual potential. If Black students are taught by educators who do not believe in their full potential, these subtle biases can accumulate over time to create a substandard educational environment that fails to prepare at-risk students to become fully contributing members of society.

In short, teacher expectations of student achievement may be swayed by implicit biases, and the manifestations of these biases can have lasting effects on students and serve as self-fulfilling prophecies (for more on self-fulfilling prophecies, see Merton, 1957). Holding lower standards for nonwhite students is particularly disheartening in light of studies that find that holding students to higher standards benefits students and actually improves test scores (Betts & Grogger, 2003; Figlio & Lucas, 2004).

Teacher Perceptions of Student Behavior

Teachers' perceptions of specific minority student behaviors can also energize implicit bias. For example, Neal et al. (2003) found that students who displayed a Black walking style (i.e., "deliberately swaggered or bent posture, with the head slightly tilted to the side, one foot dragging, and an exaggerated knee bend") were perceived by teachers as lower in academic achievement, highly aggressive, and more likely to be in need of special education services (Neal, McCray, Webb-Johnson, & Bridgest, 2003).

Researchers also found that Whites with relatively high levels of implicit racial bias perceived Blacks to be more threatening than Whites. In one study, Hugenberg and Bodenhausen (2003) explored whether implicit biases affected how subjects perceived the facial emotions displayed by others. The subjects were shown a series of faces (one series of Black and one series of White) that progressed from a scowl to a smile and asked what face in the series indicated an offset/onset of anger. "Higher implicit … [bias] was associated with a greater readiness to perceive anger in Black faces, but neither explicit nor implicit prejudice predicted anger perceptions regarding similar White faces" (Hugenberg & Bodenhausen, 2003, p. 640). These findings suggest that White teachers may incorrectly perceive Black students as angry or ag-

gressive, which could deter them from reaching out to assist these students or cause them to mislabel Black students as deviant.

Teachers' implicit biases may be further amplified by a cultural mismatch that exists between White teachers and their students of color, and this mismatch can lead to teachers misinterpreting student behavior (A. A. Ferguson, 2000). For example, Weinstein et al. (2004) recounts an anecdote wherein a European American teacher observed a lively debate occurring between African American males, and by interpreting the interaction as aggressive and contentious rather than simply verbal sparring common among African American teenagers, took the teens to the principal's office for a reprimand (Weinstein, Tomlinson-Clarke, & Curran, 2004). Often these cultural mismatch scenarios, with implicit biases fueled by negative portrayals of Black youth in the media, result in unnecessary and unequal disciplinary interventions for students of color.[1] Similarly, in their book, *The Cool Pose: The Dilemmas of Black Manhood in America*, Majors and Mancini Billson explore how many Black males exhibit a distinct persona in an effort to assert their masculine identity. They write, "Cool pose is a ritualized form of masculinity that entails behaviors, scripts, physical posturing, impression management, and carefully crafted performances that deliver a single, critical message: pride, strength, and control" (Majors & Billson, 1992, p. 4). Teachers who are not knowledgeable about Black cultural cues may misread this portrayal of masculinity as confrontational or defiant.

Students' Self-Perceptions

Stereotypes can also emerge implicitly and affect students through a subtle mechanism known as stereotype threat. Attributed to social psychologist Claude Steele, stereotype threat refers a fear of being viewed through the lens of a negative stereotype, or the fear of inadvertently confirming an existing negative stereotype of a group with which one self-identifies (Steele & Aronson, 1995). Studies have

---

1. March 2012 data released by the U.S. Department of Education reported that Black students, boys in particular, face harsher discipline in public schools. Black students were 3.5 times more likely to be suspended or expelled compared to white students in the 2009–2010 school year. (For more information, see Lewin, 2012.)

shown that these fears often manifest themselves in lower performance by the stereotyped group, even when the stereotyped group and non-stereotyped group being compared have been statistically matched in ability level (Steele & Aronson, 1995). For example, if a teacher tells students who are about to take a test that Asian students generally score higher than Whites on this test, then Whites tend to perform significantly worse than if they had not been primed to think of themselves as less capable than Asians. Moreover, stereotype threat has also been shown to impair test-taking efficiency, as individuals feeling stereotype threat complete fewer test problems and with a lower level of accuracy than those in a control group (Steele & Aronson, 1995). In short, the fear of reinforcing a negative stereotype can implicitly provoke a "disruptive apprehension" that interferes with performance (R. F. Ferguson, 2003; Steele, 1997).

Culturally Appropriate Curriculum

Implicitly biased teachers may also unknowingly use curriculum that is not culturally responsive to all members of their classroom community. This can be detrimental to students, as failing to modify curriculum in a manner that will facilitate the academic achievement of students from diverse racial, ethnic, and cultural groups creates an unequal pedagogy. Banks writes about how teachers' cultural competency and willingness to develop their students' cultural and linguistic strengths can increase academic achievement for students of color (Banks, 1995). Teachers with high levels of implicit bias may fail to make these adjustments because, as noted previously, they may subconsciously view minority cultures and linguistic styles as a sign of low academic ability and aggression rather than simply distinct cultures and linguistic styles that differ from their own. This not only reduces the academic potential of minority students but can also lead to the over-identification of students for special education and disciplinary action (for more information generally, see Arnold & Lassmann, 2003; Sherwin & Schmidt, 2003; Skiba, Michael, Nardo, & Peterson, 2000).

Implicit Bias In Criminal Justice

Implicit bias in Criminal Justice

> *"Ultimately, the decision to sentence a person to death is based not on a rational determination, a weighing of the evidence, or the finding that the particular defendant is indeed guilty of the worst of the worst offenses, but rather on a series of unconscious decisions, by prosecutors, judges, jurors, and even defense lawyers in which race affects the processing of the underlying evidence and tilts the scales of justice.*
>
> *– Smith & Cohen, 2012, p. 233 –*

Implicit bias can surface in the criminal justice system in a variety of fashions, all of which may potentially taint the prospect of fair outcomes. Before delving into the different facets of the system in which researchers have identified implicit bias in action, it is important to note that even small implicit biases can accumulate over the course of legal proceedings, thereby amplifying the effect. In the words of Jerry Kang et al., "For a single defendant, these biases may surface for various decision makers repeatedly in policing, charging, bail, plea bargaining, pretrial motions, evidentiary motions, witness credibility, lawyer persuasiveness, guilt determination, sentencing recommendations, sentencing itself, appeal, and so on. Even small biases at each stage may aggregate into a substantial effect" (Kang, et al., 2012, p. 1151).

Outside the Courtroom

Bias in Police Officers

Like all other populations, police officers are not immune to implicit bias. Eberhardt et al. studied police officers and found that when directly asked "who looks criminal?", they chose Black faces over White ones, particularly those that were more stereotypically Black (Eberhardt, Goff, Purdie, & Davies, 2004). Moreover, automatic implicit biases can cause officers to misinterpret Blacks' behavior as suspicious or aggressive, even if the actions are neutral in nature (Richardson, 2011).

Graham and Lowery examined whether police officers' unconscious racial stereotypes affected how they perceive and treat juvenile offenders. As predicted, the group of officers who were race-primed for the category Black judged the offenders

as more blameworthy and meriting harsher sanctions than the officers who were exposed to a neutral prime, and consciously held beliefs about race were not found to moderate the effects of the racial primes (Graham & Lowery, 2004). These findings have significant implications for juvenile offenders, as interactions with the criminal justice system at a youthful age can have lasting life effects.

On a more heartening note, while work by Plant and Peruche documented the existence of bias in police officers, they also concluded that racial biases are not inevitable and may be overcome (Plant & Peruche, 2005). In another study, they found that police officers who have positive intergroup contact with Black people in their personal lives outside of the workplace held more positive attitudes towards Black people and more positive beliefs with respect to the criminality and violence of Black suspects (Peruche & Plant, 2006). These findings foreshadow a larger discussion on debiasing found in chapter 7 of this document.

Shooter / Weapons Bias

Shooter/weapons bias is a well-documented phenomenon. Shooter bias refers to the strong and pervasive implicit association that exists between Blackness (as opposed to European Americans) and weapons (vs. harmless items). Some studies in this realm rely on priming, such as B. Keith Payne's work that showed that study participants identified guns more quickly than hand tools when primed with Black faces versus White faces, and they also misidentified tools as guns more frequently when a Black prime was employed (Payne, 2001). These insights are helpful for beginning to understand the shooter / weapons bias phenomenon.

One well-known study that examined implicit bias in this context was conducted by Joshua Correll and colleagues. For this four-part study, the research team constructed a videogame that allowed them to measure the effect of ethnicity on making the decision whether or not to shoot. Simulating what police officers experience, participants were instructed to "shoot" when armed individuals appeared on the screen and to "not shoot" unarmed targets. Images that then flashed on the screen displayed African American and White individuals shown in front of complex backgrounds, all of whom were holding some kind of object. In some cases the objects

were guns; however, in others they were innocuous objects such as a camera, a can, a cell phone, or a wallet. Participants were told to make the decision whether or not to shoot as quickly as possible. Correll et al. hypothesized that stereotypes that associate African Americans with violence may provoke participants to "respond with greater speed and accuracy to stereotype-consistent targets (armed African Americans and unarmed Whites) than to stereotype-inconsistent targets (armed Whites and unarmed African Americans)" (Correll, Park, Judd, & Wittenbrink, 2002, p. 1325). The first study supported this hypothesis, as participants "shot" at armed individuals more quickly when they were African American as opposed to White, and made the decision to refrain from shooting unarmed targets more quickly when they were White as opposed to African Americans. The second study required participants to make "shoot or not" decisions within even shorter time constraints, and this led to the finding that when participants neglected to shoot an armed target, the individual on the screen tended to be White rather than African American. And, in circumstances where the videogame character was displaying a harmless object, participants were more likely to mistakenly shoot African Americans versus Whites. Part three of the study considered the effects of prejudice and personal endorsement of the violent stereotype about African Americans; however, these did not predict shooter bias. In short, it seems that the shooter bias observed was a byproduct of knowledge of the cultural stereotype rather than personal endorsement thereof. Indeed, in study four, researchers found that African American and White participants displayed similar levels of bias.

Following up on this study five years later, Correll and colleagues compared actual police officers to community members to assess difference in the speed and accuracy of the "shoot or not" decisions described in their previous study. They found that officers' response times were quicker than civilians, and that officers made fewer mistakes differentiating armed from unarmed individuals (Correll, et al., 2007). While all participants displayed some degree of racial bias, police officers displayed less bias in their final "to shoot or not" decisions (Correll, et al., 2007).

It is important to emphasize, however, that weapons bias is generally regarded as occurring independent of intent, meaning that even someone who is consciously and explicitly committed to being fair and unbiased may display this bias anyways

(Kang, et al., 2012; Payne, 2006). Shooter bias is largely driven by stereotypical associations rather than blatant racial hostility (Payne, 2006). "In the policing context, implicit stereotypes can cause an officer who harbors no conscious racial animosity and who rejects using race as a proxy for criminality to unintentionally treat individuals differently based solely upon their physical appearance"(Richardson, 2011, p. 2039).

The measurement of schema consistent and inconsistent pairings, such as on the IAT, is significant to note here, as police officers may only have fractions of a second in which to decide whether to shoot (Kang, et al., 2012; Payne, 2006). Thus, the implications of this bias can be literally mean life or death (Kang, et al., 2012; Plant & Peruche, 2005; Plant, Peruche, & Butz, 2005).

Within the Courtroom

Numerous dynamics and actors within the courtroom can activate or fall susceptible to implicit biases.

Judges

Even with an avowed commitment to impartiality, judges, like the rest of the general population, display implicit biases. In one study, researchers administered the race IAT to both Black and White judges. White judges displayed a strong White preference, as this implicit bias was revealed in 87.1% of the 85 White judges studied (Rachlinski, Johnson, Wistrich, & Guthrie, 2009). In contrast, Black judges generally displayed a no clear preference (Rachlinski, et al., 2009).

While many variables may factor into this finding, one common theme in the literature is the "illusion of objectivity," which contributes to judges feeling overly confident in their ability to remain unbiased. Rachlinski et al. asked a group of judges to rank their own ability to "avoid racial prejudice in decision-making," and a stunning 97% of them placed themselves in the top half, with 50% ranking themselves in the top quartile (Rachlinski, et al., 2009, p. 1125). These unusually high percentages, that are mathematical impossibilities, were not isolated to just one study. Anoth-

er study asked administrative law judges about their capacity for avoiding bias in judging, and the numbers were extraordinarily similar. Over 97% of the judges in this study placed themselves in the top half, with 50% ranking their ability to avoid biased judging within the top quartile (Guthrie, Rachlinski, & Wistrich, 2009). Other related studies have noted how "blind spot bias" causes people to see biases in other people more clearly than in themselves (Pronin, 2006; Pronin, Lin, & Ross, 2002). These findings highlight the extent to which the illusion of objectivity can affect how judges perceive themselves and their decision making.

The illusion of objectivity not only affects judges' self-perception, it also colors their judgment. An experiment by Ulhmann and Cohen found that "when people feel that they are objective, rational actors, they act on their group-based biases more rather than less" (Uhlmann & Cohen, 2007, p. 221). Susceptible to the illusion of objectivity, judges may unintentionally act in ways that align with the implicit biases they hold. The irony that thinking of oneself as objective actually fosters more biased decisions should be a tremendous concern to judges everywhere. In the concise words of Kang et al., "believing oneself to be objective is a prime threat to objectivity" (Kang, et al., 2012, p. 1184).

Jurors

Generally speaking, research has indicated that jurors tend to show biases against defendants of a different race (Kang, et al., 2012). One meta-analysis of laboratory studies by Mitchell et al. found this ingroup bias affected both verdicts and sentencing, albeit with relatively small effect sizes (T. L. Mitchell, Haw, Pfeifer, & Meissner, 2005).

An empirical study by Levinson and Young explored how mock-jurors assess trial evidence. Introducing their Biased Evidence Hypothesis, the researchers built upon the literature on priming by asserting that the activation of racial stereotypes prompt jurors to "automatically and unintentionally evaluate ambiguous trial evidence in racially biased ways" (Levinson & Young, 2010, p. 309). Research participants were shown a slideshow of evidence from a fictitious armed robbery in which half of the participants viewed a dark-skinned perpetuator and the other half saw a light-

er-skinned perpetuator. Supporting the Biased Evidence Hypothesis, the results showed that, compared to the White-skinned perpetuator, participants who viewed the darker skinned perpetuator were more likely to consider the evidence as indicative of criminal guilt and more likely to assert that the defendant was in fact guilty of committing the armed robbery (Levinson & Young, 2010).

An empirical study by Levinson examined whether implicit racial bias affected how mock jurors recalled legal facts from fictional story involving a fistfight. While the story itself remained consistent throughout the study, in one condition, participants read about "William," who was specifically noted to be Caucasian. Other participants experienced alternate conditions in which William was replaced by Tyronne (who was explicitly listed as African American) or Kawika, a Hawaiian. After a fifteen-minute distraction task, participants were asked to recall details of the confrontation scenario. Levinson found that the reported race of the fictional defendant affected participants' recall of the story. Notably, participants were significantly more likely to remember facts about the aggressiveness of Tyronne compared to when William or Kawika were substituted in the same role (Levinson, 2007). This finding of racially biased memories did not relate to participants' explicit racial preferences (Levinson, 2007). Levinson notes how this misremembering of facts due to implicit racial biases can have crucial consequences for the enactment of justice in a legal setting.

Further work by Levinson and his colleagues Cai and Young examined whether implicit biases affect jury decisions in ways that reflect racial bias. The researchers debuted a new IAT, the Guilty/Not Guilty IAT, which measures the implicit association between African Americans and criminal guilt (Levinson, Cai, & Young, 2010). Their study tested whether the Guilty/Not Guilty IAT predicted how mock jurors responded to unclear trial evidence and found a connection between the implicit associations they measured and jurors' views of the evidence (Levinson, et al., 2010, p. 190). Studies such as this reiterate the role of implicit bias in jury deliberations and introduce skepticism to the ideal of "innocent until proven guilty."

One very important nuance to this juror bias research must be highlighted: contrary to what may be the prevailing assumption, when a case is "racially charged," jurors show less bias because they are more thoughtful about the role of race than they are

when race is not an explicit aspect of the case (Sommers & Ellsworth, 2000; Sommers & Ellsworth, 2001).

The racial composition of a jury also has a considerable impact on legal decisions (Arterton, 2008; Bowers, Sandys, & Brewer, 2004; Sommers, 2006). Sommers studied the group decision making processes and outcomes of mock juries of various racial compositions and found that compared to an all-White jury, diverse juries deliberated longer and discussed a wider range of information from the case (Sommers, 2006). Bowers looked at the jury racial composition data from 74 capital trials and concluded that juries dominated by White males was strongly associated with death penalty sentencing decisions (Bowers, et al., 2004). Supreme court Justice Sandra Day O'Connor's dissent in Georgia v. McCollum affirms this notion that jurors, their implicit racial biases, and jury composition shape trial proceedings: "[i]t is by now clear that conscious and unconscious racism can affect the way White jurors perceive minority defendants and the facts presented at their trials, perhaps determining the verdict of guilt or innocence .... [M]inority representation on the jury may help to overcome such racial bias, for there is substantial reason to believe that the distorting influence of race is minimized on a racially mixed jury" ("Georgia v. McCollum," 1992, p. 68).

Together these studies show that the likelihood of a jury convicting a defendant involves not only the strength of the case and the credibility of its evidence, but also the everyday prejudices and implicit biases the jurors bring to the deliberations.

Sentencing

Given the courtroom dynamics discussed thus far, it is not surprising that implicit racial biases may taint the sentencing process. Considering that sentences often have profound and life-altering effects on the convicted, researchers have investigated and attempted to measure the effects of these biases. The results are startling.

Adding to research that found that people hold associations between stereotypically Black physical traits and perceived criminality (Eberhardt, et al., 2004), Blair, Judd, and Chapleau explored the connection between criminal sentencing and Af-

rocentric features bias. This form of bias refers to the generally negative judgments and beliefs that many people hold regarding individuals who possess particularly Afrocentric features, notably dark skin, a wide nose, and full lips. The presence of these features may activate associations that lead to stereotypical perceptions. Blair et al. studied young Black and White male inmates in the Florida Department of Corrections database and found that Black and White inmates with equivalent criminal records tended to receive similar sentences; however, within each race, individuals with more Afrocentric features received longer sentences than those with less Afrocentric features (Blair, Judd, & Chapleau, 2004). More specifically, being one standard deviation above the mean level of Afrocentric features for a given group equated to sentences that were seven to eight months longer than individuals one standard deviation below the group mean, even when criminal records are held constant (Blair, et al., 2004). Thus, the mere perception of an offender having more Afrocentric features, even if the offender is White, activates an implicit bias that leads to longer sentences. Pizzi, Blair, and Judd posited an explanation for this phenomenon, asserting, "it is our thesis that judges have learned to be more careful to impose similar sentences between racial groups, but they have not been similarly sensitized to the possibility of discrimination based on Afrocentric features within racial categories" (Pizzi, Blair, & Judd, 2005, p. 331).

Applying this concept to death penalty sentences, Eberhardt and colleagues studied whether the extent to which Black defendants possess stereotypically Black physical traits affects their likelihood of being sentenced to the death penalty in death-eligible cases. Using data from death-eligible cases from 1979-1999 in Philadelphia, researchers found that after controlling for numerous factors, when the victim was White, Black defendants whose appearance was more stereotypically Black were more likely to be sentenced to death than those whose faces displayed fewer stereotypically Black traits (Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006). In cases where the victim was Black, however, the perceived stereotypicality of Black defendants did not predict death sentencing.

Prosecutors

Prosecutors are as susceptible to implicit racial biases as anyone else, and the unique

nature of their job provides numerous opportunities for those biases to act within criminal justice proceedings.

A recent law review article by Smith and Levinson outlines three main areas in which prosecutors may unknowingly act upon their implicit racial biases: when making charging decisions, during pretrial strategy, and throughout trial strategy (Smith & Levinson, 2012). With respect to charging decisions, implicit biases color how offenders are perceived, thereby affecting whether a suspect should be charged, and if so, what severity of a crime should be charged. These choices can have tremendous impacts on suspects' lives, particularly when prosecutors are deciding between trying someone in juvenile court vs. adult court, or determining whether to pursue the death penalty.

Then, as part of pretrial strategy, prosecutors are faced with making bail determinations. Smith and Levinson note that the assumption that suspects who have good jobs and connections in the community are generally regarded as less likely to flee; however, the stereotypes that plague African Americans as being lazier and less trustworthy may activate implicit biases that cause prosecutors to view their employment records with a greater degree of skepticism (Smith & Levinson, 2012). Also part of pretrial strategy are decisions related to whether to offer a plea bargain, and if so, what the reduced charge would be. Smith and Levinson state that implicit stereotypes can affect this process, as some people may be regarded more leniently as "troubled, but not a bad person," which increases the likelihood of being offered a plea bargain compared to an individual of a different race who may be deemed unlawful or dangerous (Smith & Levinson, 2012, p. 817-818).

Finally, prosecutorial implicit biases can crop up during aspects of trial strategy, such as jury selection and closing arguments. During closing arguments, in particular, prosecutors may activate implicit biases by referring to the accused in terms that dehumanize them, such as using animal imagery. Upon accounting for empirical studies that showed how many people still mentally connect Black people with apes (Goff, et al., 2008), the use of animal imagery is all the more alarming. Goff and colleagues found that as the number of ape references made about a Black defendant increased, so too did the likelihood of that defendant being sentenced to death

(Goff, et al., 2008). Thus, the language used by prosecutors can trigger implicit biases that dramatically affect trial outcomes.

Defense Attorneys

As with all of the other legal actors discussed so far, defense attorneys are not exempt from harboring implicit racial biases (Eisenberg & Johnson, 2004). Lyon recently identified two main areas where defense attorneys' implicit biases can affect the judicial process: within the attorney-client relationship, and during the jury selection process (Lyon, 2012). She writes about how implicit biases can influence how attorneys perceive their clients, such as seeing an "angry Black man" rather than a Black man who became frustrated when unable to understand the choices and consequences his attorney was outlining. Lyons also mentions how defense lawyers can fall into the trap of relying on humor to defuse stress; however, there is a need for caution in doing this so that the client is not dehumanized in the process.

With respect to voir dire and jury selection, defense attorneys must caution against relying on stereotypes to make assumptions about how a prospective juror may respond to the attorney's client and associated case. Lyons asserts that using stereotypes in juror selection may extend to implicit biases later in the judicial process. She warns, "If we are using this process of elimination based on stereotypes, jurors will know it. And then we cannot get angry if the jurors return the favor by making the assumption that our young male minority client is guilty, a gang member, or otherwise dangerous and not deserving of respect" (Lyon, 2012, p. 767).

Implicit Bias In
Health/Health Care

Implicit bias in Health/Health Care

The presence and prevalence of racial disparities in health and health care across a wide array of ailments have been documented extensively (for example overviews, see Elster, Jarosik, VanGeest, & Fleming, 2003; Mead, et al., 2008; Smedley, Stith, & Nelson, 2003; Stevens & Shi, 2003). Various explanations for these disparities range from individuals' lifestyle decisions, biomedical reasons, and social/environmental factors. The Kirwan Institute emphasizes this third category, noting that the social determinants of health such as where people are born, live, and work all affect health outcomes (for further information, see Daniels, Kennedy, & Kawachi, 1999; Social Determinants of Health," 2012; World Health Organization, 2008).

As discussed in this chapter, studies have documented the presence of implicit bias in a variety of facets of the health / healthcare industry.

Physicians' Implicit Biases

Like all other groups of people, health care professionals carry implicit biases that can influence their behaviors and judgments (Stone & Moskowitz, 2011). Some researchers have examined what role, if any, physicians' implicit racial biases play in the formation and perpetuation of these disparities. For example, Sabin and colleagues measured the implicit and explicit racial biases of 2,535 medical doctors and accounted for the physicians' own race and gender in their analysis. This research yielded three notable conclusions: 1) the doctors' implicit and explicit attitudes about race align well with the patterns found in large heterogeneous samples of the general population, as most doctors implicitly preferred Whites to Blacks; 2) on average, African American doctors did not display any implicit racial preferences for Whites or Blacks; 3) physician gender matters, as female doctors tended to hold fewer implicit racial biases (Sabin, et al., 2009). Other researchers have examined this phenomenon at an earlier stage, finding an implicit preference for Whites among first-year medical students (Haider, et al., 2011).

Furthering this line of inquiry, Moskowitz et al. conducted two studies aimed at exploring whether unconscious stereotypes influence the thoughts and behaviors of physicians. They identified the stereotypical maladies that medical professionals associated with African American patients and, by utilizing a reaction time procedure

in which subject received implicit primes, found that doctors implicitly associated certain diseases with African Americans (Moskowitz, Stone, & Childs, 2012). The authors articulate two main dangers of this implicit stereotyping: "(1) inaccurate components of a stereotype may be used in diagnosis and treatment without conscious knowledge of this influence, [and] (2) even an accurate stereotype may unduly influence diagnosis and treatment" (Moskowitz, et al., 2012, p. 1000).

Finally, in the realm of psychotherapy, Abreu used an experimental design to determine that therapists who were unknowingly primed with terms and stereotypes about a fictional patient rated the patient more negatively than a control group who were primed with neutral words (Abreu, 1999). This is yet another example of how the biases medical professionals carry can affect their patients.

Differential Treatment

Alarmingly, implicit biases have been shown to affect the type(s) and quality of care that patients of various races receive.

Schulman et al. examined racial variations in medical treatment using videos of actors portraying patients reporting chest pain. The patients were similar across several characteristics (e.g., socioeconomic status, type of insurance plan); however, they varied by race and sex. Results indicated that women and Blacks were less likely to be referred for cardiac catherization compared to their respective sex and race counterparts (Schulman, et al., 1999). Further analyses indicated that patients' race and sex independently influenced the doctors' recommendations, which provides insights into the differential treatment of cardiovascular disease (Schulman, et al., 1999).

A similar study by Weisse and colleagues explored whether the race and gender of patients influence doctors' decisions for treating patients who are reporting pain due to a kidney stone or back problems. Researchers presented doctors with vignettes depicting patients with these ailments. While patient race and gender varied across vignettes, the description and severity of his/her symptoms remained consistent across all cases. Researchers did not find any differences by race or gender with respect to the doctors' decision to administer treatment; however, they found that

treatments prescribed varied by patients' race and gender. With respect to race, male physicians prescribed higher doses of pain medication to White patients compared to Black patients, yet female doctors gave higher doses to Blacks (Weisse, Sorum, Sanders, & Syat, 2001). In terms of gender, male physicians prescribed higher doses of pain medicine to men, and female doctors gave higher doses of pain medication to female patients compared to males (Weisse, et al., 2001). These findings suggest that acknowledging a physicians' gender is key to understanding differential treatment of patients by race and gender (Weisse, et al., 2001).

The first study to provide compelling evidence of implicit bias among physicians using the IAT was conducted by Green et al. They sought to determine whether physicians held implicit racial biases, and if so, did the amount of implicit bias predict whether the doctors would prescribe thombolysis for Black and White patients displaying acute coronary symptoms. In terms of the results, the physicians reported not having any explicit preferences for Black or White patients; however, implicit measures recorded a preference for White patients and a belief that Black patients were less likely to cooperate with medical procedures (Green, et al., 2007). Most notably, the researchers found that increases in physicians' pro-White biases coincided with an increased likelihood of treating White patients with thrombolysis but not Black patients (Green, et al., 2007). Thus, the dissociation of implicit and explicit biases in a medical context can lead to differential treatment by race, which has obvious and important implications for patients' well-being.

Taking the study by Green et al., 2007 to a different context, Sabin and Greenwald used three different IAT tests to examine pediatricians' implicit attitudes and how they affect treatment recommendations for four pediatric conditions (pain, urinary tract infection, attention deficit hyperactivity disorder, and asthma). While there were not any significant associations between implicit attitudes and three of the diagnoses, researchers did uncover an association between unconscious biases related to patient race and prescribing narcotics for surgery-related pain (Sabin & Greenwald, 2012). Specifically, as pediatricians' pro-White implicit bias increased, so too did their inclination to prescribe pain-killing narcotics for White rather than Black patients (Sabin & Greenwald, 2012). Thus, implicit biases have been shown to influence patient treatment decisions even for youths.

Doctor – Patient Interactions

Another area affected by implicit bias in the healthcare realm is doctor-patient communication. Indeed, a study by Penner et al. concluded that White physicians' implicit racial biases led to less positive interactions with Black patients, particularly for doctors who displayed the combination of low explicit bias but high implicit bias (Penner, et al., 2010). Relatedly, a sample of physicians in another study found that "physicians were 23% more verbally dominant and engaged in 33% less patient-centered communication with African American patients than with White patients" (Johnson, Roter, Powe, & Cooper, 2004, p. 2084).

Looking at primary care clinicians, Cooper et al. examined how the implicit attitudes of primary care clinicians related to clinician-patient communication and patient ratings of care. Using an IAT that measured clinicians' race bias, the researchers found that higher implicit race bias scores generally were associated with more verbal dominance and lower patient positive affect for Black patients (Cooper, et al., 2012). From the perspective of Black patients, clinicians with higher IAT race bias were linked to Black patients feeling like they received less respect from the clinician, having less confidence in the clinician, and being less likely to recommend the clinician to other people (Cooper, et al., 2012). Conversely, White patients who interacted with clinicians who held higher levels of race bias felt that they were respected and liked (Cooper, et al., 2012).

Fostering Cultural Competency

Stone and Moskowitz define cultural competency in a medical environment as "the ability of systems to provide care to patients with diverse values, beliefs and behaviors, including their tailoring of delivery to meet patients' social, cultural and linguistic needs" (Stone & Moskowitz, 2011, p. 771). Contributing to the perpetuation of implicit biases in health care is the fact that medical professionals are not necessarily formally trained or well-versed in cultural competency (Carillo, Green, & Betancourt, 1999; White III, 2011).

Begun in 1997 and formally published in 2001, the Office of Minority Health in the

U.S. Department of Health and Human Services created the National Standards for Culturally and Linguistically Appropriate Services in Health Care (National Standards for Culturally and Linguistically Appropriate Services in Health Care, 2001). These standards "respond to the need to ensure that all people entering the health care system receive equitable and effective treatment in a culturally and linguistically appropriate manner"(National Standards for Culturally and Linguistically Appropriate Services in Health Care, 2001, p. 3). Stone and Moskowitz, however, point out that while these standards are a step in the right direction, they leave the way in which this material should be taught open to interpretation, thus medical professionals are left to "walk the thin line between the activation of cultural knowledge and the use of stereotypes" (Stone & Moskowitz, 2011, p. 772).

Several scholars have called for more cross-cultural/cultural competency curricula to educate medical professionals (Geiger, 2001; Stone & Moskowitz, 2011; White III, 2011). While these materials and teachings can be an important steps toward dismantling implicit biases, scholars also warn against stereotyping or oversimplifying a culture, as assuming that members of a given racial or ethnic group behave in a uniform and predicable manner is also problematic (Betancourt, 2004; Carillo, et al., 1999).

Concluding Thoughts

The impact of implicit biases in healthcare should not be understated. Moskowitz and colleagues capture the far-reaching effects, writing that implicit stereotypes can unintentionally affect medical professionals' "diagnoses, treatment recommendations, expectations about whether a patient will follow a prescribed treatment, and both verbal and nonverbal behavior toward patients during professional interactions, despite their intention to avoid such biases in conduct" (Moskowitz, et al., 2012). Anderson takes these concerns a step further by asserting that, "Of the four principles of bioethics, three—autonomy, non-maleficence, and justice—are most directly impacted by implicit bias"(Anderson, 2012). He stresses the need for medical providers "to increase the depth of their own understanding and to identify and utilize readily available resources to decrease both the occurrence and impact of implicit bias" (Anderson, 2012).

Debiasing

Debiasing

> *"The holy grail of implicit race bias research is to change the underlying associations that form the basis of implicit bias.*
>
> *— D. M. Amodio & Mendoza, 2010, p. 362 —*

While implicit biases are deeply entrenched in the subconscious, researchers generally agree that biases are malleable and that implicit associations may be unlearned (see, e.g., Blair, 2002; Blair, Ma, & Lenton, 2001; Dasgupta & Greenwald, 2001; Devine, 1989; Kang, 2009; Kang & Lane, 2010). As discussed in this chapter, the debiasing process can take many different forms and yield varying results depending on factors such as individual motivation and context, as these influence what associations are brought to the foreground of one's mind (Foroni & Mayr, 2005).

Debiasing is far from a simple task, as it involves the construction of new mental associations. Devine writes, "Inhibiting stereotype-congruent or prejudice-like responses and intentionally replacing them with non-prejudiced response can be likened to the breaking of a bad habit" (Devine, 1989, p. 15). She adds how "intention, attention, and time" are needed so that new responses are learned well enough to compete with the formerly automatically activated responses (Devine, 1989, p. 16). Given how strongly rooted implicit biases tend to be, debiasing efforts have to compete against stimuli that can, in effect, "re-bias" (Kang, et al., 2012, p. 1170).

The first inclination for many people who realize they hold implicit racial/ethnic biases may be to attempt to debias by repressing these biased thoughts; however, this notion generally has not been supported by the literature due to "rebound effects." Suppressing automatic stereotypes does not reduce them and may even amplify them by making them hyper-accessible (Galinsky & Moskowitz, 2000, 2007; Macrae, Bodenhausen, Milne, & Jetten, 1994). Studies have shown that "instead of repressing one's prejudices, if one openly acknowledges one's biases, and directly challenges or refutes them, one can overcome them" (bstan-'dzin-rgya-mtsho & Cuttler, 2009, p. 70). Similarly, Blair and Banaji found that conscious efforts to counter stereotypes can inhibit the activation of automatic associations (Blair & Banaji, 1996).

Numerous researchers have used the IAT to demonstrate the malleability of implicit attitudes. For example, Kawakami and colleagues found that proximity, as examined through approach-avoidance orientations, affected implicit racial bias scores across a series of four studies (Kawakami, Phills, Steele, & Dovidio, 2007). Ito and colleagues published another example of the IAT documenting the malleability of implicit racial attitudes that they achieved by manipulating participants' emotional cues. Some participants were surreptitiously induced to smile by holding a pencil in their mouths while viewing photos of unfamiliar Black or White men; others were not instructed to perform any such somatic manipulation. IAT results showed that this manipulation influenced IAT outcomes, as individuals who surreptitiously smiled while viewing Black faces earlier displayed less racial bias against Blacks (Ito, Chiao, Devine, Lorig, & Cacioppo, 2006). Richeson and Ambady considered the role of situational power (i.e., whether one is regarded as superior or subordinate in an interracial dyad) was reflected by changes in implicit racial attitudes in interracial interactions but not in same-race interactions (Richeson & Ambady, 2003). These studies, among others, declare implicit biases to be malleable.

With this in mind, it is logical that subsequent attention has been devoted to using this malleability property to counter existing biases. The following sections examine various debiasing techniques.

Interventions that may debias successfully

Counter-stereotypic training

A significant portion of the debiasing research centers on interventions that counter stereotypes and train individuals to develop new associations. One such example, advanced by Wittenbrink, Judd, and Park (2001) focused on how modifying the situational context may influence racial attitudes. By juxtaposing ordinary people in counter-stereotypic situations, such as depicting young White and Black males in scenes that included a church and a graffiti-strewn street corner, researchers found that the context condition affected participants' racial attitudes on a subsequent sequential priming task (Wittenbrink, Judd, & Park, 2001). The data from this counter-stereotypic training indicated that social category clues may affect individuals'

automatic responses and racial attitudes.

Taking the notion of countering stereotypes rather literally, Kawakami et al., 2000 studied the effects of training people to negate stereotypic associations, including racial associations. By instructing participants to verbally respond "no" when presented with a stereotypic trait that matched a category representation and "yes" when viewing non-stereotypic associations, they found that participants who received this stereotype negation training displayed diminished stereotype activation (Kawakami, et al., 2000).

Notably, this effect remained salient 24 hours after the training ended (Kawakami, et al., 2000). These findings emphasize the importance of not just counter-stereotypic instruction, but also the need for consistent repetition of this instruction. Kawakami and colleagues later extended this work to examine the effect of training on non-stereotypic traits of men and women in the context of hiring decisions and found similar results supporting the effectiveness of counter-stereotypic training (see Kawakami, Dovidio, & Kamp, 2005).

Finally, Blair and colleagues researched the strategy of using mental imagery as a way to moderate implicit stereotypes. Over the course of five experiments that used mental imagery to target gender stereotypes, they found compelling evidence that counter-stereotypical mental imagery yielded notably weaker implicit stereotypes as compared to the implicit stereotypes assessed in individuals who either engaged in other forms of mental imagining or no mental imagery whatsoever (Blair, et al., 2001).

Rather than only mental, imagery in other forms may be used to debias. For courtroom settings, Kang and colleagues suggest the use of posters, pamphlets, photographs, and similar materials that would provoke counter-typical associations in the minds of jurors and judges (Kang, et al., 2012). The effects of this would likely vary based on the amount of exposure, with the underlying intention that even momentarily activating a different association may help decrease the presence of implicit bias during legal processes.

Exposure to counter-stereotypic individuals

Another type of intervention focuses on exposing people to individuals who contradict widely-held stereotypes. One fascinating study by Dasgupta and Greenwald (2001) investigated whether exposure to counter-stereotypic exemplars could decrease automatic preferences, such as that for White over Black Americans. They found that exposure to pro-Black exemplars (e.g., Michael Jordan, Colin Powell, Martin Luther King, Jr.) as opposed to nonracial or pro-White exemplars (e.g., Tom Hanks, Jay Leno, John F. Kennedy) significantly decreased the automatic White preference effect, as measured by the IAT (Dasgupta & Greenwald, 2001). Similar to the findings of Kawakami et al., 2000, this effect had staying power, as the IAT effect of this pro-Black condition remained 24 hours after the exposure to images of admired Blacks and disliked Whites (e.g., Jeffrey Dahmer, Timothy McVeigh, Al Capone). Emphasizing the malleability of implicit biases, the authors suggest that "creating environments that highlight admired and disliked members of various groups … may, over time, render these exemplars chronically accessible so that they can consistently and automatically override preexisting biases"(Dasgupta & Greenwald, 2001, p. 807).

This scholarship aligns well with the concept of debiasing agents, which refers to individuals whose traits contrast with the stereotypes typically associated with that particular category (Kang & Banaji, 2006). The presence of debiasing agents decreases the implicit biases of those they encounter due to their unique positioning. Example debiasing agents would include male nurses, elderly athletes, and female scientists. In many cases, debiasing agents change individuals' implicit stereotypes, not just their implicit attitudes (Kang & Banaji, 2006). However, to be effective, debiasing agents must be viewed as not merely an exception bur rather connect that individual to relevant categories, regardless of any counter-stereotypical traits they may also possess (Kang & Banaji, 2006).

The success of the exposure to counter-stereotypic individuals intervention has been echoed by other studies that do not focus explicitly on race. For example, one study found that exposure to women in leadership positions at a women's college led to students being less likely to express automatic gender stereotypes about women,

compared to students from a coeducational college (Dasgupta & Asgari, 2004).

Moreover, exposure to counter-stereotypic exemplars does not even need to occur through in-person interactions. Renowned implicit bias scholar Mahzarin Banaji works to offset her own implicit biases through viewing images of counter-stereo-typical individuals on her computer screensaver (Lehrman, 2006). Photos and other wall décor can serve a similar purpose (Kang, et al., 2012; National Center for State Courts).

However, some researchers question this counter-stereotypic exemplar debiasing method. First, an extensive 2010 study by Schmidt and Nosek examined whether Barack Obama, as a high-status famous Black exemplar shifted implicit or explicit racial attitudes during his candidacy and early presidency. Results from a heterogeneous sample of nearly 480,000 individuals led to the conclusion that there was minimal evidence that implicit racial attitudes changed systematically due to Obama's presence as a counter-stereotypic exemplar (Schmidt & Nosek, 2010).

The authors suggest that the mere presence of a high-status counter-stereotypic exemplar may be inadequate to shift implicit or explicit racial attitudes (Schmidt & Nosek, 2010). Building on this work, Lybarger and Monteith's research concluded that President Obama's saliency alone did not have a debiasing effect, as one individual may be inadequate to shift long-standing implicit racial associations (Lybarger & Monteith, 2011).

Second, Joy-Gaba and Nosek provide a word of caution regarding the perceived malleability of implicit biases through exposure to counter-stereotypic exemplars. Their efforts to replicate and expand upon the work of Dasgupta and Greenwald (2001) noted earlier in this subsection yielded conspicuously less convincing results. Notably, while the degree of malleability found by Dasgupta and Greenwald was quite high on both the initial measure (d = 0.82) and follow-up 24 hours later (d = 0.71), Joy-Gaba and Nosek found significantly weaker effect magnitudes (d = 0.17 and 0.14, respectively) (Joy-Gaba & Nosek, 2010). This discrepancy leads the authors to broadly conclude that the extent to which the literature declares implicit biases to be malleable may have been overstated (Joy-Gaba & Nosek, 2010).

Intergroup Contact

Championed by American psychologist Gordon W. Allport in 1954, intergroup contact theory asserts that four key conditions are the necessary for positive effects to emerge from intergroup contact (Allport, 1954). Allport stipulated that optimal intergroup contact involves individuals of equal status, which explains why some relationships, such as that of student and teacher, do not necessarily lead to reductions in bias. Other conditions that yield positive intergroup contact effects include sharing common goals, interacting in a cooperative rather than competitive setting, and being supported by authority figures, laws, or customs. Allport's theory has been supported consistently in the literature, including through large-scale meta-analyses (see, e.g.,Pettigrew & Tropp, 2006, 2011).

Beyond simply gaining familiarity with outgroups through intergroup contact, these interactions have been shown to reduce implicit bias. For example, Thomas Pettigrew's multi-national study found that "the reduction in prejudice among those with diverse friends generalizes to more positive feelings about a wide variety of outgroups" (Pettigrew, 1997, p. 180-181). Moreover, ten years later in a meta-analytic test of intergroup contact theory, Pettigrew and Tropp examined 713 samples and concluded that ingroup contact generally reduces intergroup prejudice (Pettigrew & Tropp, 2006).

With respect to the realms of criminal justice and health care previously discussed, intergroup contact can play a debiasing role in specific contexts. Peruche and Plant studied police officers and noted that "high levels of negative contact with Black people at work were related to negative expectations regarding Black suspects and marginally more negative attitudes toward Black people generally;" however, intergroup contact with Blacks outside of the workplace countered these effects (Peruche & Plant, 2006, p. 197). Similarly, diverse clinical care teams are vital to health care, because in a diverse team where members are granted equal power, "a sense of camaraderie develops that prevents the further development of stereotypes based on race/ethnicity, gender, culture or class" (Betancourt, 2004, p. 108).

Education about Implicit Bias

Efforts aimed at raising awareness of the phenomenon of implicit bias can also debias. This education can take several forms. For example, U.S. district judge Mark W. Bennett educates potential jurors about implicit bias during his time with them during the juror selection process. Judge Bennett aims to explain implicit bias and make jurors skeptical of their own objectivity through a 25-minute lesson that concludes by asking each juror to sign a pledge against bias. The text of this pledge is prominently displayed in the jury room. Then, at the beginning of the trial, Judge Bennett reiterates how implicit bias can taint jurors' judgment by giving a short speech[2] before the lawyers' opening statements. He believes in the positive outcomes studies have documented regarding individuals' responses to awareness of their own implicit biases (Bennett, 2010).

A recent article by Anna Roberts strongly supports the idea of using the IAT to educate jurors about implicit bias while dismissing the notion that the IAT should be used to "screen" prospective jurors (Roberts, 2012). Much like Judge Mark Bennett, Roberts recommends that this educational component be integrated into juror orientation, preferably with jurors receiving hands-on experiential learning that includes taking the IAT, as this is more impactful than passively learning about the IAT and its findings (Roberts, 2012).

Judges can also benefit from implicit bias education, which in large part involves persuading them of the presence of this problem (Kang, et al., 2012; Saujani, 2003). Organizations such as the National Center for State Courts (NCSC) have begun creating resources[3], such as films and assessments, designed to raise judicial awareness of implicit biases and their implications in a courtroom setting. Results from pilot sites showed promising preliminary results (Kang, et al., 2012). Research suggests that educating judges about implicit bias is most effective under three circumstances: 1) training should start early, such as during new judge orientation when

---

2. See Appendix A for the text of this brief speech.

3. The National Center for State Courts have posted some of their educational films and materials here: http://www.ncsconline.org/D_Research/ref/implicit.html

people are most open to new ideas; 2) the training should be presented in such a way that judges do not feel defensive; it should not be accusatory in nature; and 3) judges should be encouraged to take the IAT, as the results often prompt action (Kang, et al., 2012).

An entire industry around diversity education and trainings has proliferated in recent years, offering participants promises of reduced prejudice and greater appreciation of various cultures. Studies have examined whether diversity education can counter implicit biases, though the results are mixed. One study found that Whites who volunteered for diversity education forums showed lower levels of implicit and explicit anti-Black prejudice, with the change in implicit orientations "predicted by emotion-based factors, including reduced fear of Blacks, and liking for the Black professor who taught the course" (Rudman, 2004, p. 136; Rudman, Ashmore, & Gary, 2001). Conversely, other research finds diversity training yield minimal effects, particularly from a long-term perspective (Rynes & Rosen, 1995).

Accountability

Having a sense of accountability, meaning "the implicit or explicit expectation that one may be called on to justify one's beliefs, feelings, and actions to others," can be another powerful measure to combat bias (Lerner & Tetlock, 1999, p. 255). Research finds that having a sense of accountability can decrease the influence of bias (Kang, et al., 2012; Reskin, 2005). When decision makers are not held accountable for their actions, they are less likely to self-check for how bias may affect their decision-making (National Center for State Courts). Jurors' feelings of being held accountable by the judge to produce unbiased decisions can help jurors keep their implicit biases in check (Kang, et al., 2012).

Fostering Egalitarian Motivations

Considerable research has shown that fostering egalitarian motivations can counter the activation of automatic stereotypes (Dasgupta & Rivera, 2006; Moskowitz, Gollwitzer, Wasel, & Schaal, 1999). Stone and Moskowitz write, "When activated, egalitarian goals inhibit stereotypes by undermining and counteracting the implicit

nature of stereotype activation, thereby cutting stereotypes off before they are brought to mind" (Stone & Moskowitz, 2011, p. 773). For example, work by Dasgupta and Rivera found that automatic biases are not necessarily inevitable, as the relationship between automatic antigay prejudice and discrimination was moderated by individuals' conscious holding of egalitarian beliefs (Dasgupta & Rivera, 2006).

Taking the Perspective of Others

Another debiasing strategy that has gained some traction is when individuals take the perspective of someone who is different from them. Across three experiments, Galinksy and Moskowitz found that perspective-taking was effective at debiasing, as it "tended to increase the expression of positive evaluations of the target, reduced the expression of stereotypic content, and prevented the hyperaccessibility of the stereotype construct" (Galinsky & Moskowitz, 2000, p. 720).

Benforado and Hanson support perspective-taking as a debiasing tool, noting that considering opposing perspectives and fostering recognition of multiple perspectives are good techniques for reducing automatic biases (Benforado & Hanson, 2008). They caution, however, that this approach may have limited effects if individuals believe they have taken the perspective of others when in fact they have not been as successful at this venture as they judge themselves to be.

Later empirical work by Todd et al. shed light on effects of perspective taking. The researchers employed five experiments designed to assess whether taking the perspective of others could counter automatic expressions of racial bias. Their findings found that this debiasing technique yielded "more favorable automatic interracial evaluations" (Todd, Bodenhausen, Richeson, & Galinsky, 2011, p. 1038).

Taking the perspective of others can also be used in debiasing exercises. When Stone and Moskowitz outlined the components of a cultural competency workshop for medical professionals that sought to educate them about implicit bias, the authors suggested that the medical professionals imagine themselves as a minority group patient and write a story about that person's life (Stone & Moskowitz, 2011). Finally, it is worth noting that perspective taking has benefits that extend beyond

debiasing. For example, in the realm of healthcare, studies have shown that encouraging practitioners to take the perspective of others cultivates empathy, which leads to positive outcomes for patient satisfaction and treatment (see, e.g., Blatt, LaLacheur, Galinsky, Simmens, & Greenberg, 2010; Drwecki, Moore, Ward, & Prkachin, 2011).

Deliberative Processing

Another technique that can counter implicit biases is to "engage in effortful, deliberative processing" (Kang, et al., 2012, p. 1177). This is particularly important for individuals who may be operating under time constraints or a weighty cognitive load, such as doctors and judges, because spontaneous judgments can provoke reliance on stereotypes (Burgess, 2010; Kang, et al., 2012). To that end, Betancourt suggests that medical professionals constantly self-monitor their behaviors in an effort to offset implicit stereotyping (Betancourt, 2004).

In another manner of deliberative processing, Stone and Moskowitz encourage medical professionals to rethink the standard ways that patients are classified (e.g., race/ethnicity, gender, etc.) and instead focus on a common identity that they share with each patient (Stone & Moskowitz, 2011). By activating this shared identity, the patient's other identities (e.g., race/ethnicity) are not as prevalent in the medical professional's mind, thus helping to counter the enactment of the implicit biases and stereotypes associated with those identities (Stone & Moskowitz, 2011).

The significance of deliberative processing is reinforced by research that finds that even one's emotional state can influence the activation and nature of implicit biases (Dasgupta, DeSteno, Williams, & Hunsinger, 2009). For example, DeSteno and colleagues examined how the creation of automatic outgroup prejudice can be affected by emotional states, such as anger or sadness. Using both an evaluative priming measure and the IAT, they found that an angry emotional state led to automatic prejudice against outgroups, which the researchers attributed to anger's association with intergroup competition and conflict (DeSteno, Dasgupta, Bartlett, & Cajdric, 2004). Thus, deliberate processing, including self-awareness of one's own emotional state, plays a role in individuals' ability to counter implicit biases.

Other Interventions

As discussed in this final subsection, some researchers have developed specific interventions as a means of debiasing.

As an extension of work that relied on counter-stereotypic exemplars, Foroni & Mayr showed how short fictional scenarios designed to present a counter-stereotypic example (in this case, flowers were regarded noxious while insects were positively regarded) had an immediate and automatic modulation of the IAT effect (Foroni & Mayr, 2005). This same effect was not observed when subjects were simply asked to think of flowers as negative and insects as positive. "These results suggest that a newly acquired knowledge structure targeting the abstract, category level can produce behavioral effects typically associated with automatic categorization" (Foroni & Mayr, 2005, p. 139).

A recent publication by Devine et al. highlights an intervention that is founded on the premise that "implicit bias is like a habit that can be broken through a combination of awareness of implicit bias, concern about the effects of that bias, and the application of strategies to reduce bias" (Devine, Forscher, Austin, & Cox, 2012, p. 1267). The logic here is that "breaking the habit" of implicit bias requires awareness of the contexts that can activate bias and knowledge of how to replace biased reactions with ones that reflect a non-prejudiced mindset. Devine and colleagues sought to assess whether interventions could yield long-term reductions in implicit racial bias. They used a randomly controlled experimental design in which participants assigned to the intervention group engaged in a bias education and training program that taught participants five strategies they could apply to different situations in their lives as appropriate (stereotype replacement, counter-stereotypic imaging, individuation, perspective taking, and increasing opportunities for contact). Results showed that the participants who had received this training had lower IAT scores than the control group participants, and unprecedentedly, this reduction in implicit race bias endured for at least eight weeks following the intervention (Devine, et al., 2012). Devine et al. attribute this decline in implicit bias to the multifaceted nature of the intervention rather than any specific aspect of the intervention.

Conclusion

Conclusion

> *It is well documented that conscious and unconscious race bias, even rank discrimination based on race, remain alive in our land, impeding realization of our highest values and ideals.*
>
> *– Supreme Court Justice Ruth Bader Ginsburg, "Grutter v. Bollinger," 2003 –*

The aforementioned studies underscore the devastating impact of implicit racial biases with a focus on education, criminal justice, and health care. These topics provide a mere snap shot of the range of realms affected by implicit biases. Other researchers have documented the role of implicit bias in domains such as employment (see, e.g., Bertrand & Mullainathan, 2004; Rooth, 2007) and aspects of the legal landscape, just to name a few (see, e.g., Brown, 2012). Despite a plethora of scientific evidence, not all scholars are swayed by implicit bias research. This final chapter summarizes some of the critiques implicit bias research has faced and concludes by offering some concrete steps for addressing implicit biases.

Critiques of Implicit Bias Research

Like most fields, implicit bias has been subject to criticism. The following points capture a few of these critiques, many of which are summarized in Kang & Lane 2010:

- It's "junk science" – Some critics question the scientific validity of implicit bias research, characterizing it as "motivated by junk scientists and their lawyer accomplices who manipulate data, misinterpret results, and exaggerate findings in order to snooker society into politically correct wealth transfers" (Kang & Lane, 2010, p. 504-505). Kang and Lane argue against this so-called "backlash scholarship," asserting that perfect knowledge of a concept should not be required given that the same level of scrutiny is not often inflicted upon the status quo (Kang & Lane, 2010). Others question measures of implicit bias because they are indirect in nature, thus any concept being studied relies on participants' performance in another assignment, such as the sorting procedure in the

IAT (Nosek & Riskind, 2012). Finally, other detractors remain unsatisfied with the results of scientific tests of validity (see, e.g., G. Mitchell & Tetlock, 2006), though a later meta-analysis by Greenwald and colleagues provided convincing evidence to counter many of these claims.

- Implicit biases are "hardwired" – This stance asserts that implicit biases are immutable; therefore, we are unable to change them in any way. As discussed in the previous chapter, the creation of new associations, while not easy, is generally now regarded as a feasible task (Blair, 2002; Dasgupta & Greenwald, 2001; Devine, 1989; Kang, 2009; Kang & Lane, 2010).

- Implicit biases are rational – Another objection to implicit bias research contends that because implicit biases reflect reality, it is rational to act on them accordingly (Kang & Lane, 2010). This argument hinges on the apparent accuracy of individuals' perceptions of reality.

- Unclear and /or limited policy impacts – Implicit bias research has also been subject to criticism because its connection to social policy is not immediately evident. Policies generally aim at addressing behaviors rather than individuals' thoughts, though one's actions and mental processes are certainly interconnected (Nosek & Riskind, 2012). However, Nosek and Riskind assert that implicit and explicit biases should be accounted for when crafting and implementing policy, because often policies are formulated under the erroneous assumption that people's behaviors are the result of completely conscious and intentional actions (Nosek & Riskind, 2012).

Criticism of the Implicit Association Test (IAT)

The Implicit Association Test, in particular, has been subject to a distinct set of criticisms. One argument asserts that the IAT is suspect because it is difficult to determine what exactly the IAT is measuring. While the standard explanation is that the race IAT measures the attitudes that individuals hold toward specific groups, others have contended that IAT results may actually reflect test takers' familiarity with

one group over another (e.g., greater familiarity with White people as opposed to Blacks) (Kinoshita & Peek-O'Leary, 2005). This claim of familiarity tainting IAT results directly counters earlier work by Dasgupta and colleagues who documented the persistence of a pro-White bias on the race IAT even after accounting for familiarity in their study design (Dasgupta, et al., 2000).

A second measurement-based criticism that has been leveled against the IAT questions whether it is assessing individuals' attitudes as purported, or whether it is actually measuring their cultural knowledge. Some question whether anti-Black sentiments uncovered by the IAT actually reflect negative feelings associated with African Americans' history of oppression versus personal anti-Black sentiments (Tetlock & Mitchell, 2008). Other work by Nosek and Hansen counters this notion, as their findings across seven different studies yielded the conclusion that the association between implicit attitudes and cultural knowledge is weak and inconsistent (Nosek & Hansen, 2008).

Another critical line of questioning that the IAT has faced interrogates whether the IAT measures a stable attitude or if the context of the test instead yields "an experimentally induced momentary response to the mapping project subjects face" (Tinkler, 2012, p. 992). In other words, this criticism questions the extent to which results are a methodological artifact of the IAT itself (Fiedler, et al., 2006; Tetlock & Mitchell, 2008).

Beyond the realm of measurement, another lingering criticism of the IAT is its reliance on hypothetical situations. Some argue that the contrived scenarios of the IAT calls into question the test's predictive validity in real life situations. Mitchell and Tetlock go so far as to list twelve specific ways in which implicit bias research laboratory settings contrasts with what happens in real-life workplace settings (see pgs. 1109-1110 in G. Mitchell & Tetlock, 2006). In light of this criticism, researchers have begun crafting experiments that take advantage of non-hypothetical situations to assess how implicit biases affect behavior, such as Stepanikova and colleagues' recent assessment of monetary generosity in light of IAT-measured implicit bias

scores (Stepanikova, Triplett, & Simpson, 2011).

Steps for Addressing Implicit Biases

Rudman (2004) offers several suggestions for what to do about implicit biases. They are summarized as follows:

- Increase awareness of the existence of implicit biases, because such efforts "are critical to our ability to provide descriptions of social cognition that are faithful to human complexity, as well as to endeavors to combat automatic prejudice" (p. 138). Dovidio et al. echo this mandate and add to it by asserting that "as important first step is making people aware of discrepancies between their conscious ideals and automatic negative responses" (Dovidio, et al., 1997, p. 535).

- Invest energy into advocating for policies (e.g., affirmative action) that can counter the effects of implicit bias.

- Strive for an inclusive society, because "to the extent that societal evaluations color implicit biases, a more inclusive society should reduce them" (p. 138).

- Provide people with opportunities to engage with outgroup members in contexts of mutual trust that allow for intergroup interactions to counter implicit biases.

- Recognize that we may not be able to wholly 're-wire' people's implicit biases, but "what we can do is become aware of the power of implicit partisanship" (p. 139).

A Broader, Interdisciplinary Future

In addition to permeating public discourse, as noted in the introduction, implicit bias research is also attracting increased interest from academic fields within and outside of psychology. Initiatives such as the Implicit Bias & Philosophy Inter-

national Research Project[4] seek to uncover ways in which these two fields, that are not traditionally aligned, may inform one another. Efforts such as Project Implicit Mental Heath[5] explore unconscious reactions to a range of mental health issues, including anxiety, depression, mental illness, and eating disorders. These interdisciplinary ventures are likely to help shape the field of implicit bias in the years to come.

4. For more information, please visit http://www.biasproject.org
5. For more information, please visit https://implicit.harvard.edu/implicit/user/pimh/index.jsp

APPENDIX

Judge Mark Bennett
Glossary
Works Cited

Appendix A: Judge Mark Bennett – Debiasing in the Courtroom

Judge Mark Bennett is a U.S. district judge in the Northern District of Iowa. Before opening statements, he gives jurors the following instructions regarding implicit biases:

> *Do not decide the case based on 'implicit biases.' As we discussed in jury selection, everyone, including me, has feelings, assumptions, perceptions, fears, and stereotypes, that is, "implicit biases," that we may not be aware of. These hidden thoughts can impact what we see and hear, how we remember what we see and hear, and how we make important decisions. Because you are making very important decisions in this case, I strongly encourage you to evaluate the evidence carefully and to resist jumping to conclusions based on personal likes or dislikes, generalizations, gut feelings, prejudices, sympathies, stereotypes, or biases. The law demands that you return a just verdict, based solely on the evidence, your individual evaluation of that evidence, your reason and common sense, and these instructions. Our system of justice is counting on you to render a fair decision based on the evidence, not on biases.*

– Kang & Lane, 2010, p. 500-501 –

Appendix B: Glossary

Above Average Effect
The Above Average Effect is a psychological phenomenon wherein individuals rate their own capabilities more highly than they regard those of other individuals. This rating of oneself as consistently above average is also known as illusory superiority. In the context of implicit bias, one study asked a group of 36 judges whether they were in the top half of judges who could avoid prejudiced decision-making. Researchers cited the Above Average Effect when 97% of those judges placed themselves in the top half. The Above Average Effect makes us think that we are the ones who are immune from bias.

Activation
Activation refers to how different parts of the brain display more neural activity depending on the task a subject is performing. Activation can be seen during Func-

tional Magnetic Resonance Imaging (fMRIs) when sections of the brain "light up" due to increased blood flow.

Afrocentric Features Bias

This form of bias refers to the generally negative judgments and beliefs that many people hold regarding individuals who possess particularly Afrocentric features, notably dark skin, a wide nose, and full lips. The presence of these features may activate associations that lead to stereotypical perceptions. To illustrate, one study examined the Afrocentric features and sentences of convicted felons and found that when controlling for numerous factors (e.g., seriousness of the primary offense, number of prior offenses, number of additional concurrent offenses, etc.), individuals with the most prominent Afrocentric features received sentences of up to 56-64 months more prison time than their less Afrocentrically featured counterparts. Thus, even judges, whose careers center on objectivity and fair sentencing, have been shown to possess Afrocentric features bias.

Amygdala

Considered part of the limbic system, the amygdala is located within the medial temporal lobe. It plays a key role in how we process emotions, notably fear and pleasure. The amygdala is involved with the expression of implicit attitudes; it activates when someone perceives threat or anxiety, such as when another individual appears untrustworthy. In light of this, researchers have studied amygdala activation to better understand Implicit Association Test results.

Association

In psychology, the extent to which two concepts are connected in individuals' minds is called an association. For example, many associations reflect cultural stereotypes, such as youthfulness being associated with positive terms. Associations are learned and developed over time; we are constantly absorbing and developing both positive and negative associations, even from a very young age. Notwithstanding our external commitments to equality, fairness, and similar principles, we can still hold associations that reflect our implicit biases.

Attitude

An attitude is an association between some concept (e.g., a social group) and the evaluation thereof, such as liking/disliking or favoring/disfavoring.

Bias

A bias is a prejudice that leads to a tendency to favor one entity over another, often unfairly. Biases can be explicit or implicit.

Categorical Boundaries

Categorical boundaries refer to the abstract and fluid "lines" we draw that separate one grouping from another, thereby creating categories. These "lines" are malleable; they move to fit situational context. For example, the categorical boundary that defines the distinction between fast and slow changes depending on whether a vehicle is traveling 45 mph in a school zone versus an interstate highway.

Concealed Bias

Concealed bias is explicit bias that one hides for purposes of impression management.

Debiasing

Debiasing is the process of reducing susceptibility to implicit biases or dismantling them. Starting with a conscious goal to be fair or simply being aware of implicit bias is not enough to remove, overcome, or dismantle its effects on decision-making. Instead, research studies highlight several techniques that have been shown to successfully dismantle bias, including:

1. Social contact – Meaningful intergroup interactions allow us to find connections or links to people who are different from us. This opportunity to relate to one another can help eliminate bias.
2. Counter-stereotypic Training – Explicitly stating "no" when facing a stereotype can have a debiasing effect if this technique is practiced repeatedly and reinforced over time.
3. Cuing Positive Exemplars – Exposing people to images of outstanding and admired individuals who belong to groups that typically are stereotyped negatively can reduce bias. In one study, researchers found that implicit biases

for Whites decreased when exposing individuals to both admired Black and disliked White individuals.

4. Counter-stereotypic Image Reinforcement – While it can be very hard to "unlearn" an association, teaching a new, different association can have positive effects. Being exposed to images that counter stereotypes (e.g., a female scientist) can contribute to debiasing.

Dissociations

Dissociations occur when an individual's implicit and explicit attitudes on the same topic/object differ. This suggests that implicit and explicit biases are related yet distinct mental constructs.

Dual-Process Theory

Dual-Process Theory explores how phenomena can occur as a result of individuals' reasoning occurring through two distinct processes. One process is rapid, implicit, and occurs automatically; conversely, the second process is slower, controlled, and explicit. This theory asserts that these two reasoning processes compete for control within individuals and influence their responses.

Explicit Attitudes and Beliefs

The opposite of implicit, explicit attitudes and beliefs are the ones that individuals profess publicly or express directly. These attitudes and beliefs are conscious and acknowledged by the individual who holds them.

Extinction

Also known as reversal, extinction is the process of unlearning mental associations. Extinction conditions the brain to no longer respond in the same way to specific cues. It can be used to dismantle implicit biases and is particularly effective when paired with counter-stereotypic image reinforcement, which refers to the viewing of images that are concrete examples designed to counter stereotypes (e.g., a female construction worker).

Functional Magnetic Resonance Imaging (fMRI)

fMRI is a type of MRI that takes structural pictures of the brain. More specifically,

fMRIs measure and record brain activity by detecting blood flow, thus allowing scientists to understand the brain's neural activity in a continuous and dynamic manner. In implicit bias studies involving fMRIs, researchers ask subjects to think specific thoughts or perform certain tasks and then examine which portions of the brain image "light up" or are activated, thus providing insights into the subject's mind processes.

Illusion of Objectivity

Built on the idea of detached impartiality, the illusion of objectivity refers to the false impression that one may be free from biases, opinions, and other subjective influences. Under the illusion of objectivity, we think that we are the only people who are objective, whereas everyone else is biased. In short, the illusion of objectivity is a bias that makes us think we are not actually biased. Psychological research has shown that we all harbor implicit biases, regardless of how strong our desires are to be unbiased or neutral. In legal settings, judges should acknowledge the illusion of subjectivity lest they become overconfident in their ability to make impartial judgments.

Implicit Association Text (IAT)

The Implicit Association Test (IAT) is a computer-based test that measures the strength of associations individuals hold on an unconscious (or implicit) level. It measures associations and stereotypes that people often do not even realize they possess. In the IAT, participants pair two concepts (e.g., men and career vs. men and family) and the test measures the speed with which participants are able to classify words and images related to these concepts. The more closely associated the concepts are in the participants' minds (e.g., women and family), the more rapidly they will be able to respond; conversely, concepts that are not as strongly associated (e.g., women and career) take longer to classify. The speed with which individuals make these classifications sheds light on the associations they hold and the strength of those associations. Researchers contend that association strengths are at the heart of individuals' attitudes and stereotypes.

Implicit Bias

Also known as unconscious or hidden bias, implicit biases are negative associations that people unknowingly hold. They are expressed automatically without conscious

awareness. Many studies have indicated that implicit biases affect individuals' attitudes and actions, thus creating real-world implications even though individuals may not even be aware that these biases exist within themselves. Notably, implicit biases have been shown to trump individuals' stated commitments to equality and fairness, thereby producing behavior that diverges from the explicit attitudes that many people profess. The Implicit Association Test (IAT) is often used to measure implicit biases with regard to race, gender, sexual orientation, age, religion, and other topics.

Ingroup
An ingroup is a group with which one feels a sense of membership, solidarity, or shared interest. Ingroup membership may be established along numerous identities, such as race, religion, sexual orientation, etc.

Ingroup Bias
Ingroup bias refers to the positive attitudes that people tend to feel towards member of their ingroup. Feelings of safety and familiarity are often associated with ingroup members. Research has shown that people are inclined to forgive members of their ingroup more quickly than they are members of outgroups.

Insula / Insular Cortex
Located deep within the brain between the temporal lobe and the parietal lobe, the insula is involved in several processes, including risky decision-making, bodily awareness, and certain aspects of motor control. It is also the part of the brain that is involved in the experience of disgust with respect to both smells and sights, even if these cues are only imagined. As such, neuroscientists often look for activation in the insula when studying implicit biases that prompt strong feelings of disgust in research subjects.

Outgroup
In contrast to ingroups, outgroups are groups with which one does not belong or associate. As a byproduct of ingroup favoritism, some people feel a sense of dislike or contempt toward members of outgroups.

Positivity Bias

When we evaluate our abilities to be better than average, or better than they actually are, we are engaging in positivity bias. This form of bias involves making predominantly favorable judgments. More specifically, positivity bias can also refer to the attribution of individuals' successes to internal factors while explaining their failures through citing external factors.

Schema

A schema is a mental shortcut we automatically use to organize and categorize the vast amounts of information we encounter in our daily lives. As mental "blueprints," schemata streamline information processing by creating generalized categories and expectations for objects, events, places, people, and so on. These abstract representations are often but not necessarily always accurate. Schemata unfortunately can contribute to stereotyping and prejudice when they cause people to misinterpret situations or exclude information in order to conform to pre-existing schemata. This can perpetuate implicit biases.

Shooter Bias

Shooter bias is a term that emerged from research studies that examined subjects' reactions to a series of images of Black or White men. In the images, the men were holding either a gun or a benign object such as a can of soda pop. On the computerized test, participants were instructed to "shoot" if the man was holding a gun, refrain if he was holding any other object, and make this decision as quickly as possible. Shooter bias reflects the finding that participants made faster decisions and were more likely to shoot when the image depicted a Black man, even when this was an erroneous decision and the Black man was not actually wielding a gun.

Stereotype

A stereotype is a standardized and simplified belief about the attributes of a social group. Although not always accurate, stereotypes are often widely held and can have both positive and negative impacts of individuals. The act of stereotyping involves both activation (making the cognitive association) and application (using that association to make a judgment about a person or group).

Appendix C: Works Cited

Abreu, J. M. (1999). Conscious and non-conscious African American Stereotypes: Impact on First Impression and Diagnostic Ratings by Therapists. Journal of Consulting and Clinical Psychology, 67(3), 387-393.

Allport, G. W. (1954). The Nature of Prejudice. Cambridge, MA: Addison-Wesley.

Amodio, D., & Devine, P. G. (2009). On the Interpersonal Functions of Implicit Stereotyping and Evaluative Race Bias: Insights from Social Neuroscience. In R. E. Petty, R. H. Fazio & P. Briñol (Eds.), Attitudes: Insights from the New Implicit Measures (pp. 193-226). New York, NY: Psychology Press.

Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual Differences in the Activation and Control of Affective Race Bias as Assessed by Startle Eyeblink Response and Self-Report. Journal of Personality and Social Psychology, 84(4), 738-753.

Amodio, D. M., & Mendoza, S. A. (2010). Implicit Intergroup Bias: Cognitive, Affective, and Motivational Underpinnings. In B. Gawronski & B. K. Payne (Eds.), Handbook of Implicit Social Cognition (pp. 353-274). New York, NY: The Guilford Press.

Anderson, J. (2012). Should PAs Assess and Address the Impact of Implicit Bias on Patient Care? Journal of the American Academy of Physician Assistants. Retrieved from http://www.jaapa.com/should-pas-assess-and-address-the-impact-of-implicit-bias-on-patient-care/article/235552/

Arnold, M., & Lassmann, M. E. (2003). Overrepresentation of Minority Students in Special Education. Education, 124(2), 230-236.

Arterton, J. B. (2008). Unconscious Bias and the Impartial Jury. Connecticut Law Review, 40(4), 1023-1033.

Ashburn-Nardo, L., Knowles, M. L., & Monteith, M. J. (2003). Black Americans' Implicit Racial Associations and their Implications for Intergroup Judgment. Social Cognition, 21(1), 61-87.

Ashburn-Nardo, L., Voils, C. I., & Monteith, M. J. (2001). Implicit Associations as the Seeds of Intergroup Bias: How Easily Do They Take Root? Journal of Personality and Social Psychology, 81(5), 789-799.

Banks, J. A. (1995). Multicultural Education and Curriculum Transformation. The Journal of Negro Education, 64(4), 390-400.

Bargh, J. A. (1997). The Automaticity of Everyday Life. In R. S. Wyler Jr. (Ed.), Advances in Social Cognition (Vol. 10, pp. 1-61). Mahwah, NJ: Lawrence Erlbaum Associates.

Bargh, J. A. (1999). The Cognitive Monster: The Case Against the Controllability of Automatic Stereotype Effects. In S. Chaiken & Y. Trope (Eds.), Dual-Process Theories in Social Psychology (pp. 361-382). New York, NY: The Guilford Press.

Baron, A. S., & Banaji, M. R. (2006). The Development of Implicit Attitudes: Evidence of Race Evaluations From Ages 6 and 10 and Adulthood. Psychological Science, 17(1), 53-58.

Baron, R. M., Tom, D. Y. H., & Cooper, H. M. (1985). Social Class, Race, and Teacher Expectancies. In J. B. Dusek (Ed.), Teacher Expectancies (pp. 251-270). Hillsdale: Erlbaum.

Benforado, A., & Hanson, J. (2008). The Great Attributional Divide: How Divergent Views of Human Behavior Are Shaping Legal Policy. Emory Law Journal, 57(2), 311-408.

Bennett, M. W. (2010). Unraveling the Gordian Knot of Implicit Bias in Jury Selection: The Problems of Judge-Dominated Voir Dire, the Failed Promise of Batson, and Proposed Solutions. Harvard Law and Policy Review, 4(1), 149-171.

Bertrand, M., Chugh, D., & Mullainathan, S. (2005). Implicit Discrimination. The American Economic Review, 95(2), 94-98.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. The American Economic Review, 94(4), 991-1013.

Betancourt, J. R. (2004). Not Me!: Doctors, Decisions, and Disparities in Health Care. Cardiovascular Reviews and Reports, 25(3), 105-109.

Betts, J. R., & Grogger, J. (2003). The Impact of Grading Standards on Student Achievement, Educational Attainment, and Entry-Level Earnings. Economics of Education Review, 22(4), 343-352.

Blair, I. V. (2002). The Malleability of Automatic Stereotypes and Prejudice. Personality and Social Psychology Review, 6(3), 242-261.

Blair, I. V., & Banaji, M. R. (1996). Automatic and Controlled Processes in Stereotype Priming. Journal of Personality and Social Psychology, 70(6), 1142-1163.

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The Influence of Afrocentric Facial Features in Criminal Sentencing. Psychological Science, 15(10), 674-679.

Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imaging Stereotypes Away: The Moderation of Implicit Stereotypes Through Mental Imagery. Journal of Personality and Social Psychology, 81(5), 828-841.

Blanton, H., Jaccard, J., Klick, J., Mellers, B., Mitchell, G., & Tetlock, P. E. (2009). Strong Claims and Weak Evidence: Reassessing the Predictive Validity of the IAT. Journal of Applied Psychology, 94(3), 567-582.

Blascovich, J., Mendes, W. B., Hunter, S. B., Lickel, B., & Kowai-Bell, N. (2001). Perceiver Threat in Social Interactions with Stigmatized Others. Journal of Personality and Social Psychology, 80(2), 253-267.

Blatt, B., LaLacheur, S. F., Galinsky, A. D., Simmens, S. J., & Greenberg, L. (2010). Does Perspective-Taking Increase Patient Satisfaction in Medical Encounters? Academic Medicine, 85(9), 1445-1452.

Bogardus, E. S. (1933). A Social Distance Scale. Sociology & Social Research, 17(3), 265-271.

Bosson, J. K., William B. Swann, J., & Pennebaker, J. W. (2000). Stalking the Perfect Measure of Implicit Self-Esteem: The Blind Men and the Elephant Revisited? Journal of Personality and Social Psychology, 79(4), 631-643.

Bowers, W. J., Sandys, M., & Brewer, T. W. (2004). Crossing Racial Boundaries: A Closer Look at the Roots of Racial Bias in Capital Sentencing When the Defendant is Black and the Victim is White. DePaul Law Review, 53(4), 1497-1538.

Brown, D. A. (2012). Tax Law: Implicit Bias and the Earned Income Tax Credit. In J. D. Levinson & R. J. Smith (Eds.), Implicit Racial Bias Across the Law (pp. 164-178). Cambridge, MA: Cambridge University Press.

bstan-'dzin-rgya-mtsho, & Cuttler, H. C. (2009). The Art of Happiness: A Handbook for Living. New York, NY: Penguin Group.

Burgess, D. J. (2010). Are Providers More Likely to Contribute to Healthcare Disparities Under High Levels of Cognitive Load? How Features of the Healthcare Setting May Lead to Baises in Medical Decision Making. Medical Decision Making, 30(2), 246-257.

Carillo, J. E., Green, A. R., & Betancourt, J. R. (1999). Cross-Cultural Primary Care: A Patient-Based Approach. Annals of Internal Medicine, 130(10), 829-834.

Christian, D. (1997). Vernacular Dialects in U.S. Schools. Cal Digest. Retrieved from http://www.cal.org/resources/digest/digest_pdfs/christ01.pdf

Cooper, L. A., Roter, D. L., Carson, K. A., Beach, M. C., Sabin, J. A., Greenwald, A. G., et al. (2012). The Associations of Clinicians' Implicit Attitudes About Race with Medical Visit Communication and Patient Ratings of Interpersonal Care. American Journal of Public Health, 102(5), 979–987.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The Police Officer's Dilemma: Using Ethnicity to Disambiguate Potentially Threatening Individuals. Journal of Personality and Social Psychology, 83(6), 1314-1329.

Correll, J., Park, B., Judd, C. M., Wittenbrink, B., Sadler, M. S., & Keesee, T. (2007). Across the Thin Blue Line: Police Officers and Racial Bias in the Decision to Shoot. Journal of Personality and Social Psychology, 92(6), 1006-1023.

Cross, J. B., DeVaney, T., & Jones, G. (2001). Pre-Service Teacher Attitudes Toward Differing Dialects. Linguistics and Education, 12(2), 211-227.

Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable Neural Components in the Processing of Black and White Faces. Psychological Science, 15(12), 806-813.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit Attitude Measures: Consistency, Stability, and Convergent Validity. Psychological Science, 12(2), 163-170.

Cvencek, D., Greenwald, A. G., Brown, A. S., Gray, N. S., & Snowden, R. J. (2010). Faking of the Implicit Association Test is Statistically Detectable and Partly Correctable. Basic and Applied Social Psychology, 32(4), 302-314.

Daniels, N., Kennedy, B. P., & Kawachi, I. (1999). Why Justice is Good for Our Health: The Social Determinants of Health Inequalities. Daedalus, 128(4), 215-251.

Dasgupta, N. (2004). Implicit Ingroup Favoritism, Outgroup Favoritism, and Their Behavioral Manifestations. Social Justice Research, 17(2), 143-169.

Dasgupta, N., & Asgari, S. (2004). Seeing is Believing: Exposure to Counter-stereotypic Women Leaders and Its Effect on the Malleability of Automatic Gender Stereotyping. Journal of Experimental Social Psychology, 40(5), 642-658.

Dasgupta, N., DeSteno, D., Williams, L. A., & Hunsinger, M. (2009). Fanning the Flames of Prejudice: The Influence of Specific Incidental Emotions on Implicit Prejudice. Emotion, 9(4), 585-591.

Dasgupta, N., & Greenwald, A. G. (2001). On the Malleability of Automatic Attitudes: Combating Automatic Prejudice With Images of Admired and Disliked Individuals. Journal of Personality and Social Psychology, 81(5), 800-814.

Dasgupta, N., Greenwald, A. G., & Banaji, M. R. (2003). The First Ontological Challenge to the IAT: Attitude or Mere Familiarity? Psychological lnquiry, 14(3/4), 238-243.

Dasgupta, N., McGhee, D. E., Greenwald, A. G., & Banaji, M. R. (2000). Automatic Preference for White Americans: Eliminating the Familiarity Explanation. Journal of Experimental Social Psychology, 36(3), 316-328.

Dasgupta, N., & Rivera, L. M. (2006). From Automatic Antigay Prejudice to Behavior: The Moderating Role of Conscious Beliefs About Gender and Behavioral Control. Journal of Personality and Social Psychology, 91(2), 268-280.

Davis, M., & Whalen, P. (2001). The Amygdala: Viligance and Emotion. Molecular Psychiatry, 6(1), 13-34.

DeSteno, D., Dasgupta, N., Bartlett, M. Y., & Cajdric, A. (2004). Prejudice From Thin Air: The Effect of Emotion on Automatic Intergroup Attitudes. Psychological Science, 15(5), 319-324.

Devine, P. G. (1989). Stereotypes and Prejudice: Their Automatic and Controlled Components. Journal of Personality and Social Psychology, 56(1), 5-18.

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. L. (2012). Long-Term Reduction in Implicit Bias: A Prejudice Habit-Breaking Intervention. Journal of Experimental Social Psychology, 48(6), 1267-1278.

Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The Regulation of Explicit and Implicit Race Bias: The Role of Motivations to Respond Without Prejudice. Journal of Personality and Social Psychology, 82(5), 835-848.

Dixon, T. L., & Linz, D. (2000). Race and the Misrepresentation of Victimization on Local Television News. Communication Research, 27(5), 547-573.

Dovidio, J. F., & Fazio, R. H. (1992). New Technologies for the Direct and Indirect Assessment of Attitudes. In J. M. Tanur (Ed.), Questions about Questions: Inquiries into the Cognitive Bases of Surveys (pp. 204-237). New York, NY: Russell Sage Foundation.

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and Explicit Prejudice and Interracial Interaction. Journal of Personality and Social Psychology, 82(1), 62-68.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the Nature of Prejudice: Automatic and Controlled Processes. Journal of Experimental Social Psychology, 33, 510-540.

Dovidio, J. F., Kawakami, K., Smoak, N., & Gaertner, S. L. (2009). The Nature of Contemporary Racial Prejudice. In R. E. Petty, R. H. Fazio & P. Briñol (Eds.), Attitudes: Insights from the New Implicit Measures (pp. 165-192). New York, NY: Psychology Press.

Drwecki, B. B., Moore, C. F., Ward, S. E., & Prkachin, K. M. (2011). Reducing Racial Disparities in Pain Treatment: The Role of Empathy and Perspective-Taking. Pain, 152(5), 1001-1006.

Dunton, B. C., & Fazio, R. H. (1997). An Individual Difference Measure of Motivation to Control Prejudiced Reactions. Personality and Social Psychology Bulletin, 23(3), 316-326.

Dusek, J. B., & Joseph, G. (1983). The Bases of Teacher Expectancies: A Meta-Analysis. Journal of Educational Psychology, 75(3), 327-346.

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking Deathworthy: Perceived Stereotypicality of Black Defendants Predicts Capital-Sentencing Outcomes. Psychological Science, 17(5), 383-386.

Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing Black: Race, Crime, and Visual Processing. Journal of Personality and Social Psychology, 87(6), 876-893.

Egloff, B., & Schmukle, S. C. (2002). Predictive Validity of an Implicit Association Test for Assessing Anxiety. Journal of Personality and Social Psychology, 83(6), 1441-1455.

Egloff, B., Schwerdtfeger, A., & Schmukle, S. C. (2005). Temporal Stability of the Implicit Association Test - Anxiety. Journal of Personality Assessment, 84(1), 82-88.

Eisenberg, T., & Johnson, S. L. (2004). Implicit Racial Attitudes of Death Penalty Lawyers. DePaul Law Review, 53(4), 1539-1556.

Elster, A., Jarosik, J., VanGeest, J., & Fleming, M. (2003). Racial and Ethnic Disparities in Health Care for Adolescents: A Systemic Review of the Literature. Archives of Pediatrics and Adolescent Medicine, 157(9), 867-874.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in Automatic Activation as an Unobtrustive Measure of Racial Attitudes: A Bona Fide Pipeline? Journal of Personality and Social Psychology, 69(6), 1013-1027.

Fazio, R. H., & Olson, M. A. (2003). Implicit Measures in Social Cognition Research: Their Meaning and Use. Annual Review of Psychology, 54(1), 297-327.

Feistritzer, C. E. (2011). Profile of Teachers in the U.S. in 2011. Washington, DC: National Center for Education Information.

Ferguson, A. A. (2000). Bad Boys: Public Schools in the Making of Black Masculinity. Ann Arbor, MI: University of Michigan Press.

Ferguson, R. F. (2003). Teachers' Perceptions and Expectations and the Black-White Test Score Gap. Urban Education, 38(4), 460-507.

Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and Unaided Response Control on the Implicit Association Tests. Basic and Applied Social Psychology, 27(4), 307-316.

Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved Problems with the 'I', the 'A', and the 'T': A Logical and Psychometric Critique of the Implicit Association Test (IAT). European Review of Social Psychology, 17(3), 74-147.

Figlio, D. N., & Lucas, M. E. (2004). Do High Grading Standards Affect Student Performance? Journal of Public Economics, 88, 1815-1834.

Foroni, F., & Mayr, U. (2005). The Power of a Story: New, Automatic Associations from a Single Reading of a Short Scenario. Psychonomic Bulletin & Review, 12(1), 139-144.

Gaertner, S. L., & McLaughlin, J. P. (1983). Racial Stereotypes: Associations and Ascriptions of Positive and Negative Characteristics. Social Psychology Quarterly, 46(1), 23-30.

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-Taking: Decreasing Stereotype Expression, Stereotype Accessibility, and In-Group Favoritism. Journal of Personality and Social Psychology, 78(4), 708-724.

Galinsky, A. D., & Moskowitz, G. B. (2007). Further Ironies of Suppression: Stereotype and Counterstereotype Accessibility. Journal of Experimental Social Psychology, 43(5), 833-841.

Geiger, H. J. (2001). Racial Stereotyping and Medicine: The Need for Cultural Competence. Canadian Medical Association Journal, 164(12), 1699-1700.

Georgia v. McCollum, 505 42 (U.S. 1992).

Gilbert, D. T., & Hixon, J. G. (1991). The Trouble of Thinking: Activation and Application of Stereotypic Beliefs. Journal of Personality and Social Psychology, 60(4), 509-517.

Goff, P. A., Eberhardt, J. L., Williams, M. J., & Jackson, M. C. (2008). Not Yet Human: Implicit Knowledge, Historical Dehumanization, and Contemporary Consequences. Journal of Personality and Social Psychology, 94(2), 292-306.

Graham, S., & Lowery, B. S. (2004). Priming Unconscious Racial Stereotypes about Adolescent Offenders. Law and Human Behavior, 28(5), 483-504.

Green, A. R., Carney, D. R., Pallin, D. J., Ngo, L. H., Raymond, K. L., Iezzoni, L. I., et al. (2007). Implicit Bias among Physicians and its Prediction of Thrombolysis Decisions for Black and White Patients. Journal of General Internal Medicine, 22(9), 1231-1238.

Greenwald, A. G. Implicit Association Test: Validity Debates. from http://faculty. washington.edu/agg/iat_validity.htm

Greenwald, A. G., & Banaji, M. R. (1995). Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes. Psychological Review, 102(1), 4-27.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A Unified Theory of Implicit Attitudes, Stereotypes, Self-Esteem, and Self-Concept. Psychological Review, 109(1), 3-25.

Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to Measure Self-Esteem and Self-Concept. Journal of Personality and Social Psychology, 79(6), 1022-1038.

Greenwald, A. G., & Krieger, L. H. (2006). Implicit Bias: Scientific Foundations. California Law Review, 94(4), 945-967.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring Individual Differences in Implicit Cognition: The Implicit Association Test. Journal of Personality and Social Psychology, 74(6), 1464-1480.

Greenwald, A. G., & Nosek, B. A. (2001). Health of the Implicit Association Test at Age 3. Zeitschrift für Experimentelle Psychologie, 48(2), 85-93.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm. Journal of Personality and Social Psychology, 85(2), 197-216.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity. Journal of Personality and Social Psychology, 97(1), 17-41.

Greenwald, A. G., Smith, C. T., Sriram, N., Bar-Anan, Y., & Nosek, B. A. (2009). Implicit Race Attitudes Predicted Vote in the 2008 U.S. Presidential Election. Analyses of Social Issues & Public Policy, 9(1), 241-253.

Grutter v. Bollinger, 539 306 (U.S. 2003).

Guthrie, C., Rachlinski, J. J., & Wistrich, A. J. (2009). The "Hidden Judiciary": An Empirical Examination of Executive Branch Justice. Duke Law Journal, 58(7), 1477-1530.

Haider, A. H., Sexton, J., Sriram, N., Cooper, L. A., Efron, D. T., Swoboda, S., et al. (2011). Association of Unconscious Race and Social Class Bias With Vignette-Based Clinical Assessments by Medical Students. Journal of the American Medical Association, 306(9), 942-951.

Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., & Rauch, S. L. (2000). Differential Response in the Human Amygdala to Racial Outgroup vs Ingroup Face Stimuli. NeuroReport, 11(11), 2351-2355.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A Meta-Analysis on the Correlation Between the Implicit Association Test and Explicit Self-Report Measures. Personality and Social Psychology Bulletin, 31(10), 1369-1385.

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing Prejudice: Implicit Prejudice and the Perception of Facial Threat. Psychological Science, 14(6), 640-643.

Iowa: Ruling for State in 'Implicit Bias' Suit. (2012, April 18). The New York Times, p. A19. Retrieved from http://www.nytimes.com/2012/04/18/us/iowa-ruling-for-state-in-implicit-bias-suit.html

Ito, T. A., Chiao, K. W., Devine, P. G., Lorig, T. S., & Cacioppo, J. T. (2006). The Influence of Facial Feedback on Race Bias. Psychological Science, 17(3), 256-261.

Johnson, R. L., Roter, D., Powe, N. R., & Cooper, L. A. (2004). Patient Race/Ethnicity and Quality of Patient-Physician Communication During Medical Visits. American Journal of Public Health, 94(12), 2084-2090.

Jones, E. E., & Sigall, H. (1971). The Bogus Pipeline: A New Paradigm for Measuring Affect and Attitude. Psychological Bulletin, 76(5), 349-364.

Jost, J. T., Rudman, L. A., Blair, I. V., Carney, D. R., Dasgupta, N., Glaser, J., et al. (2009). The existence of implicit bias is beyond reasonable doubt: A refutation of ideological and methodological objections and executive summary of ten studies that no manager should ignore. Research in Organizational Behavior, 29, 39-69.

Joy-Gaba, J. A., & Nosek, B. A. (2010). The Surprisingly Limited Malleability of Implicit Racial Evaluations. Social Psychology 41(3), 137-146.

Kang, J. (2005). Trojan Horses of Race. Harvard Law Review, 118(5), 1489-1593.

Kang, J. (2008). Memorandum on Implicit bias.

Kang, J. (2009). Implicit Bias: A Primer for the Courts: Prepared for the National Campaign to Ensure the Racial and Ethnic Fairness of America's State Courts.

Kang, J. (2010). Implicit Bias and the Pushback from the Left. St. Louis University Law Journal, 54, 1139-1150.

Kang, J. (2012). Communications Law: Bits of Bias. In J. D. Levinson & R. J. Smith (Eds.), Implicit Racial Bias Across the Law (pp. 132-145). Cambridge, MA: Cambridge University Press.

Kang, J., & Banaji, M. (2006). Fair Measures: A Behavioral Realist Revision of 'Affirmative Action'. California Law Review, 94, 1063-1118.

Kang, J., Bennett, M., Carbado, D., Casey, P., Dasgupta, N., Faigman, D., et al. (2012). Implicit Bias in the Courtroom. UCLA Law Review, 59(5), 1124-1186.

Kang, J., & Lane, K. (2010). Seeing Through Colorblindness: Implicit Bias and the Law. UCLA Law Review, 58(2), 465-520.

Kawakami, K., Dovidio, J. F., & Kamp, S. v. (2005). Kicking the Habit: Effects of Nonstereotypic Association Training and Correction Processes on Hiring Decisions. Journal of Experimental Social Psychology, 41(1), 68-75.

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just Say No (to Stereotyping): Effects of Training in the Negation of Stereotypic Associations on Stereotype Activation. Journal of Personality and Social Psychology, 78(5), 871-888.

Kawakami, K., Phills, C. E., Steele, J. R., & Dovidio, J. F. (2007). (Close) Distance Makes the Heart Grow Fonder: Improving Implicit Racial Attitudes and Interracial Interactions Through Approach Behaviors. Journal of Personality and Social Psychology, 92(6), 957-971.

Kim, D.-Y. (2003). Voluntary Controllability of the Implicit Association Test (IAT). Social Psychology Quarterly, 66(1), 83-96.

Kinoshita, S., & Peek-O'Leary, M. (2005). Does the Compatibility Effect in the Race Implicit Association Test Reflect Familiarity or Affect? Psychonomic Bulletin & Review, 12(3), 442-452.

Kruglanski, A. W., & Freund, T. (1983). The Freezing and Unfreezing of Lay-Inferences: Effects on Impressional Primacy, Ethnic Stereotyping, and Numerical Anchoring. Journal of Experimental Social Psychology, 19(5), 448-468.

Lehrman, S. (2006, June). The Implicit Prejudice. Scientific American, 294, 32-34.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the Effects of Accountability. Psychological Bulletin, 125(2), 255-275.

Levinson, J. D. (2007). Forgotten Racial Equality: Implicit Bias, Decisionmaking, and Misremembering. Duke Law Journal, 57(2), 345-424.

Levinson, J. D., Cai, H., & Young, D. (2010). Guilty By Implicit Racial Bias: The Guilty/Not Guilty Implicit Association Test. Ohio State Journal of Criminal Law, 8(1), 187-208.

Levinson, J. D., & Young, D. (2010). Different Shades of Bias: Skin Tone, Implicit Racial Bias, and Judgments of Ambiguous Evidence. West Virginia Law Review, 112(2), 307-350.

Lewin, T. (2012, March 6). Black Students Face More Discipline, Data Suggests. The New York Times. Retrieved from http://www.nytimes.com/2012/03/06/education/black-students-face-more-harsh-discipline-data-shows.html

Lieberman, M. D., Hariri, A., Jarcho, J. M., Eisenberger, N. I., & Bookheimer, S. Y. (2005). An fMRI Investigation of Race-Related Amygdala Activity in African-American and Caucasian-American Individuals. Nature Neuroscience, 8(6), 720-722.

Lybarger, J. E., & Monteith, M. J. (2011). The Effect of Obama Saliency on Individual-Level Racial Bias: Silver Bullet or Smokescreen? Journal of Experimental Social Psychology, 47(3), 647-652.

Lyon, A. D. (2012). Race Bias and the Importance of Consciousness for Criminal Defense Attorneys. Seattle University Law Review, 35, 755-768.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of Mind but Back in Sight: Stereotypes on the Rebound. Journal of Personality and Social Psychology, 67(5), 808-817.

Mahajan, N., Martinez, M. A., Gutierrez, N. L., Diesendruck, G., Banaji, M. R., & Santos, L. R. (2011). The Evolution of Intergroup Bias: Perceptions and Attitudes in Rhesus Macaques. Journal of Personality and Social Psychology, 100(3), 387-405.

Majors, R., & Billson, J. M. (1992). Cool Pose: The Dilemmas of Black Manhood in America. New York, NY: Touchstone.

McConnell, A. R., & Liebold, J. M. (2001). Relations among the Implicit Association Test, Discriminatory Behavior, and Explicit Measures of Attitudes. Journal of Experimental Social Psychology, 37(5), 435-442.

McElroy, M. (2012). Unconscious Racial Attitudes Playing Large Role in 2012 Presidential Vote. The Cutting Edge. Retrieved from http://www.thecuttingedgenews.com/index.php?article=73355&pageid=37&pagename=Page+One

McKown, C., & Weinstein, R. S. (2002). Modeling the Role of Child Ethnicity and Gender in Children's Differential Response to Teacher Expectations. Journal of Applied Social Psychology, 32(1), 159-184.

Mead, H., Cartwright-Smith, L., Jones, K., Ramos, C., Woods, K., & Siegel, B. (2008). Racial and Ethnic Disparities in U.S. Health Care: A Chartbook. New York, NY: The Commonwealth Fund.
Merton, R. K. (1957). Social Theory and Social Structure. New York, NY: Free Press.

Mitchell, G., & Tetlock, P. E. (2006). Antidiscrimination Law and the Perils of Mindreading. Ohio State Law Journal, 67(5), 1023-1121.

Mitchell, T. L., Haw, R. M., Pfeifer, J. E., & Meissner, C. A. (2005). Racial Bias in Mock Juror Decision-Making: A Meta-Analytic Review of Defendant Treatment. Law and Human Behavior, 29(6), 621-637.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious Control of Stereotype Activation Through Chronic Egalitarian Goals. Journal of Personality and Social Psychology, 77(1), 167-184.

Moskowitz, G. B., Stone, J., & Childs, A. (2012). Implicit Stereotyping and Medical Decisions: Unconscious Stereotype Activation in Practitioners' Thoughts About African Americans. American Journal of Public Health, 102(5), 996-1001.

Myers, E. M., & Zeigler-Hill, V. (2012). How Much Do Narcissists Really Like Themselves? Using the Bogus Pipeline Procedure to Better Understand the Self-Esteem of Narcissists. Journal of Research in Personality, 46(1), 102-105.

National Center for State Courts. Strategies to Reduce the Influence of Implicit Bias. Williamsburg, VA.

National Standards for Culturally and Linguistically Appropriate Services in Health Care. (2001). Washington, D.C.: U.S. Department of Health and Human Services, Office of Minority Health.

Neal, L. V. I., McCray, A. D., Webb-Johnson, G., & Bridgest, S. T. (2003). The Effects of African American Movement Styles on Teachers' Perceptions and Reactions. The Journal of Special Education, 37(1), 49-57.

Nier, J. A. (2005). How Dissociated Are Implicit and Explicit Racial Attitudes? A Bogus Pipeline Approach. Group Processes & Intergroup Relations, 8(1), 39-52.

Nisbett, R. E., & Wilson, T. D. (1977). Telling More Than We Can Know: Verbal Reports on Mental Processes. Psychological Review, 84(3), 231-259.

Nosek, B. A., & Banaji, M. R. (2002). (At Least) Two Factors Moderate the Relationship Between Implicit and Explicit Attitudes. In R. K. Ohme & M. Jarymowicz (Eds.), Natura Automatyzmow (pp. 49-56). Warszawa: WIP PAN & SWPS.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting Implicit Group Attitudes and Beliefs From a Demonstration Web Site. Group Dynamics: Theory, Research, and Practice, 6(1), 101-115.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and Using the Implicit Association Test: II. Method Variables and Construct Validity. Personality and Social Psychology Bulletin, 31(2), 166-180.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at Age 7: A Methodological and Conceptual Review. In J. A. Bargh (Ed.), Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes (pp. 265-292). New York: Psychology Press.

Nosek, B. A., & Hansen, J. J. (2008). The Associations in Our Heads Belong to Us: Searching for Attitudes and Knowledge in Implicit Evaluation. Cognition and Emotion, 22(4), 553-594.

Nosek, B. A., & Riskind, R. G. (2012). Policy Implications of Implicit Social Cognition. Social Issues and Policy Review, 6(1), 113-147.

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Linder, N. M., Ranganath, K. A., et al. (2007). Pervasiveness and Correlates of Implicit Attitudes and Stereotypes. European Review of Social Psychology, 18, 36-88.

NPR (Producer). (2012, June 18) Why Race Could Color the Vote Against Obama. Science Friday. Radio broadcast retrieved from http://www.npr.org/2012/05/11/152508354/why-race-could-color-the-vote-against-obama

Oliver, M. B. (2003). African American Men as 'Criminal and Dangerous': Implications of Media Portrayals of Crime on the 'Criminalization' of African American Men. Journal of African American Studies, 7(2), 3-18.

Olson, M. A., & Fazio, R. H. (2004). Trait Inferences as a Function of Automatically Activated Racial Attitudes and Motivation to Control Prejudiced Reactions Basic and Applied Social Psychology, 26(1), 1-11.

Olson, M. A., & Fazio, R. H. (2007). Discordant Evaluations of Blacks Affect Nonverbal Behavior. Personality and Social Psychology Bulletin, 33(9), 1214-1224.

Orne, M. T. (1962). On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and their Implications. American Psychologist, 17(11), 776-783.

Payne, B. K. (2001). Prejudice and Perception: The Role of Automatic and Controlled Processes in Misperceiving a Weapon. Journal of Personality and Social Psychology, 81(2), 181-192.

Payne, B. K. (2006). Weapon Bias: Split-Second Decisions and Unintended Stereotyping. Current Directions in Psychological Science, 15(6), 287-291.

Payne, B. K., Krosnick, J. A., Pasek, J., Lelkes, Y., Akhtar, O., & Tompson, T. (2010). Implicit and Explicit Prejudice in the 2008 American Presidential Election. Journal of Experimental Social Psychology, 46(2), 367-374.

Penner, L. A., Dovidio, J. F., West, T. V., Gaertner, S. L., Albrecht, T. L., Dailey, R. K., et al. (2010). Aversive Racism and Medical Interactions with Black Patients: A Field Study. Journal of Experimental Social Psychology, 46(2), 436-440.

Peruche, B. M., & Plant, E. A. (2006). The Correlates of Law Enforcement Officers' Automatic and Controlled Race-Based Responses to Criminal Suspects. Basic and Applied Social Psychology, 28(2), 193-199.

Pettigrew, T. F. (1997). Generalized Intergroup Contact Effects on Prejudice. Personality and Social Psychology Bulletin, 23(2), 173-185.

Pettigrew, T. F., & Tropp, L. R. (2006). A Meta-Analytic Test of Intergroup Contact Theory. Journal of Personality and Social Psychology, 90(5), 751-783.

Pettigrew, T. F., & Tropp, L. R. (2011). When Groups Meet: The Dynamics of Intergroup Contact. Philadelphia, PA: Psychology Press.

Petty, R. E., Fazio, R. H., & Briñol, P. (2009). The New Implicit Measures: An Overview. In R. E. Petty, R. H. Fazio & P. Briñol (Eds.), Attitudes: Insights from the New Implicit Measures (pp. 3-18). New York, NY: Psychology Press.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., et al. (2000). Performance on Indirect Measures of Race Evaluation Predicts Amygdala Activation. Journal of Cognitive Neuroscience, 12(5), 729-738.

Pichon, S., Gelder, B. d., & Grèzes, J. (2009). Two Different Faces of Threat: Comparing the Neural Systems for Recognizing Fear and Anger in Dynamic Body Expressions. NeuroImage, 47(4), 1873-1883.

Pizzi, W. T., Blair, I. V., & Judd, C. M. (2005). Discrimination in Sentencing on the Basis of Afrocentric Features. Michigan Journal of Race and Law, 10(2), 327-353.

Plant, E. A., & Devine, P. G. (1998). Internal and External Motivation to Respond Without Prejudice. Journal of Personality and Social Psychology, 75(3), 811-832.

Plant, E. A., & Peruche, B. M. (2005). The Consequences of Race for Police Officers' Responses to Criminal Suspects. Psychological Science, 16(3), 180-183.

Plant, E. A., Peruche, B. M., & Butz, D. A. (2005). Eliminating Automatic Racial Bias: Making Race Non-Diagnostic for Responses to Criminal Suspects. Journal of Experimental Social Psychology, 41(2), 141-156.

powell, j., & Godsil, R. (2011, September / October). Implicit Bias Insights as Preconditions to Structural Change. Poverty & Race, 20, 4-6.

Pronin, E. (2006). Perception and misperception of bias in human judgment. Trends in Cognitive Sciences, 11(1), 37-43.

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(3), 369-381.

Rachlinski, J. J., Johnson, S. L., Wistrich, A. J., & Guthrie, C. (2009). Does Unconscious Racial Bias Affect Trial Judges? Notre Dame Law Review, 84(3), 1195-1246.

Reskin, B. (2005). Unconsciousness Raising. Regional Review, 14(3), 32-37.

Richardson, L. S. (2011). Arrest Efficiency and the Fourth Amendment. Minnesota Law Review, 95(6), 2035-2098.

Richeson, J. A., & Ambady, N. (2003). Effects of Situational Power on Automatic Racial Prejudice. Journal of Experimental Social Psychology, 39(2), 177-183.

Roberts, A. (2012). (Re)forming the Jury: Detection and Disinfection of Implicit Juror Bias. Connecticut Law Review, 44(3), 827-882.

Ronquillo, J., Denson, T. F., Lickel, B., Lu, Z.-L., Nandy, A., & Maddox, K. B. (2007). The Effects of Skin Tone on Race-Related Amygdala Activity: An fMRI Investigation. Social Cognitive and Affective Neuroscience, 2(1), 39-44.

Rooth, D.-O. (2007). Implicit Discrimination in Hiring: Real World Evidence. Bonn, Germany: Institute for the Study of Labor.

Rosenthal, R., & Jacobson, L. (1968). Pygmalion in the Classroom. New York, NY: Rinehart & Winston.

Rudman, L. A. (2004). Social Justice in Our Minds, Homes, and Society: The Nature, Causes, and Consequences of Implicit Bias. Social Justice Research, 17(2), 129-142.

Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" Automatic Biases: The Malleability of Implicit Prejudice and Stereotypes. Journal of Personality and Social Psychology, 81(5), 856-868.

Rudman, L. A., Feinberg, J., & Fairchild, K. (2002). Minority Members' Implicit Attitudes: Automatic Ingroup Bias as a Function of Group Status. Social Cognition, 20(4), 294-320.

Rutland, A., Cameron, L., Milne, A., & McGeorge, P. (2005). Social Norms and Self-Presentation: Children's Implicit and Explicit Intergroup Attitudes. Child Development, 76(2), 451-466.

Rynes, S., & Rosen, B. (1995). A Field Survey of Factors Affecting the Adoption and Perceived Success of Diversity Training. Personnel Psychology, 48(2), 247-270.

Sabin, J. A., & Greenwald, A. G. (2012). The Influence of Implicit Bias on Treatment Recommendations for 4 Common Pediatric Conditions: Pain, Urinary Tract Infection, Attention Deficit Hyperactivity Disorder, and Asthma. American Journal of Public Health, 102(5), 988-995.

Sabin, J. A., Nosek, B. A., Greenwald, A. G., & Rivara, F. P. (2009). Physicians' Implicit and Explicit Attitudes About Race by MD Race, Ethnicity, and Gender. Journal of Health Care for the Poor and Underserved, 20(3), 896-913.

Sanbonmatsu, D. M., & Fazio, R. H. (1990). The Role of Attitudes in Memory-Based Decision Making. Journal of Personality and Social Psychology, 59(4), 614-622.

Saujani, R. M. (2003). 'The Implicit Association Test': A Measure of Unconscious Racism in Legislative Decision-making. Michigan Journal of Race and Law, 8(2), 395-423.

Schmidt, K., & Nosek, B. A. (2010). Implicit (and Explicit) Race Attitudes Barely Changed During Barack Obama's Presidential Campaign and Early Presidency. Journal of Experimental Social Psychology, 46(2), 308-314.

Schulman, K. A., Berlin, J. A., Harless, W., Kerner, J. F., Sistrunk, S., Gersh, B. J., et al. (1999). The Effect of Race and Sex on Physicians' Recommendations for Cardiac Catherization. The New England Journal of Medicine, 340(8), 618-626.

Sekaquaptewa, D., Espinoza, P., Thompson, M., Vargas, P., & Hippel, W. v. (2003). Stereotypic Explanatory Bias: Implicit Stereotyping as a Predictor of Discrimination. Journal of Experimental Social Psychology, 39(1), 75-82.

Sherwin, G. H., & Schmidt, S. (2003). Communication Codes Among African American Children and Youth - The Fast Track from Special Education to Prison? Journal of Correctional Education, 54(2), 45-52.

Sigall, H., & Page, R. (1971). Current Stereotypes: A Little Fading, A Little Faking. Journal of Personality and Social Psychology, 18(2), 247-255.

Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. (2000). The Color of Discipline: Sources of Racial and Gender Disproportionality in School Punishment. Bloomington, IN: The Indiana Education Policy Center.

Smedley, B. D., Stith, A. Y., & Nelson, A. R. (Eds.). (2003). Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Washington, D.C.: National Academies Press.

Smith, R. J., & Cohen, G. B. (2012). Capital Punishment: Choosing Life or Death (Implicitly). In J. D. Levinson & R. J. Smith (Eds.), Implicit Racial Bias Across the Law (pp. 229-243). Cambridge, MA: Cambridge University Press.

Smith, R. J., & Levinson, J. D. (2012). The Impact of Implicit Racial Bias on the Exercise of Prosecutorial Discretion. Seattle University Law Review, 35(3), 795-826.

Social Determinants of Health. (2012). Columbus, OH: The Kirwan Institute for the Study of Race and Ethnicity.

Sommers, S. R. (2006). On Racial Diversity and Group Decision Making: Identifying Multiple Effects of Racial Composition on Jury Deliberations. Journal of Personality and Social Psychology, 90(4), 597-612.

Sommers, S. R., & Ellsworth, P. C. (2000). Race in the Courtroom: Perceptions of Guilt and Dispositional Attributions. Personality and Social Psychology Bulletin, 26(11), 1367-1379.

Sommers, S. R., & Ellsworth, P. C. (2001). White Juror Bias: An Investigation of Prejudice Against Black Defendants in the American Courtroom. Psychology, Public Policy, and Law, 7(1), 201-229.

Stanley, D., Phelps, E., & Banaji, M. (2008). The Neural Basis of Implicit Attitudes. Current Directions in Psychological Science, 17(2), 164-170.

Steele, C. M. (1997). A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance. American Psychologist, 52(6), 613-629.

Steele, C. M., & Aronson, J. (1995). Stereotype Threat and the Intellectual Test Performance of African Americans. Journal of Personality and Social Psychology, 69(5), 797-811.

Steffens, M. C. (2004). Is the Implicit Association Test Immune to Faking? Experimental Psychology, 51(3), 165-179.

Stepanikova, I., Triplett, J., & Simpson, B. (2011). Implicit Racial Bias and Prosocial Behavior. Social Science Research, 40(4), 1186-1195.

Stevens, G. D., & Shi, L. (2003). Racial and Ethnic Disparities in the Primary Care Experiences of Children: A Review of the Literature. Medical Care Research and Review, 60(1), 3-30.

Stone, J., & Moskowitz, G. B. (2011). Non-Conscious Bias in Medical Decision Making: What Can Be Done to Reduce It? Medical Education, 45(8), 768-776.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behavior. European Journal of Social Psychology, 1(2), 149-178.

Tenenbaum, H. R., & Ruck, M. D. (2007). Are Teachers' Expectations Different for Racial Minority Than for European American Students? A Meta-Analysis. Journal of Educational Psychology, 99(2), 253-273.

Tetlock, P. E., & Mitchell, G. (2008). Calibrating Prejudice in Milliseconds. Social Psychology Quarterly, 71(1), 12-16.

Tinkler, J. E. (2012). Controversies in Implicit Race Bias Research. Sociology Compass, 6(12), 987-997.

Todd, A. R., Bodenhausen, G. V., Richeson, J. A., & Galinsky, A. D. (2011). Perspective Taking Combats Automatic Expressions of Racial Bias. Journal of Personality and Social Psychology, 100(6), 1027-1042.

Towles-Schwen, T. (2003). Choosing Social Situations: The Relation Between Automatically Activated Racial Attitudes and Anticipated Comfort Interacting with African Americans. Personality and Social Psychology Bulletin, 29(2), 170-182.

Uhlmann, E. L., & 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(2), 207-223.

van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R. W. (2010). The Implicit Prejudiced Attitudes of Teachers: Relations to Teacher Expectations and the Ethnic Achievement Gap. American Educational Research Journal, 47(2), 497-527.

Vanman, E. J., Saltz, J. L., Nathan, L. R., & Warren, J. A. (2004). Racial Discrimination by Low-Prejudiced Whites: Facial Movements as Implicit Measures of Attitudes Related to Behavior. Psychological Science, 15(11), 711-714.

von Hippel, W., Sekaquaptewa, D., & Vargas, P. (1997). The Linguistic Intergroup Bias as an Implicit Indicator of Prejudice. Journal of Experimental Social Psychology, 33(5), 490-509.

Weber, S. J., & Cook, T. D. (1972). Subject Effects in Laboratory Research: An Examination of Subject Roles, Demand Characteristics, and Valid Inference. Psychological Bulletin, 77(4), 273-295.

Wegener, D. T., & Petty, R. E. (1995). Flexible Correction Processes in Social Judgment: The Role of Naive Theories in Corrections for Perceived Bias. Journal of Personality and Social Psychology, 68(1), 36-51.

Weinstein, C. S., Tomlinson-Clarke, S., & Curran, M. (2004). Toward a Conception of Culturally Responsive Classroom Management. Journal of Teacher Education, 55(1), 25-38.

Weisbuch, M., Pauker, K., & Ambady, N. (2009). The Subtle Transmission of Race Bias via Televised Nonverbal Behavior. Science, 326(5960), 1711-1714.

Weisse, C. S., Sorum, P. C., Sanders, K. N., & Syat, B. L. (2001). Do Gender and Race Affect Decisions About Pain Management? Journal of General Internal Medicine, 16(4), 211-217.

Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., Wright, C. I., & Rauch, S. L. (2001). A Functional MRI Study of Human Amygdala Responses to Facial Expressions of Fear Versus Anger. Emotion, 1(1), 70-83.

White III, A. A. (2011). Seeing Patients: Unconscious Bias in Health Care. Cambridge, MA: Harvard University Press.

Wilson, T. D., & Dunn, E. W. (2004). Self-Knowledge: Its Limits, Value, and Potential for Improvement. Annual Review of Psychology, 55, 493-518.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A Model of Dual Attitudes. Psychological Review, 107(1), 101-126.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for Racial Prejudice at the Implicit Level and Its Relationship With Questionnaire Measures. Journal of Personality and Social Psychology, 72(2), 262-274.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous Prejudice in Context: Variability in Automatically Activated Attitudes. Journal of Personality and Social Psychology, 81(5), 815-827.

World Health Organization. (2008). Closing the Gap in a Generation: Health Equity Through Action on the Social Determinants of Health. Geneva: Final report of the Commission on Social Determinants of Health.