

The Evolving Science on Implicit Bias

An Updated Resource for the State Court Community



FUNDED BY



WRITTEN BY

Jennifer K. Elek
Senior Court Research Associate

Andrea L. Miller
Senior Court Research Associate

National Center for State Courts

This report was developed under grant number SJI-17-T-130 from the State Justice Institute. The points of view expressed are those of the authors and do not necessarily represent the official position or policies of the State Justice Institute.

Executive Summary

The judiciary is regarded by the public as a legitimate authority largely because of the perception of independence and impartiality. That perception is under threat. During the past turbulent year, public trust in government declined across the globe. As public trust declines, the ability of the judiciary to skillfully and effectively demonstrate the ideals of fairness and impartiality under law becomes ever more critical. In recognition of the need for leadership during such times, on July 30, 2020, the Conference of Chief Justices (CCJ) and the Conference of State Court Administrators (COSCA) passed a resolution “in support of racial equality and justice for all.” The resolution noted, in part, that “courts in many states, with the encouragement, support, and guidance of CCJ and COSCA, have initiated efforts... to identify and address unconscious bias, and facilitate the uncomfortable conversations that arise from the recognition of such bias.”

The terms *unconscious bias* and *implicit bias* emerged from research in the psychological and brain sciences. In everyday vernacular, they serve as shorthand labels for the notion widely supported by research evidence that social discrimination is like a virus: It can be easily and rapidly “caught” by a person from the social environment. This infection triggers an immune response: It influences the person’s thinking and behavior in that environment to reinforce existing patterns of social discrimination, often in ways the person does not fully appreciate or understand. Implicit bias both results from and reinforces different forms of inequality at multiple levels of society. Research on implicit biases addresses how they can arise in individual information processing, decision-making, and behavior in ways that reproduce, reinforce, and are reinforced by dynamics that are historical, cultural, institutional, and interpersonal in nature. A comprehensive and successful approach to implicit bias intervention must be one that considers the importance of this broader social context and addresses the full array of forces that contribute to observed inequities.

The present report defines commonly used terms originating from the science of implicit bias; explains how the concept of implicit bias fits into broader conversations underway across the country about equity and fairness; and summarizes what is currently known from research in the psychological and brain sciences, including implicit bias strategies generally found to be effective and ineffective. This report concludes with some implications of this knowledge for state court leaders and other court practitioners who seek to better understand and address the reproduction and perpetuation of systemic biases through this lens.



Acknowledgements

This report was made possible through the generous support of the State Justice Institute and the National Center for State Courts. We thank the numerous state judicial educators, judges, academic researchers, and others who shared their expertise and insights from their professional experiences to help inform the authors' approach to this report. Finally, authors thank Pamela Casey, David Rottman, Bob Phipps, Tracey Johnson, and Deirdre Roesch for their individual contributions along the way. Authors would especially like to thank Caisa Royer for her assistance with earlier drafts of portions of this report, particularly Appendix B.



Preface

An independent and impartial judiciary is a cornerstone of government in the United States. Judicial officers are duty-bound to uphold these ideals of fairness. The American Bar Association has promulgated a model Code of Judicial Conduct to provide guidance on these matters, which most states have adopted in some form.¹ Specific ethics rules address issues of fairness, such as Rule 2.2, which explicitly instructs judges to “perform all duties of judicial office fairly and impartially.” Rule 2.3 (A) calls for judges to perform these duties “without bias or prejudice.” Rule 2.3 (B) reads: “A judge shall not, in the performance of judicial duties, by words or conduct manifest bias or prejudice, or engage in harassment, including but not limited to bias, prejudice, or harassment based upon race, sex, gender, religion, national origin, ethnicity, disability, age, sexual orientation, marital status, socioeconomic status, or political affiliation, and shall not permit court staff, court officials, or others subject to the judge’s direction and control to do so.” Judges are also required to hold attorneys in proceedings accountable for conduct that does not meet this ethical standard.

Moreover, the judiciary is regarded by the public as a legitimate authority largely because of the perception of independence and impartiality. That perception is under threat. There is a growing distrust of government and one another in the United States, and this distrust makes problem-solving in the public interest harder.²

In 2020, civic life was fundamentally altered by the global COVID-19 pandemic, its impact on the economy, and a cultural awakening to systemic racism. This further challenged traditional assumptions about fairness in criminal justice and healthcare systems, employment, housing, and other social institutions, buoyed by a steady stream of inequities laid bare by and exacerbated

during the public health crisis.³ During this turbulent year, public trust in government declined across the globe.⁴ At a time when citizens and governments are called upon to cooperate and mobilize coordinated responses to a variety of global challenges, an erosion of trust risks devastating consequences.⁵ As public trust declines, the ability of the judiciary to skillfully and effectively demonstrate the ideals of fairness and impartiality under law becomes ever more critical.

In recognition of the need for leadership during such times, on July 30, 2020, the Conference of Chief Justices (CCJ) and the Conference of State Court Administrators (COSCA) passed a resolution “in support of racial equality and justice for all.”⁶ This resolution came in the wake of dozens of official statements by state courts in response to the deaths of Breonna Taylor, Ahmaud Arbery, and George Floyd.⁷ The resolution urged those organizations “to continue and to intensify efforts to combat racial prejudice within the justice system, both explicit and implicit, and to recommit to examine what systemic change is needed to make equality under the law an enduring reality for all”. It acknowledged that “current events have underscored the persistence in our society of institutional and structural racism resulting in policies and practices that disproportionately impact persons of color.” The resolution further noted that “courts in many states, with the encouragement, support, and guidance of CCJ and COSCA, have initiated efforts... to identify and address unconscious bias, and facilitate the uncomfortable conversations that arise from the recognition of such bias.”

This report was written to update previous work on implicit bias and assist the state courts in responding to this national call to action.

Table of Contents

I. Introduction.....	1
II. Key Terms and Definitions.....	2
III. Understanding Implicit Bias.....	4
Foundations.....	4
Origins.....	5
Pervasiveness and Impact.....	6
Situational Triggers.....	7
IV. The Role of Implicit Bias in Understanding Equality.....	9
V. The State of Science on Bias Interventions.....	13
General Interventions to Reduce Prejudice and Discrimination.....	13
Implicit Bias Interventions.....	14
Key Takeaways.....	18
VI. Implications for Courts and Their Communities.....	19
VII. Conclusion.....	26
Endnotes.....	27
Appendix A: Glossary of Terms.....	38
Appendix B: Implicit Measures.....	40

Introduction

By 2009, the terms *implicit bias* and *unconscious bias* were emerging from the psychological and brain sciences, entering the public conversation about how social and cultural biases can subtly influence the everyday behavior of individuals. The National Center for State Courts (NCSC) considered whether this information on subtle forms of bias would be of interest to the court community as part of the *Helping Courts Address Implicit Bias* pilot project, conducted in 2009-2012 with support from the State Justice Institute and the Open Society Foundations. Of the 108 justice system professionals participating in one of three state judicial education pilot programs conducted during the project, approximately 90% expressed satisfaction with their implicit bias program and believed it applicable to their work. Participants used words like “valuable,” “relevant,” “informative,” “worthwhile,” and “eye-opening” to describe their reactions to their respective programs.⁸ Participants demonstrated knowledge gains as measured using pre- and post-training surveys. The report called for further research to evaluate educational programs designed around this material to determine their efficacy in achieving their varied goals and understand the conditions under which this type of content may be effectively delivered to court community audiences.

Much has transpired since the original NCSC (2012) report was published, underscoring the need to update the court community on the current state of the science and practice related to unconscious or implicit biases. The present report aims to:

- a. equip court practitioners with current definitions of commonly used terms;
- b. explain how the concept of implicit biases fits into broader conversations underway across the country about equity and fairness;
- c. summarize what is currently known from implicit social cognition research in the psychological and brain sciences, including implicit bias strategies generally found to be effective and ineffective; and
- d. identify some implications of this knowledge for state court leaders and other court practitioners who seek to better understand and address the reproduction and perpetuation of systemic biases through this lens.

Key Terms & Definitions

Reflecting how ubiquitous the term *implicit bias* has become, many who use this language in public discourse today do not define it. Among those who do define it, many rely on imprecise definitions that are difficult to understand or that vary from one another. To communicate clearly about the state of the science of implicit bias, it is useful to first define several key terms as understood by scientists today. These terms are also listed in the **Glossary of Terms** (Appendix A of this report).

- **Bias:** The unintended influence of factors that are not meant to be considered on a final decision or result.⁹ Bias can occur either when relevant information does not influence the decision or when irrelevant information influences the decision. The particular situation or legal context surrounding a decision determines which factors are considered relevant or irrelevant.
- **Conscious:** Mental processes involving both awareness and volition.¹⁰
- **Unconscious:** Mental processes that lack either full awareness or full volition.¹¹
- **Explicit bias:** A bias that is measured using an explicit, or direct, measure.¹² **Explicit measures** require participants to self-report their responses. They rely on the assumption that individuals are aware of their responses and are willing to express them.¹³
- **Implicit bias:** A bias that is measured using an implicit, or indirect, measure.¹⁴ **Implicit measures** capture participants' responses in ways that do not rely on individuals' awareness or willingness to respond, such as by measuring reaction time to different groups of stimuli.¹⁵ The scientific field of study that uses these implicit or indirect measures in research

on attitudes, stereotypes, and self-esteem is classified as **implicit social cognition**.¹⁶

In contrast to current prevailing scientific definitions, the term *implicit bias* is rarely used in public discourse to refer to a specific measurement of bias. Instead, *implicit bias* and *unconscious bias* are often used synonymously to refer to an attitude, stereotype, or prejudice that a person is unaware of possessing but which may operate automatically to influence thinking or behavior. (Similarly, the term *explicit bias* is sometimes used to refer to a biased attitude, stereotype, or prejudice that a person is consciously aware of.) This can create confusion because researchers have concluded that there is “no evidence that people are unaware of the mental contents underlying their implicit biases.”¹⁷ In fact, when asked, people are able to predict the pattern of their implicit biases “to a high degree of accuracy.”¹⁸

Although people appear to be generally aware of their personal beliefs and cultural stereotypes (referred to as *content awareness*), they may not be aware of or fully understand how they developed this knowledge (referred to as *source awareness*), or how and to what extent that knowledge influences their everyday thinking and behavior (referred to as *impact awareness*).¹⁹ Researchers are still working to fully test relevant hypotheses and to develop a precise scientific understanding of the differences between biases documented using indirect versus direct measures.

In the meantime, the public (including the media and educators) continues to find *implicit bias* and *unconscious bias* useful terms.²⁰ They serve as shorthand labels for the notion widely supported by research evidence that social discrimination is like a virus: It can be easily and rapidly “caught” by a person from the social

environment. This infection triggers an immune response: It influences the person's thinking and behavior in that environment to reinforce existing patterns of social discrimination, often in ways the person does not fully appreciate or understand. For consistency with public understanding and clarity in this report, the term *implicit bias* will be used hereafter to refer to this shorthand label, rather than the more restrictive technical definition focused on its scientific measurement.

Understanding Implicit Bias

Foundations

We all accumulate a unique set of experiences in our lives that shape our perspectives about the world around us. But we are each limited in the information available to us about our world. Science tells us that what we experience is not what objectively exists, but what we are able to interpret based on the information we collect through our bodily senses. We do not have direct access to information about what others are feeling or thinking, but we use our observations about their facial expressions, tone of voice, choice of words, mannerisms, and other behavioral information to deduce what we can about them – and decide whether and how to interact with them. So when the human brain processes information, it is making predictions about, or a best guess at, what is going on in our external reality so we can decide how to act within it.²¹ These predictions are far from perfect, but they help us survive.

In addition to the subjective point of view that we cultivate through our experiences, our cognitive capacity to observe, think, and act is a finite resource. But research in the psychological and brain sciences paints a picture of a cognitive system with astonishing efficiencies built in. As we interact with the world, our mental machinery is designed to quickly search for patterns (e.g., certain types of small, spherical objects are *apples*) and make associations (e.g., *apples* are red, sweet, juicy, and are edible). Our brains do this between groups of people (e.g., older adults) and characteristics (e.g., slow, frail) as well. These associations occur, to some degree, automatically. Unlike **controlled** mental processes, which require at least some intention, effort, or conscious awareness to be enacted, **automatic** associations are formed without apparent mental effort; we may not be

consciously aware of or intend to make these associations.²² This automaticity in the human mind frees up our limited cognitive resources to perform other tasks. Because of this, we are generally not always fully aware of all the activity our minds are undertaking to help us detect, process, and act on information.

Although automatic associations make navigating the world possible, they are sometimes incorrect or even harmful. The problem is that when the brain automatically associates certain characteristics with specific groups, the association is not accurate for all members of the group. Following the above examples, not all apples are red; not all older adults are slow. Kang (2009) describes the problem this presents for the justice system:

Though our shorthand schemas of people may be helpful in some situations, they also can lead to discriminatory behaviors if we are not careful. Given the critical importance of exercising fairness and equality in the court system, lawyers, judges, jurors, and staff should be particularly concerned about identifying such possibilities. Do we, for instance, associate aggressiveness with Black men, such that we see them as more likely to have started the fight than to have responded in self-defense?²³

Our minds are constantly classifying incoming information into categories that have meaning to us. These categories may be meaningful because they are categories that society has defined for us or that we have learned from others over time. Embedded in the architecture of our daily lives, many of these associations can be, or have become, invisible to us. We may not endorse these associations, but they can nevertheless contaminate our choices and leak out through our behavior to impact others in ways that we do not intend.

Is it Implicit Bias?

Automaticity and control occur on a continuum. There are many forms of bias that may not fit neatly into a category of purely automatic or purely controlled. For example, **microaggressions** are brief, everyday exchanges

that send denigrating messages to certain individuals because of their membership in certain groups.²⁴ Microaggressions can be subtle, as they include verbal speech, non-verbal cues, and outward behaviors, but they can do substantial harm to their targets.²⁵ Because microaggressions are defined by how they are *experienced by the receiver*, as opposed to being defined by the *intentions of the actor*, they vary widely in the extent to which they involve any intent or knowledge on the part of the *actor* that he or she may be manifesting a bias. In other words, an individual can engage in a microaggression with full knowledge and intent to harm the receiver, or with a complete lack of awareness of the harm being done, or with some level of intent and knowledge that falls between these two extremes. Depending on these factors, a microaggression may or may not be an instance of *implicit bias*.



Researchers use several scientific methods to measure implicit bias.²⁶ For more information about various implicit measures used in research and educational settings, see **Appendix B** of this report.

Origins

Researchers believe that implicit biases have a few common origins, including the following.

1. Ingroup favoritism: We favor the familiar.

People tend to demonstrate preferences for their *ingroup*, or members of the groups to which they belong.²⁷ Favoritism can benefit, for example, the decision-maker's family members and friends, those who share the same political or religious ideology, or fans of the same sports teams. Although we tend to favor those who we think share our values, favoritism can result from any perceived similarity between the decision-maker and the person being judged – even when the similarity is superficial or coincidental. Why?

Scientists believe this occurs because we tend

to like things that are familiar, and nothing is more familiar to us than ourselves.²⁸ People demonstrate consistent and strongly positive attitudes toward themselves, and this positive attitude can transfer easily to other things, people, and groups when they bear a resemblance to those attributes.²⁹ For example, when choosing between products, people tend to prefer brands that resemble their own names.³⁰ Similarly, ingroup favoritism is often observed even among strangers in artificial research settings, based on seemingly random similarities (e.g., the person being judged was randomly assigned a code number that matched the decision-maker's birthdate, or was assigned to wear the same color shirt as the decision-maker in the research study).³¹ Bottom line: People categorize others very easily and quickly to determine how to interact with them – that is, whether they are “in” or “out.”

2. Social learning: We are taught, but also “catch,” biases from others.

Implicit biases can develop and strengthen over time with the accumulation of personal experience. Personal experiences include

not only direct learning experiences between ourselves and the object, person, or group (i.e., classical conditioning), but also by observing the behavior of parents, friends, bosses, coworkers, and other influential people in our lives (i.e., social learning).³² For example, children observing the behavior of adults interacting with one another will (a) indicate a preference for the adult who received positive treatment from the main speaker vs. the adult who received negative treatment; (b) choose to share resources (i.e., a teddy bear) with adults who received the positive treatment; (c) systematically imitate the adults who received positive treatment and shun those who received negative treatment from the main speaker; and (d) generalize these approach-avoid preferences to similar-looking others, illustrating the rapid and unintentional intergenerational transmission of social group bias.³³ Implicit biases in children are positively correlated with the implicit biases of their parents; however, consistent with social learning theory, this is found only among children who have a positive attachment relationship with their parents.³⁴ Implicit biases can develop relatively quickly through such experiences and have been found in children as young as 5 years old.³⁵

3. Cultural knowledge: Our beliefs are shaped by our environment.

Cultural preferences and expectations, including stereotypes, are communicated in a variety of ways. They are embedded in a society's laws;



upheld by government leaders; highlighted in the news; and reproduced in entertainment media such as movies, television, and video games. Society is structured around these cultural beliefs and values, which are baked into the formal and informal rules, social scripts, language, and symbols that people encounter, follow, or use every day. As a result, people develop a shared understanding of the social norms and stereotypes that are pervasive in their culture, and this cultural knowledge can foster the development of automatic associations.³⁶ Even if the attitudes we personally endorse differ or change over time, the implicit biases that arise from cultural knowledge can be resistant to change if those cultural stereotypes continue to be reproduced and reinforced throughout our social environment.³⁷ So long as representations of cultural stereotypes persist in our environment, people will have implicit biases reflecting those communicated preferences.

Pervasiveness and Impact

Implicit biases can influence a number of judgments and actions in professional settings, where they have significant impacts on people's lives.³⁸ In the legal domain, for example, researchers have demonstrated correlations between judges' implicit biases and their sentencing decisions,³⁹ as well as between labor arbitrators' implicit biases and their decision, in real arbitration cases.⁴⁰ Police officers' implicit biases correspond to their decisions to shoot criminal suspects of different races.⁴¹ In medicine, researchers have found correlations between medical providers' implicit biases toward their patients and the quality of care that patients receive,⁴² as well as between nurses' implicit biases and their likelihood of remaining in their jobs.⁴³ Employers' implicit biases correspond to their hiring decisions.⁴⁴

In addition to professional decision-making, implicit associations correspond to a variety of behaviors outside of the laboratory that affect people's experiences, behaviors, and life outcomes. For example, implicit biases have been linked to self-reported racially hostile behavior, such as the use of verbal slurs and

physical harm against people of color.⁴⁵ Voters' implicit attitudes about electoral candidates have been shown to predict election outcomes.⁴⁶ Scores on implicit measures can distinguish between adolescents who are likely to engage in suicide ideation or suicide attempts and those who are not,⁴⁷ and between sex offenders who commit crimes against children and those who commit crimes against adults.⁴⁸ Implicit associations have also been shown to correspond to substance use among those with addictions to drugs or alcohol.⁴⁹

Finally, there is a large body of research demonstrating other less-than-intentional biases and disparities in professional decision-making. These studies do not use implicit measures of bias, but they use experimental methods to identify disparities in decision-making that participants most likely are either not fully aware of or do not fully intend. For example, trial court judges in one research study decided a series of hypothetical cases; the facts of the cases were identical for all participants, except the social categories (e.g., gender, race) of the litigants involved.⁵⁰ Although most judges in the sample stated that they were confident in their abilities to make case decisions free from gendered and racial biases, these litigant characteristics had significant effects on case

outcomes. A wide variety of social and cognitive decision-making biases have been demonstrated using similar methods in other studies of judges,⁵¹ as well as in studies of professionals in other fields.⁵²

Situational Triggers

People are more likely to act in biased ways under certain conditions. Although a comprehensive review is beyond the scope of this report, a few common examples follow.

1. Situational incentives encourage speed over fairness and accuracy.

The performance that organizational leaders pay attention to and reward has an influence on what employees prioritize in their work. Some organizations do not provide employees with any meaningful feedback on their performance. In absence of feedback, people are less likely to remain vigilant for possible biases in their decision-making processes over time.⁵³

Many organizations provide some performance feedback to employees, but this is limited to what can be easily measured. What is easily measured (e.g., productivity) may not be what matters most to the organization or the



community it serves (e.g., quality). For example, organizations that emphasize efficiency measures over quality measures are motivating their employees to work faster, potentially at the expense of at least some degree of accuracy and fairness. People can process information faster, and produce more decisions, when they rely more on automatic associations and stereotypes. In these instances, the decision-maker develops inferences and expectations about the person or people being judged earlier on in the information-gathering process. However, those expectations bias the decision-maker's attention and memory in favor of stereotype-confirming evidence.⁵⁴ Moreover, biased expectations can influence how the decision-maker interacts with the person or people being judged, creating a self-fulfilling prophecy. That is, the decision-maker may then act in ways that elicit from others the very behaviors that would confirm his or her own biased expectations.⁵⁵

involve time pressure, that force a decision maker to form complex judgments relatively quickly, or in which the decision maker is distracted and cannot fully attend to incoming information all limit the ability to fully process information.⁵⁸ Decision makers who are rushed, stressed, distracted, or pressured are more likely to apply stereotypes – recalling facts in ways biased by stereotypes and making more stereotypic judgments – than decision makers whose cognitive capacities are not similarly constrained.

2. Clear criteria for making a good decision are absent.

Decision-making environments vary in the extent to which they provide structure and clarity for the person making the decision. When the basis for judgment is somewhat vague (e.g., situations that call for discretion; cases that involve the application of new, unfamiliar laws; decisions for which there is not a clear decision-making process laid out in advance), biased judgments are more likely. Without more explicit, concrete criteria for decision making, individuals tend to disambiguate the situation using whatever information is most easily accessible—including stereotypes.⁵⁶

3. Decisions are made in a distracting or otherwise stressful environment.

Tiring (e.g., long hours, fatigue), stressful (e.g., heavy, backlogged, or very diverse caseloads; loud construction noise; threats to physical safety; popular or political pressure about a particular decision; emergency or crisis situations), or otherwise distracting circumstances can adversely affect judicial performance.⁵⁷ Specifically, situations that

The Role of Implicit Bias in Understanding Inequality

Implicit biases are measured at the level of the individual person, so implicit bias education and interventions often focus on the individual. However, this phenomenon exists in a rich social and historical context. Implicit associations are both formed and expressed within that context, so taking context into account is crucial for the development of interventions. This section describes the different levels of inequality within which implicit biases operate.

1. Systemic inequality

Bias and inequality generally operate in ways that permeate multiple facets of society in multiple ways, and they tend to reinforce and re-create themselves over time.⁵⁹ The concept of systemic inequality captures this dynamic. **Systemic inequality** is the combination of a diverse array of discriminatory and inequitable practices in society, including the unjustly gained economic and political power of some groups over others, ongoing resource inequalities, ideologies and attitudes that regard some groups as superior to others, and the set of institutions that preserve the advantages of some groups over others.⁶⁰

2. Cultural inequality

Within a society, certain social groups have the power to define the culture's value system.⁶¹ **Cultural inequality** is the inequality that is "built into our literature, art, music, language, morals, customs, beliefs, and ideology" to such an extent that it defines "a generally agreed-upon way of life."⁶² Dominant culture dictates what is regarded as "good, bad, just, natural, desirable, and possible,"⁶³ while being presumed to be neutral and inclusive.⁶⁴ For example, the people who appear in movies, television shows, and advertisements in the United States are disproportionately slim, White,

able-bodied people with Eurocentric facial features. The fact that this is the agreed-upon standard of beauty in our culture is an example of cultural inequality.⁶⁵ Another example of cultural inequality is that Black men are over-represented as violent criminals in works of fiction, such as movies and television shows.

3. Institutional inequality

Institutional inequality refers to the network of institutional structures, policies and practices that create advantages and benefits for some groups over others.⁶⁶ Institutions can be defined broadly to include any collective body that influences social norms and the allocation of resources to individuals and social groups. Institutions can include the justice system, schools, media, banks, business, health care, governmental bodies, family units, religious organizations, and civic groups.⁶⁷ Institutional inequality can be intentional or unintentional; it often occurs as a result of decisions that are neutral on their face but have disparate impact with regard to race, gender, and other categories. Whether institutional inequality occurs in subtle ways or as a result of overt practices that limit the rights, mobility, or access of certain groups, the actions that lead to the disparity are sanctioned by the institution.⁶⁸ One example of institutional inequality in criminal justice is the use of different punishments for crack and powder cocaine, which are chemically similar but disproportionately used by different racial groups. This policy appears race-neutral on its face, but it results in African Americans receiving more severe punishments than White Americans for, effectively, the same drug use behavior.

4. Organizational inequality

Organizational inequality exists when

the practices, rules and policies of formal organizations (such as corporations or government agencies) result in different outcomes for different groups.⁶⁹ Like institutional inequality, organizational inequality can occur as a result of intentional decisions designed to produce different outcomes for different groups or as a result of policies and practices that appear neutral on their face.⁷⁰ For example, a retail corporation might have a policy that requires newly promoted supervisors to relocate to a new branch. Because women bear a disproportionate amount of housekeeping and caregiving duties and experience significant wage inequality compared to men, they are less likely to be able to relocate their families for their jobs. This policy appears gender-neutral on its face, but it results in greater promotion potential for men than for women, and it exacerbates the gender wage gap.

5. Interpersonal inequality

Interpersonal inequality exists when inequality manifests at the individual, person-to-person level. The term “bias” gets used in ways that conflate what are actually different psychological processes: stereotypes, prejudice, and discrimination. Any of these three processes can exist on a continuum from highly automatic (generally measured with implicit measures) to highly controlled (generally measured with explicit measures).

- **Stereotypes** are beliefs and opinions about the characteristics, attributes, and behaviors of members of a group (e.g., “soccer moms are energetic”).⁷¹ In other words, stereotyping is cognitive in nature. When one engages in the act of stereotyping, one assumes that because an individual belongs to a particular social group, the individual must share the characteristics of the group. When an individual automatically associates a particular trait with a particular social group in long-term memory (largely outside of conscious awareness), we can measure this association with an

implicit measure, and we refer to this association as an implicit stereotype.

- **Prejudice** is the emotion, attitude, or evaluation that a person feels about members of a particular social group (e.g., “I don’t like soccer moms”).⁷² In other words, prejudice is affective, or emotional, in nature. When an individual automatically associates a particular attitude or evaluation with a particular social group in long-term memory (largely outside of conscious awareness), we can measure this association with an implicit measure, and we refer to this association as implicit prejudice.
- **Discrimination** consists of treating people differently from others, based on their membership in a particular social group (e.g., “soccer moms cannot attend my party”).⁷³ In other words, whereas stereotyping is *cognitive* and prejudice is *affective*, discrimination is *behavioral* in nature. This differential treatment can range from fully intentional, controlled behavior, to fully automatic, unconscious behavior. It can occur as a result of stereotypes about the other group, prejudicial attitudes about the other group, or both.



Bringing It All Together

The different levels of inequality, from systemic inequality at the societal level to interpersonal inequality at the individual level, are closely intertwined. Although we conceptualize these levels separately for the purpose of defining them, the boundaries between these levels are not always clear-cut. Furthermore, many forms of bias operate across multiple levels. For example, microaggressions, described above, can take the form of verbal speech by an individual person (e.g., telling an Asian American who was born and raised in the U.S. that they speak English so well), or they can take the form of environmental conditions created by an organization (e.g., adorning the walls of a courthouse with portraits of influential figures from history who are exclusively White men).⁷⁴

Although court professionals are often taught to think about implicit bias as an individual, interpersonal phenomenon, implicit bias exists in the broader context of inequality and discrimination at multiple levels in society. Implicit associations form as a result of repeated exposure to certain stereotypes and attitudes about different social groups. In other words, implicit biases at the individual level are shaped by the history and culture of the society in which the person lives and the experiences and social interactions that the person has as a result

of that culture and history. For example, most White Americans hold automatic associations in memory between the concept of Black men and the concept of violent crime. The widespread existence of this automatic association is not a coincidence; it stems from the fact that Americans grow up in a society characterized by racial inequality at systemic, cultural, institutional and organizational levels, and in which popular media disproportionately depict Black men as violent criminals.

Thus, although implicit bias is typically measured at the level of the individual person, it is important to consider how it both results from and reinforces different forms of inequality at multiple levels of society. The field of implicit social cognition addresses how biases can arise in individual information processing, decision-making, and behavior in ways that reproduce and reinforce, and are reinforced by, dynamics that are historical, cultural, institutional, and interpersonal in nature. A comprehensive and successful approach to implicit bias intervention must be one that takes into account the importance of this broader social context and addresses the full array of forces that contribute to observed inequities.

The Multiple Levels of Inequality: Privilege

Privilege is a lens through which people view the world they live in, and it operates at multiple levels of inequality. **Privilege** is an unearned favored state conferred simply because of one's group membership.⁷⁵ Everyone has privilege in at least some domains of life, but some people experience privilege on more dimensions of their group identities than others (e.g., race, gender, sexual orientation, class, ability, nationality). When a person experiences privilege in a certain identity dimension (e.g., ability and disability), it does not mean that the person has not worked hard, has not suffered, or does not deserve what he or she has. It simply means that any hardships the person experiences are not exacerbated by a particular form of oppression (e.g., ableism).

Privilege is, by its nature, largely invisible to the people who hold it. Essentially, privilege is the collection of the things a person doesn't have to think about, that others do, simply because of their group identities. When individuals belong to a dominant group in society, their experiences and actions are regarded as normal and natural and are taken for granted. Their experiences are more likely to be reflected in popular culture



and more likely to be represented in government and other seats of power. They are less likely to think about their group identity throughout their day-to-day lives, because their group identity does not create barriers and difficulties for them on an ongoing basis. For this reason, privilege has also been defined as the “luxury of obliviousness.”⁷⁶ Privilege operates at multiple levels of inequality: it originates in the overarching group disparities that are characterized by systemic inequality; it is reflected, reinforced, and reproduced at the levels of culture, institutions, and organizations; and it forms the lens through which individuals view their world and their interpersonal interactions.

The State of the Science on Bias Interventions

Psychologists have made significant strides over the last 10 years toward understanding which intervention strategies are effective in reducing the expression of biases. The section that follows summarizes what was learned from recent implicit social cognition research and from consideration of relevant ideas drawn from the relevant broader research literature on prejudice and discrimination.

General Interventions to Reduce Prejudice and Discrimination

Over the course of several decades, psychology research has produced evidence to clearly support the effectiveness of some general interventions in reducing prejudice or discrimination at the interpersonal level. These interventions are not necessarily targeted at implicit bias, but they offer important insights regarding the factors that are likely to make an *implicit* bias intervention more successful.

1. Intergroup Contact

Intergroup contact is one of the most thoroughly researched prejudice interventions in social psychology. Originally articulated as a research hypothesis in 1954,⁷⁷ the contact hypothesis received support over several decades and hundreds of research studies.⁷⁸ The findings show that when members of different social groups interact with each other, reductions in prejudice and discrimination follow.⁷⁹

In this regard, researchers have also shown that not all intergroup contact is equally effective in reducing explicit prejudice. Contact situations that include all four of the following features have the greatest impact: 1) the groups are working toward a common goal, 2) the groups have equal status within the contact situation,

3) the situation allows individuals to get to know each other on an individual basis, and 4) the contact situation receives institutional support or support from the relevant authority figures.⁸⁰ Contact situations that include two or three of these factors also tend to reduce prejudice, but to a lesser degree. Research suggests that intergroup contact is effective because it increases a person's knowledge of the outgroup, decreases the level of anxiety that a person feels about interacting with members of the outgroup, and increases a person's empathy for members of the outgroup.

Researchers have begun to examine the effects of intergroup contact on *implicit* bias, but this area of research is relatively young. Findings suggest that intergroup contact may also be effective for reducing implicit prejudice.⁸¹ For example, one recent study showed that among non-Black physicians, the extent of interracial contact over several years in medical school predicted lower anti-Black prejudice (measured both explicitly and implicitly), while the number of hours spent in diversity training did not.⁸² In contrast to explicit prejudice, which is influenced by the *quality* of intergroup contact, implicit prejudice seems to be influenced only by the *quantity* of intergroup contact.⁸³



The Jigsaw Classroom

One specific intervention aimed at providing meaningful intergroup contact that has received significant attention from researchers in psychology is the “Jigsaw Classroom.” It has primarily been studied in educational settings but it can be implemented in any situation involving collaborative group work. In a Jigsaw Classroom set-up, students are divided into groups to complete a project, and each student is responsible for a particular portion of the final product.⁸⁴ In the first phase of the project, students work independently on their own portion of the project. In the second phase, students gather together with each of the students from the other groups who are assigned to the same portion of the project (these are called “expert groups”). Finally, in the third stage, students report their progress back to their own group, sharing their new-found expertise. The group then works together to finish the final product.



The Jigsaw Classroom is a particular type of intergroup contact that emphasizes some of the factors that are known to make intergroup contact more effective. Specifically, it puts students in a situation where they are working toward a common goal. It also puts each student on equal footing with the others, allowing each individual to serve as an “expert” in a particular portion of the project and requiring students to be interdependent on one another.⁸⁵ Research suggests that, in addition to providing some educational benefits, the Jigsaw Classroom increases individuals’ evaluations of outgroup members and decreases the extent to which individuals engage in the stereotyping of outgroup members.⁸⁶

2. Structure versus Discretion

A second major area of research on bias interventions has focused on the context in which individuals are acting and making decisions. A substantial body of research in social cognition shows that individual decision-making discretion makes room for bias and prejudice to manifest as discrimination and inequality.⁸⁷ Specifically, when individuals make decisions under conditions of limited structure, ambiguous decision-making procedures, or subjective criteria, they are more likely to make decisions that manifest their biases.⁸⁸ These effects can include changing the relative weights of the decision criteria (depending on the social group membership of the people targeted by the decision),⁸⁹ applying available options differently to members of different groups,⁹⁰ holding members of different groups to different standards,⁹¹ or changing the decision-making

procedures from one decision to the next.⁹²

The major implication of this research for bias interventions is that one way to reduce group disparities in decision-making is to limit individual discretion as much as possible.⁹³ Embedding structure in the decision-making process, specifying decision-making procedures as clearly as possible, and relying more extensively on criteria that can be measured objectively may limit the extent to which an individual’s biases can leak out into the final decision outcome.⁹⁴

Implicit Bias Interventions

In addition to these more general bias interventions, researchers have developed several interventions aimed specifically at implicit bias. Implicit bias interventions tend to fall into one of two categories, which this

section will discuss in turn. The first category includes interventions that attempt to retrain the underlying implicit association in memory. The interventions in this group tend to be impractical for most purposes outside of the laboratory, and they have demonstrated limited success. The second category includes interventions that leave the underlying association in memory intact but attempt to interrupt its outward expression (in other words, limit the extent to which the implicit association can leak out into the individual's decisions or behavior). The interventions in this group vary in the extent to which they are practical to implement outside the lab, and they show more promise in their effectiveness.

1. Interventions that Attempt to Change Associations in Memory

Some implicit bias interventions attempt to retrain the brain by changing the implicit association that exists in the individual's long-term memory. For example, if an individual automatically associates members of a particular racial group (e.g., Black) with a negative evaluation (e.g., bad, lazy) in memory, the intervention would attempt to re-write this automatic association as a relationship between the racial group and an opposing, positive evaluation (e.g., good, hardworking). From this body of research on *change* interventions, three main lessons have been learned.

Lesson #1: Some change strategies may slightly reduce the strength of negative implicit associations.

Researchers have examined several different approaches to re-writing the automatic associations in long-term memory. It is possible to reduce the strength of a negative implicit association, but the effectiveness of this strategy is inconsistent.

For example, evaluative conditioning is a procedure that involves repeated exposure to a new idea that runs counter to the person's automatic association. It teaches people to automatically link concepts together in memory

that were not linked together previously. For example, one study placed participants in front of a computer screen and showed them a series of images of Black faces paired with positive words, as well as White faces paired with negative words.⁹⁵ The idea is that prior to the study, many participants had a pre-existing automatic association between Black faces and negative evaluations (and, conversely, White faces and positive evaluations). Having these participants repeatedly view images presenting the opposite idea might lead their pre-existing associations to weaken, or, with enough exposure, become negated. In some, but not all, studies, researchers have found that participants' negative implicit associations with the social group in question decreased in strength after an evaluative conditioning activity.⁹⁶

Other methods for re-writing a negative automatic association in memory include exposing research participants to positive, counter-stereotypical exemplars (i.e., example members of the social group in question),⁹⁷ asking participants to imagine the perspective of members of the other group (imagined perspective-taking),⁹⁸ or inducing a positive emotion while participants consider members of the other group (emotion induction).⁹⁹

Meta-analyses (i.e., research studies that measure the effects of many research studies combined) of these intervention strategies have generally found that evaluative conditioning and counter-stereotypical exemplars are sometimes effective in reducing the strength of implicit racial biases to a small degree; in contrast, imagined perspective-taking, emotion induction, and other strategies are generally not effective.¹⁰⁰

Lesson #2: Reductions in implicit bias resulting from change interventions typically don't last long.

Even when researchers successfully reduce the strength of a person's negative implicit associations in memory, these changes typically do not last long enough to have an effect outside

the lab. Over the past few years, researchers have devoted more attention to measuring how long changes in implicit associations last after a single experimental intervention. Unfortunately, meta-analyses reveal that most reductions in negative implicit associations following these discrete interventions do not last longer than one or two days.¹⁰¹ There are a few notable exceptions to this time limit, but in each of these exceptions, the intervention itself took place over the course of weeks or months, rather than in a single experimental session.¹⁰²

Lesson #3: With change interventions, reductions in implicit bias typically don't alter downstream behavior.

Even when researchers are able to reduce the strength of a person's negative implicit associations in memory, this change typically does not affect downstream thoughts or behaviors. In order for implicit bias interventions to have a meaningful effect outside the lab, they must cause changes in the individual's decisions about, or behaviors toward, members of the social group in question. Unfortunately, meta-analyses of implicit bias intervention studies find that very few studies measure the impact of changes in implicit associations on downstream behaviors; those that do measure the impact of interventions after a delay tend to find no effects.¹⁰³ Recent research suggests that because implicit bias is intertwined with the culture and environment within which individuals are acting, the stability of our social environments makes it unlikely that small reductions in implicit associations in memory will manifest as noticeable reductions in prejudice and discrimination.¹⁰⁴

2. Interventions that Affect the Expression of Implicit Bias

Another broad class of implicit bias interventions are ones that attempt to bypass or disrupt the expression of the implicit association. Unlike the first type of strategy discussed above, this class of expression interventions leaves the underlying implicit association in memory intact. Instead of re-writing the association, the goal

is to limit the extent to which it can leak out in decisions and behaviors. For example, bypassing interventions may teach people how to prevent implicit associations from getting activated in the first place; disruption interventions may teach people how to override their automatic gut reactions or decisions with a more egalitarian response. Researchers have examined several different interventions that address the expression of implicit bias.¹⁰⁵ Generally, this class of interventions shows more promise than interventions that try to retrain the brain.

a) Bypassing interventions

One well-known bypassing intervention is commonly referred to as a *blinding* procedure. Blinding procedures are structural practices that block the transmission of information that would trigger decision-makers' implicit biases. This is the principle behind now-commonplace practices such as blind auditions, blind peer-review, and double-blind clinical studies.¹⁰⁶ Blinding procedures are already used at other decision points in the justice system process. For example, as part of the National Research Council's recommended best practices for conducting eyewitness identification lineups, the administering police officer should not know the identity of the suspect in the lineup.¹⁰⁷ Jurisdictions are already experimenting with race-blinding procedures as a technique to reduce disparate treatment in prosecutorial charging decisions.¹⁰⁸ Blinding procedures can be helpful, but it can sometimes be difficult or impractical to blind for all factors that may activate implicit bias.

b) Disruption interventions: harnessing intrinsic motivation

One approach to disrupting the influence of implicit associations involves activating the individual's egalitarian goals. Researchers have shown, for example, that people who are *intrinsically* motivated to avoid prejudiced responding are more successful at overriding their implicit biases in favor of more egalitarian responses.¹⁰⁹ An intrinsic motivation, as opposed to an *extrinsic* motivation, is one that

comes from within the person, as part of his or her personality or sense of self. Reminding these individuals of their intrinsic motivation to promote equality in the moment can help them override their implicit associations. However, this strategy is risky and can backfire. If the intervention produces an extrinsic motivation to promote equality (i.e., a motivation that is guided by the chance to earn external rewards, such as social approval, prestige, or financial/material gains), it can result in greater implicit bias.

c) Disruption interventions: harnessing interpersonal motives

A second approach to teaching people strategies for overriding their automatic associations involves interpersonal motives. Research suggests that people adapt their thoughts and behaviors in subtle ways to fit into the context they are in. In the domain of implicit bias, this can mean that the mere presence of people who belong to other social groups or who support egalitarian norms can result in less implicit bias and more egalitarian behavior.¹¹¹ For example, participants in one study either interacted with a Black experimenter or a White experimenter before completing a measure of implicit racial prejudice; those who interacted with the Black experimenter exhibited lower implicit bias.¹¹² Similar effects have been shown in research studies that did not measure implicit associations directly. A study of jury decision-making, for example, found that juries composed of White and Black jurors engaged in higher-quality deliberations and made more egalitarian verdict decisions than juries composed of only White members.¹¹³ Importantly, these effects were not limited to the Black jurors; White jurors engaged in better decision-making when they were in the presence of Black jurors.

d) Disruption interventions: forming new decision-making habits

A third approach to teaching people to override their automatic associations involves breaking the “habit” of a person’s automatic response. Treating implicit bias like a bad habit that can be broken involves the same kinds of

strategies a person would use to break any other bad habit (such as smoking or biting one’s fingernails).¹¹⁴ There are two types of strategies that researchers have found to be effective in breaking the habit of automatic biases. First, researchers have developed long-term educational experiences (often weeks or months long) that teach people to become aware of situations when they are most vulnerable to implicit biases, replace their automatic responses with a more egalitarian response, and practice the new egalitarian response until it becomes habit. Several versions of this approach have demonstrated effectiveness, often resulting in behavioral changes months or years after the intervention has ended.¹¹⁵

The second type of habit-breaking intervention involves getting individuals to establish a behavioral plan for deciding or responding in a future situation in which they may be prone to bias. *Implementation intentions* have shown promise in this context. An implementation intention is an “if-then” statement that lays out contingencies between a situation and a response (e.g., “If situation X is encountered, then I will initiate egalitarian response Y”).¹¹⁶ Researchers have found that participants who commit to an implementation intention in advance (for example, a plan to think “good” after seeing a Black face) are more likely to be able to override their automatic responses in favor of a more egalitarian response.¹¹⁷ Reasons why this strategy is effective include that it increases the individual’s commitment to the response,¹¹⁸ makes the response more accessible in the individual’s mind,¹¹⁹ makes the response more automatic and less effortful,¹²⁰ and helps shield the individual from the intrusion of unwanted thoughts.¹²¹

Conclusions and Takeaways

Meta-analyses of these intervention strategies suggest they show promise as tools for reducing the influence of implicit associations on decisions and behavior.¹²² Interventions that prevent the activation of implicit associations, leverage individuals' intrinsic egalitarian motivations, create diverse decision-making contexts with shared norms of equality, and help people break the habit of their automatic, biased responses have been shown to reduce disparities in subsequent decisions and behaviors, even when they are not meant to change the underlying implicit association in memory.

Although psychological research on bias interventions is still in a state of rapid change and advancement, it points to three key takeaways that have practical implications for courts and their communities.

Key Takeaway #1: General interventions that attempt to reduce prejudice and discrimination through positive, meaningful intergroup contact and by structuring discretionary decisions are still some of the most effective strategies for courts. Intergroup contact has been widely studied as a bias reduction strategy for over half a century. Engagement activities that include the following features have the greatest impact: 1) different groups are working toward a common goal, 2) the

groups have equal status in the activity, 3) the activity allows individuals to get to know each other on an individual basis, and 4) the activity receives institutional support or support from the relevant authority figures. In addition, researchers and practitioners have long known that greater structure in decision-making processes can limit opportunities for bias to infect decision outcomes.

Key Takeaway #2: Implicit bias interventions that attempt to change implicit associations in memory are not consistently effective.

Although some *change* interventions can reduce the strength of implicit associations in some contexts, they are difficult to implement outside the lab, have inconsistent effects that do not last longer than a few days, and tend not to change subsequent decisions and behaviors.

Key Takeaway #3: Implicit bias interventions that bypass or disrupt biased responding show more promise. Specifically, there is evidence to support *expression* interventions that prevent the activation of implicit associations, leverage individuals' intrinsic egalitarian motivations, create diverse decision-making contexts with shared norms of equality, and give people tools to break the habit of their automatic, biased responses.



Implications for Courts and Their Communities

Recent developments in the science of implicit bias have implications for the state courts and their communities, where court leaders and other practitioners seek to better understand and address the reproduction and perpetuation of systemic biases through this lens. Four interrelated implications are discussed below.

1. **Lead by example - and know where you're headed.**

The science of implicit bias highlights the influence of the social environment on our thinking, decisions, and behavior. The court is a specific type of social environment, with a unique institutional culture, formal rules, and informal social norms that create expectations about appropriate behavior. Interpersonal influences, such as the conduct of leadership, play an important role in constructing that social environment.¹²³ Attitudes, preferences, and behavior may be readily learned or “caught” by observing the conduct of respected authorities and peers. And, as previously noted, strategies (such as intergroup contact) may be more effective at reducing prejudice and discrimination when implemented under certain conditions. One of those conditions is clear institutional support or support from relevant authority figures.¹²⁴

Court leaders can influence the social environment of the court in a variety of ways. In the wake of the killing of George Floyd, one of the most common ways organizational leaders across the country responded was by issuing a public statement to demonstrate social accountability. Dozens of state court leaders

also issued public statements, reaffirming commitments to identifying and addressing systemic injustices.¹²⁵ Over the years, the Conference of Chief Justices and the Conference of State Court Administrators have passed numerous policy resolutions on issues pertaining to equal justice.¹²⁶ States have established leadership teams and task forces charged with overseeing such activities.¹²⁷ Some states have created centralized Diversity, Equity, & Inclusion (DE&I) offices to, for example, create DE&I goals and implementation plans, coordinate activities and programs, establish metrics for measuring progress, and monitor DE&I goals.¹²⁸ Such leadership efforts can be valuable first steps for establishing a court culture that supports DE&I initiatives and produces concrete reforms that materially improve the lives of court users who have not historically been equally served.



A State Supreme Court Speaks with One Voice

"As members of the judicial branch, we are cautious – always careful not to prejudge situations. But we cannot ignore the risks that African Americans, Blacks, and other people of color face as each day dawns. The urgency for action has long been upon us, but the immediacy of the need is even more apparent today. We must ensure that the lives of African Americans, Blacks, and people of color are valued and respected and that the color of peoples' skin does not affect their rights to justice or the treatment they are afforded by our system of justice... Our courts are an integral part of the justice system and have an essential role to play in ensuring justice for all. We must stand firm against racism and oppression. We must be intentional in our efforts to move in a different direction. We must examine our individual thoughts and beliefs, as well as our professional approaches, processes, and environments to address the impact of our own biases. We must examine, anew, what we are doing, or failing to do, to root out conscious and unconscious bias in our legal system."

Oregon Supreme Court, June 5, 2020¹²⁹



Ultimately, judicial leadership must determine the goals of institutional efforts to address systemic and implicit biases. Terms such as diversity, inclusion, equality, and equity are often used interchangeably, but these terms represent different goals that implicate different strategies.

- **Diversity** refers to the presence of individuals who represent a variety of groups or perspectives. It captures the quantitative representation of different groups, but it does not capture how much each group is heard or how much influence each group has.¹³⁰
- **Inclusion** refers to the meaningful involvement of people from different groups, or the extent to which diverse perspectives are incorporated into systems, processes, and decisions.¹³¹ In contrast to diversity, which reflects the quantity of representation, inclusion reflects the quality of representation.
- **Equality** refers to the equal treatment of different individuals or groups; it occurs when people receive the same treatment or distribution of resources,

regardless of their needs or starting positions.¹³² An equality mindset assumes that everyone will benefit from the same supports to meet their needs.¹³³

- **Equity** refers to the state that exists when we cannot predict outcomes based on a person's group membership, and outcomes for all groups are improved. Equity often involves the differential treatment of different individuals, based on their needs and starting positions, with the goal that everyone will arrive at the same outcome.¹³⁴

Keeping in mind these different end goal states, court leaders will need to articulate objectives using appropriate terminology. Depending on the nature of the problem and the stated needs of the people who are most affected by the problem, the court might strive to achieve diversity, inclusion, equality, or equity. In recent years, members of the public, justice partners, and stakeholders are increasingly pushing to achieve equity.¹³⁵

2. Educate not just to raise awareness, but to build capacity for change.

As succinctly noted by Ivuoma Onyeodor and colleagues, “No one-size-fits-all solution addresses organizational diversity. Although implicit bias trainings can help address diversity, equity, and inclusion, they are not sufficient. Organizations must also change structures and improve climate.”¹³⁶ On their own, implicit bias educational sessions can realistically achieve only so much. They can add value when used to raise awareness and educate staff about bias and inequality, and train staff on specific individual and organizational strategies to be implemented in pursuit of organizational goals.

Much ink has been spilled over the efficacy of corporate diversity training efforts in recent years. As these trainings have proliferated, researchers have studied various programs for evidence of effectiveness. Historically, quality field research on diversity training has been relatively scarce. Research on educational interventions can be challenging to conduct for a variety of reasons, limiting conclusions researchers can draw about them collectively (e.g., substantial differences in program design, including program goals, featured content, methods, trainers used, type of audience, and more, make comparisons difficult; barriers in data collection, such as the inability to locate participants for follow-up evaluations, can preclude meaningful evaluations; other intervening factors can complicate interpretation of results). The evidence to date suggests that most corporate diversity training programs, including various implicit bias programs, are generally ineffective at reducing bias and inconsistent at changing behavior.¹³⁷ Part of this is because diversity training is often offered as a single, standalone training session or workshop. Experts suggest the efficacy of diversity trainings can be improved by incorporating training into a broader comprehensive strategy aimed at building capacity for change.¹³⁸ Specific recommendations include: tailoring training programs to match institutional goals, linking content to desired outcomes; preparing trainers

to manage participant discomfort as part of the learning process, rather than trying to avoid discomfort; training attendees on how to use a limited number of concrete strategies for managing bias (i.e., 2-3 specific strategies) that are most relevant to their work; and, importantly, developing a plan for evaluating the efficacy of the training.¹³⁹

Although it is unlikely that an implicit bias educational program will change attendees' implicit biases, it may offer other benefits.¹⁴⁰ For example, as a result of participation, people may become more aware of, concerned about, and motivated to address discrimination, become more sensitive to the biases of others and more likely to label biases as wrong, and have more confidence in their ability to effectively engage in equity-promoting behaviors. Studies of one educational curriculum designed by academic researchers at the University of Wisconsin-Madison have found that participants, when contacted several weeks after completing a version of the course, expressed greater concern about discrimination and reported feeling more comfortable discussing the issues. Two years after their seminars, those who completed the program were better able to identify biased behavior in others or were more likely to publicly object to others' expressions of bias. Such changes may be valuable to those seeking to motivate and engage professional staff in DE&I efforts to advance specific organizational goals.

3. Gather information to understand what is really happening in your court and community.

Disparities in the justice system vary by jurisdiction and decision point. Because of meaningful variations in the social environment (such as the composition of local communities, local history and politics, and other factors), it is important to collect data that can shed light on the specific types, direction, and magnitude of disparities – and their root causes – in a particular jurisdiction.¹⁴¹

Understanding the nature of the problem is crucial for determining which types of interventions are needed and which are likely to be successful. The more the court can use data to inform its strategy, the better positioned it will be to channel resources toward the interventions with the biggest impact. The following are some broad questions court personnel might consider when defining the problem:

1. What is the specific disparity or hardship we are trying to address?
2. Do we have enough information about the size and scope of the problem, or do we need more information?
3. What kinds of data are needed (e.g., case processing or case outcome data, employment data, anonymous surveys of the affected populations, town halls or targeted listening sessions with the affected populations)?
4. At which points in the justice process does this disparity or hardship emerge? Which components of the

problem are under the direct control of the court, which components could the courts influence through its role as a convener of important social institutions, and which components are outside the control of the court?

5. Who are the people most affected by this disparity or hardship?
6. What do the people in this group say they need?

After considering these questions, the court should determine whether the problem is one of internal organizational culture, public outreach and communication, individual decision-making, or policy. Often, more than one of these domains will be in play simultaneously. Depending on the nature of the problem, an implicit bias intervention may or may not offer the best possible solution.

Executive Order on Advancing Racial Equity and Support for Underserved Communities Through the Federal Government

On January 20, 2021, President Joseph Biden signed an executive order establishing that the “Federal Government should pursue a comprehensive approach to advancing equity for all, including people of color and others who have been historically underserved, marginalized, and adversely affected by persistent poverty and inequality.” In the executive order, the President acknowledged that “[b]ecause advancing equity requires a systematic approach to embedding fairness in decision-making processes, executive departments and agencies [...] must recognize and work to redress inequities in their policies and programs that serve as barriers to equal opportunity.” The executive order continued, calling on all federal agencies to “assess whether, and to what extent, its programs and policies perpetuate systemic barriers to opportunities and benefits for people of color and other underserved groups. Such assessments will better equip agencies to develop policies and programs that deliver resources and benefits equitably for all.”¹⁴²

Examining data is an essential step in uncovering disparities and bias, but many courts face challenges when seeking to document disparities in the justice system. This typically involves the lack of quality data and the fragmented nature of data systems. Common reported barriers to court collection of race and ethnicity data, for example, include the lack of staff time, limitations in technology systems, confusion about race and ethnicity categories, and concerns about data being misused or misinterpreted.¹⁴³ However, courts have undertaken significant efforts in this area.

Court leaders in some states have conducted or commissioned disparity analyses using administrative data. For example, the late Massachusetts Supreme Court Chief Justice Ralph Gants recently commissioned a report from researchers at Harvard Law School focused on documenting racial disparities in the state criminal court process using administrative data from multiple state agencies and survey data from the U.S. Census Bureau.¹⁴⁴ Other states have commissioned other types of reviews. For example, the New York State Courts recently charged an independent commission with comprehensively examining and documenting institutional racism in the state court system to inform and guide current improvement efforts.¹⁴⁵ This commission adopted a multimethod approach involving document reviews of institutional policies, programs, and practices and historical reports on issues of racial bias; reviews of employment statistics; engagement with various court and community stakeholders through numerous interviews and solicitation of written submissions; in-person court observations; and more to generate recommendations for improvement. Finally, scientists have started using the audit method to conduct large scale field experiments to document discrimination. This method has been widely used to provide the bulk of the evidence of discrimination in housing, employment, healthcare, education, and the delivery of other public services.¹⁴⁶ This may be another useful method to consider.

It is important to note that bias can also influence how data are interpreted and how decisions are made based on those interpretations. Some justice system leaders have sought to address this challenge by engaging diverse perspectives, including the voices of directly impacted communities, in collaborative efforts to determine what information to collect, identify disparities and understand their root causes, brainstorm potential solutions, and decide how best to address problems and allocate resources to promising strategies.¹⁴⁷

4. Experiment: Design interventions based on the evidence and evaluate interventions for efficacy.

Once the court identifies specific disparities that may benefit from a bias intervention, interventions should be customized to address the problem at the targeted decision point. The custom intervention should be pilot tested and evaluated to determine its effectiveness.



On the Importance of Evaluation: An Illustration

In a recent book, former Google data scientist Seth Stephens-Davidowitz summarized some of his research on what Google searches reveal about our stereotypes and biases:¹⁴⁸

Consider what happened shortly after the mass shooting in San Bernardino, California, on December 2, 2015. That morning, Rizwan Farook and Tashfeen Malik entered a meeting of Farook's coworkers armed with semiautomatic pistols and semiautomatic rifles and murdered fourteen people. That evening, literally minutes after the media first reported one of the shooters' Muslim-sounding name, a disturbing number of Californians had decided what they wanted to do with Muslims: kill them.

The top Google search in California with the word "Muslims" in it at the time was "kill Muslims." And overall, Americans searched for the phrase "kill Muslims" with about the same frequency that they searched for "martini recipe," "migraine symptoms," and "Cowboys roster." In the days following the San Bernardino attack, for every American concerned with "Islamophobia," another was searching for "kill Muslims." [...]

Four days after the shooting, then-president Obama gave a prime-time address to the country. He wanted to reassure Americans that the government could both stop terrorism and, perhaps more important, quiet this dangerous Islamophobia.

Obama appealed to our better angels, speaking of the importance of inclusion and tolerance. The rhetoric was powerful and moving. The Los Angeles Times praised Obama for "[warning] against allowing fear to cloud our judgment." The New York Times called the speech both "tough" and "calming." The website ThinkProgress praised it as "a necessary tool of good governance, geared towards saving the lives of Muslim Americans." Obama's speech, in other words, was judged a major success. But was it?

Google search data suggests otherwise. [...] In his speech, the president said, "It is the responsibility of all Americans – of every faith – to reject discrimination." But searches calling Muslims "terrorists," "bad," "violent," and "evil" doubled during and shortly after the speech [and] searches for "kill Muslims" tripled [...]. In fact, just about every negative search we could think of to test regarding Muslims shot up during and after Obama's speech, and just about every positive search we could think of to test declined.

In other words, Obama seemed to say all the right things. All the traditional media congratulated Obama on his healing words. But new data from the internet [...] suggested that the speech actually backfired in its main goal. Instead of calming the angry mob, as everybody thought he was doing, the internet data tells us that Obama actually inflamed it. Things that we think are working can have the exact opposite effect from the one we expect. Sometimes we need [...] data to correct our instinct to pat ourselves on the back.

Depending on the nature of the problem, one or more implicit bias intervention strategies may be appropriate. Approaches may include the following.

- Redesign the decision-making environment to **remove or minimize situational triggers** of implicit bias and create conditions for success.
- **Structure decision-making processes** to bypass or disrupt the expression of implicit biases, such as by establishing clear decision-making criteria before evidence is presented and a decision is made or incorporating blinding procedures.
- Cultivate opportunities for staff to **engage in positive, meaningful intergroup contact**. In addition to the discrimination-reduction benefits of these activities, under certain conditions, the sharing of diverse perspectives can produce other performance-enhancing benefits in the form of greater creativity, more innovation, and better decisions.¹⁴⁹ Intergroup contact may be increased, for example, by:
 - » being intentional about the composition of program committees, task forces, and other decision-making bodies;
 - » fostering workforce diversity by improving recruitment, hiring, retention, and promotion processes; and
 - » building staff communication skills to navigate crucial conversations and prepare them to conduct community outreach and engagement work.¹⁵⁰
- Equip individuals with the tools for improved decision making, such as by training them to develop **new decision-making habits**, such as stereotype replacement, implementation intentions, or other techniques.¹⁵¹

Counterstereotypes and Fostering Curiosity about Others: An Illustration

Seth Stephens-Davidowitz (2017) writes:¹⁵²

Let's return to Obama's speech about Islamophobia. Recall that every time Obama argued that people should respect Muslims more, the very people he was trying to reach became more enraged.

Google searches, however, reveal that there was one line that did trigger the type of response then-president Obama might have wanted. He said, 'Muslim Americans are our friends and neighbors, our coworkers, our sports heroes and, yes, they are our men and women in uniform, who are willing to die in defense of our country.'

After this line, for the first time in more than a year, the top Googled noun after 'Muslim' was not 'terrorists,' 'extremists,' or 'refugees.' It was 'athletes,' followed by 'soldiers.' And, in fact, 'athletes' kept the top spot for a full day afterward.

When we lecture angry people, the search data implies that their fury can grow. But subtly provoking people's curiosity, giving new information, and offering new images of the group that is stoking their rage may turn their thoughts in different, more positive directions.

Conclusion

The science on implicit bias is still evolving, and researchers and practitioners continue to assess its implications for broader efforts aimed at reducing prejudice and discrimination and improving equality, equity, diversity, and inclusion. As observed by neuroscientist and psychologist Lisa Feldman Barrett:¹⁵³

And now we get to the toughest issue of all: what it means to control your behavior and therefore be responsible for your actions. The law (like much of psychology) usually considers responsibility in two parts: actions caused by you, where you have more responsibility, and actions caused by the situation, where you have less.

[But] the concepts in your head are not purely a matter of personal choice. [...] They are forged by the social reality you live in. [...] You learn from the environment like any other animal. Nevertheless, all animals shape their own environment. So as a human being, you have the ability to shape your environment to modify your conceptual system, which means that you are ultimately responsible for the concepts that you accept and reject.

In the meantime, courts leaders forge ahead, continuing to make the best decisions they can with the knowledge we have today. On June 9, 2020, days after the death of George Floyd, Connecticut Chief Justice Richard A. Robinson wrote:¹⁵⁴

The existing imperfections in our justice systems have profound and lasting effects on all of us, but it is more severe on those of us who are the most vulnerable. There is a need for real and immediate improvement. America can - and must - do a better job of

providing “equal justice under law,” the very words that are engraved on the front of the United States Supreme Court Building in Washington, D.C. [...]

Many of you have heard me talk about race, implicit bias and my own life experiences facing these issues. Many of you have attended Judicial Branch training and programs that were designed to help us deal with these issues in our own lives and in order to fulfill the mission of the Branch to serve the interests of justice and the public by resolving matters brought before it in a fair, timely, efficient and open manner.

I am proud of the work that we have started, but there is so much more to do. I know that I am asking a lot of you. I know that you are tired, you are weary and maybe even rightfully disillusioned, but this is a battle for the nation's soul. We must double and even triple our efforts to provide equal justice for all those whom we serve. We have but two choices: to keep working hard and succeed; or to quit and fail. As for me, the latter is not an option.

Endnotes

- ¹ American Bar Association (2020). *ABA Model Code of Judicial Conduct*. Chicago: Authors. Available at https://www.americanbar.org/groups/professional_responsibility/publications/model_code_of_judicial_conduct/. For information about the jurisdictional adoption of a revised model code of judicial conduct, see https://www.americanbar.org/groups/professional_responsibility/resources/judicial_ethics_regulation/map/.
- ² Rainie, L., & Perrin, A. (July 22, 2019). Key findings about Americans' declining trust in government and each other. Available at <https://www.pewresearch.org/fact-tank/2019/07/22/key-findings-about-americans-declining-trust-in-government-and-each-other/>.
- ³ Engel, L., Abaroa-Ellison, J., Jemison, E., & Cumberbatch, K. (November, 2020). Racial disparities and COVID-19. Washington, D.C.: Council on Criminal Justice.
- ⁴ Edelman (2021). 2021 Edelman Trust Barometer. Available at <https://www.edelman.com/trust/2021-trust-barometer>
- ⁵ For example, Americans who had less trust in social institutions were less willing to get vaccinated during the global COVID-19 pandemic. The COVID States Project (September 2020). The state of the nation: A 50-state COVID-19 survey. Report #13: Public trust in institutions and vaccine acceptance. Available at <https://covidstates.org/>.
- ⁶ Conference of Chief Justices and Conference of State Court Administrators (July 30, 2020). Resolution 1: *In support of racial equality and justice for all*. Available at: https://www.ncsc.org/_data/assets/pdf_file/0029/42869/07302020-Racial-Equality-and-Justice-for-All.pdf
- ⁷ National Center for State Courts, *State court statements on racial justice*. Available at: <https://www.ncsc.org/newsroom/state-court-statements-on-racial-justice>.
- ⁸ p. 21 in National Center for State Courts (2012), *Helping courts address implicit bias: resources for education*. Williamsburg VA: Authors.
- ⁹ Evans, J. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 225-278. Available at <https://www.annualreviews.org/doi/10.1146/annurev.psych.59.103006.093629>.
- ¹⁰ Evans, J. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 225-278. Available at <https://www.annualreviews.org/doi/10.1146/annurev.psych.59.103006.093629>.
- ¹¹ Evans, J. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 225-278. Available at <https://www.annualreviews.org/doi/10.1146/annurev.psych.59.103006.093629>.
- ¹² Greenwald, A., & Lai, C. (2020). Implicit social cognition. *Annual Review of Psychology*, 71, 419-445. Available at <https://www.annualreviews.org/doi/10.1146/annurev-psych-010419-050837>.
- ¹³ Olson, M. & Zabel, K. (2016). Measures of prejudice. In Nelson, T. (Ed.), *Handbook of Prejudice, Stereotyping, and Discrimination* (2nd ed.). Oxfordshire: Taylor & Francis Group.
- ¹⁴ Greenwald, A., & Lai, C. (2020). Implicit social cognition. *Annual Review of Psychology*, 71, 419-445. Available at <https://www.annualreviews.org/doi/10.1146/annurev-psych-010419-050837>.
- ¹⁵ Olson, M. & Zabel, K. (2016). Measures of prejudice. In Nelson, T. (Ed.), *Handbook of Prejudice, Stereotyping, and Discrimination* (2nd ed.). Oxfordshire: Taylor & Francis Group.
- ¹⁶ Greenwald, A., & Banaji, M. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4-27. In this seminal publication, the authors proposed the term and explained that the "signature of implicit social cognition is that traces of past experience affect some performance, even though the influential earlier experience is not remembered in the usual sense – that is, it is unavailable to self-report or introspection" (p. 4).
- ¹⁷ Gawronski, B. (2019). Six lessons for a cogent science of implicit bias and its criticism. *Perspectives on Psychological Science*, 14, 574-595.
- ¹⁸ Hahn, A., Judd, C., Hirsh, H., & Blair, I. (2014). Awareness of implicit attitudes. *Journal of Experimental Psychology: General* (43), 1369-1392.

- 19 Gawronski, B. (2019). Six lessons for a cogent science of implicit bias and its criticism. *Perspectives on Psychological Science*, 14, 574-595.
- 20 E.g., Benson, T., & Fiarman, S. (2020). *Unconscious bias in schools: A developmental approach to exploring race and racism*. Cambridge, MA: Harvard Education Press.
- 21 E.g., see Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Boston: Houghton Mifflin Harcourt.
- 22 Correll, J., Judd, C.M., Park, B., & Wittenbrink, B. (2010). Measuring Prejudice, stereotypes, and discrimination. In *The SAGE Handbook of Prejudice, Stereotyping, and Discrimination*. Thousand Oaks: SAGE Publications. Devine, P. G., & Sharp, L. B. (2009). Automaticity and control in stereotyping and prejudice. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (p. 61-87). Psychology Press.
- 23 Kang, J. (2009). *Implicit bias: A primer for courts*. Williamsburg, VA: National Center for State Courts.
- 24 Sue, D., Bucceri, J., Lin, A., Nadal, K., & Torino, G. (2007). Racial microaggressions and the Asian American experience. *Cultural Diversity and Ethnic Minority Psychology*, 13, 72-81. Sue, D., Capodilupo, C., & Holder, A. (2008). Racial microaggressions in the life experience of Black Americans. *Professional Psychology: Research and Practice*, 39, 329-336.
- 25 Blume, A., Lovato, L., Thyken, B., & Denny, N. (2012). The relationship of microaggressions with alcohol use and anxiety among ethnic minority college students in a Historically White Institution. *Cultural Diversity and Black Minority Psychology*, 18, 45-54. Harwood, S., Hunt, M., Mendenhall, R., & Lewis, J. (2012). Racial microaggressions in the residence halls: Experiences of students of color at a Predominantly White University. *Journal of Diversity in Higher Education*, 5, 159-173. Huynh, V. (2012). Ethnic microaggressions and the depressive and somatic symptoms of Latino and Asian American Adolescents. *Journal of Youth and Adolescence*, 41, 831-846. Nadal, K., Wong, Y., Griffin, K., Davidoff, K., & Sriken, J. (2014). The adverse impact of racial microaggressions on college students' self-esteem. *Journal of College Student Development*, 55, 461-474. Sue, D., Capodilupo, C., & Holder, A. (2008). Racial microaggressions in the life experience of Black Americans. *Professional Psychology: Research and Practice*, 39, 329-336. Torres, L., Driscoll, M., & Burrow, A. (2010). Racial microaggressions and psychological functioning among highly achieving African-American: A mixed-methods approach. *Journal of Social and Clinical Psychology*, 29, 1074-1099.
- 26 For reviews, see Hahn, A., & Gawronski, B. (2018). Implicit social cognition. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, 4, 1-33. 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. Gawronski, B. (2009). Ten frequently asked questions about implicit measures and their frequently supposed, but not entirely correct answers. *Canadian Psychology/Psychologie canadienne*, 50(3), 141. Wittenbrink, B., & Schwarz, N. (Eds.). (2007). *Implicit measures of attitudes*. Guilford Press.
- 27 Azevedo, R. T., Macaluso, E., Avenanti, A., Santangelo, V., Cazzato, V., & Aglioti, S. M. (2013). Their pain is not our pain: brain and autonomic correlates of empathic resonance with the pain of same and different race individuals. *Human brain mapping*, 34(12), 3168-3181.
- 28 E.g., Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, 33, 1-39. Turner, J.C., Brown, R.J., & Tajfel, H. (1979). Social comparison and group interest in ingroup favouritism. *European Journal of Social Psychology*, 9, 187-204. Tajfel, H., Billig, M., Bundy, R., & Flament, C. (1971). Social categorization and intergroup behavior. *European Journal of Social Psychology*, 1, 149-177. Zajonc, R. (2001). Mere exposure: a gateway to the subliminal. *Current Directions in Psychological Science*, 10, 224-228. Zajonc, R. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9, 1-27.
- 29 For a review, see Banaji, M. R., & Heiphetz, L. (2010). Attitudes. In: S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (pp. 348-388). New York: John Wiley & Sons.
- 30 Brendl, C. M., Chattopadhyay, A., Pelham, B. W., & Carvallo, M. (2005). Name letter branding: Valence transfers when product specific needs are active. *Journal of Consumer Research*, 32(3), 405-415.
- 31 E.g., see Billig, M. & Tajfel, H. (1973). Social categorization and similarity in intergroup behaviour. *European Journal of Social Psychology*, 3, 27-52. Molenberghs, P. (2013). The neuroscience of in-group bias. *Neuroscience & Biobehavioral Reviews*, 37, 1530-1536.
- 32 Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science*, 12(5), 413-417. Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review*, 102(1), 4.
- 33 Skinner, A., Meltzoff, A., & Olson, K. (2017) "Catching" social group bias: Exposure to biased nonverbal signals creates social biases in preschool children. *Psychological Science*, 28, 216-224. Skinner, A., Olson, K., & Meltzoff, A. (2020). Acquiring group bias: Observing other people's nonverbal signals can create social group biases. *Journal of Personality and Social Psychology: Interpersonal Relations and Group Processes*, 119, 824-838. Skinner, A., & Meltzoff, A. (2017). How societal prejudices seep into the minds of our children. *UNESCO IBE In Focus: Education & the Future*, 98-101.
- 34 Bandura, A. (1997). *Self-efficacy: The exercise of control*. Macmillan. Sinclair, S., Dunn, E., & Lowery, B. (2005). The relationship between parental racial attitudes and children's implicit prejudice. *Journal of Experimental Social Psychology*, 41(3), 283-289.

- 35 Williams, A., & Steele, J. R. (2019). Examining children's implicit racial attitudes using exemplar and category-based measures. *Child development*, 90(3), e322-e338. See also Mandalaywala, T., Ranger-Murdoch, G., Amodio, D., & Rhodes, M. (2018). The nature and consequences of essentialist beliefs about race in early childhood. *Child Development*, 90, e437-e453.
- 36 Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of personality and social psychology*, 56(1), 5. Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of personality and social psychology*, 69(6), 1013. Karpinski, A., & Hilton, J. L. (2001). Attitudes and the implicit association test. *Journal of personality and social psychology*, 81(5), 774. Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. *Trends in cognitive sciences*, 15(4), 152-159.
- 37 Payne, B., Vuletic, H., & Brown-Iannuzzi, J. (2019). Historical roots of implicit bias in slavery. *Proceedings of the National Academy of Sciences, USA*, 116, 11693-11698. Payne, B., Vuletic, H., & Lundberg, K. (2017). The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry*, 28, 233-248. Vuletic, H., & Payne, B. (2019). Stability and change in implicit bias. *Psychological Science*, 30, 854-862.
- 38 Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the Implicit Association Test can have societally large effects. *Journal of Personality and Social Psychology*, 108(4), 553-561.
- 39 Rachlinski, J., Johnson, S., Wistrich, A., & Guthrie, C. (2009). Does Unconscious Racial Bias Affect Trial Judges? *Notre Dame Law Review*, 84, 1195-1246.
- 40 Girvan, E., Deason, G., & Borgida, E. (2015). The generalizability of gender bias: Testing the effects of contextual, explicit, and implicit sexism on labor arbitration decisions. *Law and Human Behavior*, 39(5), 525-537.
- 41 Correll, J., Park, B., Judd, C., 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, 1006-1023. Plant, E. A., & Peruche, B. M. (2005). The consequences of race for police officers' responses to criminal suspects. *Psychological Science*, 16, 180-183.
- 42 Green, A. R., Carney, D. R., Pallin, D. J., Ngo, L. H., Raymond, K. L., Iezzoni, L., & Banaji, M. (2007). Implicit bias among physicians and its prediction of thrombolysis decisions for black and white patients. *Journal of General Internal Medicine*, 22, 1231-1238. Onyeador, I., Wittlin, N., Burke, S., Dovidio, J., Perry, S., Hardeman, R., Dyrbye, L., Herrin, J., Phelan, S., & van Ryn, M., (2020). The value of interracial contact for reducing anti-Black bias among non-Black physicians: A cognitive habits and growth evaluation (CHANGE) study report. *Psychological Science*, 31(1), 18-30. Ashford, R.D., Brown, A.M., & Curtis, B. (2019). The language of substance use and recovery: novel use of the Go/No-Go Association Task to measure implicit bias. *Health Communication*, 34(11), 1296-1302. Liang, J., Wolsiefer, K., Zestcott, C., Chase, D., & Stone, J. (2019). Implicit bias toward cervical cancer: Provider and training differences. *Gynecologic Oncology*, 153(1), 80-86. FitzGerald, C. & Hurst, S. (2017). Implicit bias in healthcare professionals: A systematic review. *BMC Medical Ethics*, 18(19). Schwartz, M., Chambliss, H., Brownell, K., Blair, S., & Billington, C. (2003). Weight bias among health professionals specializing in obesity. *Obesity Research*, 11(9), 1033-1039. Tomiyama, A., Finch, L., Belsky, A., Buss, J., Finley, C., Schwartz, M. & Daubenmier, J. (2015). Weight bias in 2001 versus 2013: Contradictory attitudes among obesity researchers and health professionals. *Obesity*, 23(1), 46-53.
- 43 von Hippel, W., Brener, L., & von Hippel, C. (2008). Implicit prejudice toward injecting drug users predicts intentions to change jobs among drug and alcohol nurses. *Psychological Science*, 19, 7-11.
- 44 Rooth, D. (2007). *Implicit discrimination in hiring: Real world evidence* (IZA Discussion Paper No. 2764). Bonn, Germany: Forschungsinstitut zur Zukunft der Arbeit (Institute for the Study of Labor). Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues*, 57, 743-762. McDonnell, M.C., Cmar, J.L., Antonelli, K. & Markoski, K.M. (2019). Professionals' implicit attitudes about the competence of people who are blind. *Journal of Visual Impairment and Blindness*, 113(4), 341-354.
- 45 Rudman, L. A., & Ashmore, R. D. (2007). Discrimination and the Implicit Association Test. *Group Processes and Intergroup Relations*, 10, 359-372.
- 46 Arcuri, L., Castelli, L., Galdi, S., Zogmaister, C., & Amadori, A. (2008). Predicting the vote: Implicit attitudes as predictors of the future behavior of decided and undecided voters. *Political Psychology*, 29, 369-387.
- 47 Nock, M. K., & Banaji, M. R. (2007). Prediction of suicide ideation and attempts among adolescents using a brief performance-based test. *Journal of Clinical and Consulting Psychology*, 75, 707-715.
- 48 Gray, N. S., Brown, A. S., MacCulloch, M. J., Smith, J., & Snowden, R. J. (2005). An implicit test of the associations between children and sex in pedophiles. *Journal of Abnormal Psychology*, 114, 304-308.

- 49 Palfai, T. P., & Ostafin, B. D. (2003). Alcohol-related motivational tendencies in hazardous drinkers: Assessing implicit response tendencies using the modified-IAT. *Behaviour Research and Therapy*, 41, 1149–1162. Wiers, R. W., Houben, K., & de Kraker, J. (2007). Implicit cocaine associations in active cocaine users and controls. *Addictive Behaviors*, 32, 1284–1289. Wiers, R. W., Van Woerden, N., Smulders, F. T. Y., & De Jong, P. J. (2002). Implicit and explicit alcohol related cognitions in heavy and light drinkers. *Journal of Abnormal Psychology*, 111, 648–658. Thush, C., & Wiers, R. W. (2007). Explicit and implicit alcohol-related cognitions and the prediction of future drinking in adolescents. *Addictive Behaviors*, 32, 1367–1383. Thush, C., Wiers, R. W., Ames, S. L., Grenard, J. L., Sussman, S., & Stacy, A. W. (2007). Apples and oranges? Comparing indirect measures of alcohol-related cognition predicting alcohol use in at-risk adolescents. *Psychology of Addictive Behaviors*, 21, 587–591.
- 50 Miller, A. (2019). Expertise fails to attenuate gendered biases in judicial decision-making. *Social Psychological and Personality Science*, 10(2), 227–234.
- 51 Guthrie, C., Rachlinski, J., & Wistrich, A. (2001). Inside the Judicial Mind. *Cornell Law Review*, 86, 777–830. Guthrie, C., Rachlinski, J., & Wistrich, A. (2007). Blinking on the bench: how judges decide cases. *Cornell Law Review*, 93, 1–44. Guthrie, C., Rachlinski, J., & Wistrich, A. (2009). The “Hidden Judiciary”: An Empirical Examination of Executive Branch Justice. *Duke Law Journal*, 58, 1477–1530. Lassiter, G., Diamond, S., Schmidt, H., & Elek, J. (2007). Evaluating videotaped confessions: Expertise provides no defense against the camera-perspective effect. *Psychological Science*, 18(3), 224–226. Rachlinski, J., Guthrie, C., & Wistrich, A. (2007). Heuristics and Biases in Bankruptcy Judges. *Journal of Institutional and Theoretical Economics*, 163, 167–186. Rachlinski, J., Guthrie, C., & Wistrich, A. (2011). Probable Cause, Probability, and Hindsight. *Journal of Empirical Legal Studies*, 8, 72–98. Rachlinski, J., Wistrich, A., & Guthrie, C. (2015). Can Judges Make Reliable Numeric Judgments? Distorted Damages and Skewed Sentences. *Indiana Law Journal*, 90, 695–739. Robbenmolt, J. (2002). Punitive damage decision making: The decisions of citizens and trial court judges. *Law and Human Behavior*, 26, 315–341. Wistrich, A., Guthrie, C., & Rachlinski, J. (2005). Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding. *University of Pennsylvania Law Review*, 153, 1251–1345.
- 52 Ayres, I. (2001). *Pervasive prejudice? Unconventional evidence of race and gender discrimination*. Chicago: University of Chicago Press. Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? *American Economic Review*, 94, 991–1013. Pager, D. (2003). The mark of a criminal record. *American Journal of Sociology*, 108, 937–975. Riach, P. A., & Rich, J. (2002). Field experiments of discrimination in the market place. *Economic Journal*, 112, F480–F518.
- 53 Neuberg, S., & Fiske, S. (1987). Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology*, 53, 431–444. Tetlock, P. (1983). Accountability and complexity of thought. *Journal of Personality and Social Psychology*, 45, 74–83.
- 54 Bodenhausen, G., & Wyer, R. (1985). Effects of stereotypes in decision making and information processing strategies. *Journal of Personality and Social Psychology*, 46, 267–282. Darley, J., & Gross, P. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, 44, 20–33.
- 55 e.g., Word, C., Zanna, M., & Cooper, J. (1973). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology*, 10, 102–120.
- 56 e.g., Dovidio, J., & Gaertner, S. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological Science*, 11, 319–323. Johnson, J. Whitestone, E., Jackson, L., & Gatto, L. (1995). Justice is still not colorblind: Differential racial effects of exposure to inadmissible evidence. *Personality and Social Psychology Bulletin*, 21, 893–898.
- 57 Eells, T., & Showalter, C. (1994). Work-related stress in American trial judges. *Bulletin of the American Academy of Psychiatry & the Law*, 22, 71–83. Hartley, L., & Adams, R. (1974). Effect of noise on the Stroop test. *Journal of Experimental Psychology*, 102, 62–66. Keinan, G. (1987). Decision making under stress: Scanning of alternatives under controllable and uncontrollable threats. *Journal of Personality and Social Psychology*, 52, 639–644.
- 58 E.g., see Bodenhausen, G., & Lichtenstein, M., (1987). Social stereotypes and information-processing strategies: The impact of task complexity. *Journal of Personality and Social Psychology*, 52, 871–880. Gilbert, D., & Hixon, J. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, 60, 509–517. Sherman, J., Lee, A., Bessenoff, G., & Frost, L. (1998). Stereotype efficiency reconsidered: Encoding flexibility under cognitive load. *Journal of Personality and Social Psychology*, 75, 589–606. van Knippenberg, A., Dijksterhuis, A., & Vermeulen, D. (1999). Judgment and memory of a criminal act: The effects of stereotypes and cognitive load. *European Journal of Social Psychology*, 29, 191–201.
- 59 Jones, J. (1972). *Prejudice and Racism*. New York: McGraw-Hill.
- 60 Feagin, J. R. 2001. *Racist America: Roots, Current Realities, and Future Reparations*. New York: Routledge.
- 61 Jones, J. M. (1997). *Prejudice and racism* (2nd ed.). New York, NY: McGraw-Hill.
- 62 Benokraitis, N. V., & Feagin, J. R. (1995). *Modern sexism: Blatant, subtle, and covert discrimination*. Pearson College Division.
- 63 Charlton, J. I. (1998). *Nothing About Us Without Us: Disability Oppression and Empowerment*, Berkeley, CA: University of California Press (p. 51).

- 64 Schmidt, S.L. (2005). More than men in white sheets: Seven Concepts critical to the teaching of racism as systemic inequality. *Equity and Excellence in Education*, 38, 110-122.
- 65 Kite, M. & Whitley, B. (2016). *The Psychology of Prejudice and Discrimination* (3rd ed.). New York, NY: Routledge.
- 66 Wijeyesinghe, C. L., Griffin, P. and Love, B. (1997). Racism curriculum design. In *Teaching for diversity and social justice: A sourcebook* (Eds. Adams, M., Bell, L. A., & Griffin), pp. 82-109. New York: Routledge. Feagin, J. R. 2001. *Racist America: Roots, Current Realities, and Future Reparations*. New York: Routledge. Carmichael, S. [now Ture, K.] & Hamilton, C.V. (1967). *Black Power: The Politics of Liberation in America*. New York: Vintage Books.
- 67 Schmidt, S.L. (2005). More than men in white sheets: Seven Concepts critical to the teaching of racism as systemic inequality. *Equity and Excellence in Education*, 38, 110-122.
- 68 Jones, J. M. (1997). *Prejudice and racism* (2nd ed.). New York, NY: McGraw-Hill.
- 69 Benokraitis, N. V., & Feagin, J. R. (1995). *Modern sexism: Blatant, subtle, and covert discrimination*. Pearson College Division.
- 70 Kite, M. & Whitley, B. (2016). *The Psychology of Prejudice and Discrimination* (3rd ed.). New York, NY: Routledge.
- 71 Kite, M. & Whitley, B. (2016). *The Psychology of Prejudice and Discrimination* (3rd ed.). New York, NY: Routledge.
- 72 Kite, M. & Whitley, B. (2016). *The Psychology of Prejudice and Discrimination* (3rd ed.). New York, NY: Routledge.
- 73 Kite, M. & Whitley, B. (2016). *The Psychology of Prejudice and Discrimination* (3rd ed.). New York, NY: Routledge.
- 74 Kite, M. & Whitley, B. (2016). *The Psychology of Prejudice and Discrimination* (3rd ed.). New York, NY: Routledge.
- 75 Kite, M. & Whitley, B. (2016). *The Psychology of Prejudice and Discrimination* (3rd ed.). New York, NY: Routledge. See also McIntosh, P. (1989). Unpacking the invisible knapsack. *Peace and Freedom Magazine*, July/August, 1989, pp. 10-12.
- 76 Johnson, A. (2006). *Privilege, Power, and Difference*. New York, NY: McGraw Hill Education.
- 77 Allport, G. (1954). *The Nature of Prejudice*. Boston: Addison-Wesley.
- 78 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. & Tropp, L. (2005). Allports' intergroup contact hypothesis: Its history and influence. In Dovidio, J.F., Glick, P., & Rudman, L., (Eds.), *Reflecting on the nature of prejudice: Fifty years after Allport*. Malden, MA: Blackwell (pp. 262-277). Pettigrew, T. F., & Tropp, L. R. (2011). Essays in social psychology. When groups meet: The dynamics of intergroup contact. Psychology Press. Tropp, L. & Prenovost, M. (2008). The role of intergroup contact in predicting children's interethnic attitudes: Evidence from meta-analytic and field studies. In Levy, S. & Killen, M. (Eds.), *Intergroup Attitudes and Relations in Childhood Through Adulthood*. New York: Oxford University Press, pp. 236-248.
- 79 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. Research also suggests that the causal relationship flows in both directions: intergroup contact reduces prejudice, and prejudice reduces intergroup contact. Binder, J., Zagefka, H., Brown, R., Funke, F., Kessler, T., Mummendey, A., Maquil, A., Demoulin, S., & Leyens, J.-P. (2009). Does contact reduce prejudice or does prejudice reduce contact? A longitudinal test of the contact hypothesis among majority and minority groups in three European countries. *Journal of Personality and Social Psychology*, 96(4), 843-856. See also Mousa, S. (2020). Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq. *Science*, 369, 866-870. Paluck, E., Green, S., & Green, D. (2019). The contact hypothesis reevaluated. *Behavioral Public Policy*, 3, 129-158.
- 80 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.
- 81 Aberson, C. L., Porter, M. K., & Gaffney, A. M. (2008). Friendships predict Hispanic student's implicit attitudes toward Whites relative to African Americans. *Hispanic Journal of Behavioral Sciences*, 30, 544-556. Dasgupta, N. & Rivera, L. M. (2008). When social context matters: The influence of long-term contact and short-term exposure to admired group members on implicit attitudes and behavioral intentions. *Social Cognition*, 26, 112-123.
- 82 Onyeador, I., Wittlin, N., Burke, S., Dovidio, J., Perry, S., Hardeman, R., Dyrbye, L., Herrin, J., Phelan, S., & van Ryn, M., (2020). The value of interracial contact for reducing anti-Black bias among non-Black physicians: A cognitive habits and growth evaluation (CHANGE) study report. *Psychological Science*, 31(1), 18-30.
- 83 Tam, T., Hewstone, M., Harwood, J., Voci, A., & Kenworthy, J. (2006). Intergroup contact and grandparent-grandchild communication: The effects of self-disclosure on implicit and explicit biases against older people. *Group Processes and Intergroup Relations*, 9, 413-429. Turner, R. N. Hewstone, M., & Voci, A. (2007). Reducing explicit and implicit outgroup prejudice via direct and extended contact: The mediating role of self-disclosure and intergroup anxiety. *Journal of Personality and Social Psychology*, 93, 369-388.

- 84 Aronson, E. (2000). The jigsaw classroom. Retrieved from: <http://www.jigsaw.org/>. Aronson, E. & Patnoe, S. (1997). *The Jigsaw Classroom: Building Cooperation in the Classroom*. London: Longman. Kite, M. & Whitley, B. (2016). *The Psychology of Prejudice and Discrimination* (3rd ed.). New York, NY: Routledge.
- 85 Aronson, E. & Bridgeman, D. (1979). Jigsaw groups and the desegregated classroom: In pursuit of common goals. *Personality and Social Psychology Bulletin*, 5(4), 438-446.
- 86 Aronson, E. & Bridgeman, D. (1979). Jigsaw groups and the desegregated classroom: In pursuit of common goals. *Personality and Social Psychology Bulletin*, 5(4), 438-446. Aronson, E. (2002). Building empathy, compassion, and achievement in the jigsaw classroom. In J. Aronson (Ed.), *Improving academic achievement* (pp. 209-225). New York, NY: Academic Press. Walker, A. & Crogran, M. (1998). Academic performance, prejudice, and the jigsaw classroom: New pieces to the puzzle. *Community and Applied Social Psychology*, 8, 381-393.
- 87 Heilman, M. & Haynes, M. (2008). Subjectivity in the appraisal process: A facilitator of gender bias in work settings. In E. Borgida & S. Fiske (Eds.), *Beyond Common Sense: Psychological Science in the Courtroom* (pp. 127-155). Oxford: Blackwell Publishing.
- 88 Heilman, M. & Haynes, M. (2008). Subjectivity in the appraisal process: A facilitator of gender bias in work settings. In E. Borgida & S. Fiske (Eds.), *Beyond Common Sense: Psychological Science in the Courtroom* (pp. 127-155). Oxford: Blackwell Publishing. Nieva, V. & Gutek, B. (1980). Sex effects on evaluation. *Academy of Management Review*, 5(2), 267-276.
- 89 Foschi, M., Lai, L., & Sigerson, S. (1994). Gender and double standards in the assessment of job applicants. *Social Psychology Quarterly*, 57, 326-339. Heilman, M., Wallen, A., Fuchs, D., & Tamkins, M. (2004). Penalties for success: Reactions to women who succeed at male gender-typed tasks. *Journal of Applied Psychology*, 89(3), 416-427. Phelan, J., Moss-Racusin, C., & Rudman, L. (2008). Competent yet out in the cold: Shifting criteria for hiring reflect backlash toward agentic women. *Psychology of Women Quarterly*, 32, 406-413.
- 90 Bobbitt-Zeher, D. (2011). Gender discrimination at work: Connecting gender stereotypes, institutional policies, and gender composition of workplace. *Gender and Society*, 25(6), 764-786.
- 91 Biernat, M. (2003). Toward a broader view of social stereotyping. *American Psychologist*, 58, 1019-1027. Biernat, M. (2012). Stereotypes and shifting standards: Forming, communicating, and translating person impressions. In P. Devine & A. Plant (Eds.), *Advances in Experimental Social Psychology*, vol. 45 (p. 1-59). San Diego: Elsevier. Biernat, M., Crandall, C. S., Young, L. V., Kobrynowicz, D., & Halpin, S. M. (1998). All that you can be: Stereotyping of self and others in a military context. *Journal of Personality and Social Psychology*, 75, 301-317. Biernat, M., Fuegen, K., & Kobrynowicz, D. (2010). Shifting standards and the inference of incompetence: Effects of formal and informal evaluation tools. *Personality and Social Psychology Bulletin*, 36(7), 855-868. Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, 66, 5-20. Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology*, 60, 485-499. Biernat, M., Tocci, M. J., & Williams, J. C. (2011). The language of performance evaluations: Gender-based shifts in content and consistency of judgment. *Social Psychological and Personality Science*, 3, 186-192. Biernat, M. & Vescio, T. K. (2002). She swings, she hits, she's great, she's benched: Implications of gender-based shifting standards for judgment and behavior. *Personality and Social Psychology Bulletin*, 28, 66-77.
- 92 Biernat, M., Fuegen, K., & Kobrynowicz, D. (2010). Shifting standards and the inference of incompetence: Effects of formal and informal evaluation tools. *Personality and Social Psychology Bulletin*, 36(7), 855-868. Bobbitt-Zeher, D. (2011). Gender discrimination at work: Connecting gender stereotypes, institutional policies, and gender composition of workplace. *Gender and Society*, 25(6), 764-786.
- 93 Heilman, M. & Haynes, M. (2008). Subjectivity in the appraisal process: A facilitator of gender bias in work settings. In E. Borgida & S. Fiske (Eds.), *Beyond Common Sense: Psychological Science in the Courtroom* (pp. 127-155). Oxford: Blackwell Publishing. Heilman, M. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, 32, 113-135.
- 94 E.g., Baltes, B. B., Bauer, C. B., & Frensch, P. A. (2007). Does a structured free recall intervention reduce the effect of stereotypes on performance ratings and by what cognitive mechanism? *Journal of Applied Psychology*, 92(1), 151-164. Bauer, C.C. & Baltes, B.B. (2002). Reducing the effects of gender stereotypes on performance evaluations. *Sex Roles*, 47, 465-476.
- 95 Olson, M. A., & Fazio, R. H. (2006). Reducing automatically-activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin*, 32, 421-433.
- 96 See 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, 957-971. Wennekers, A. M., Holland, R. W., Wigboldus, D. H. J., & van Knippenberg, A. (2011). First see, the nod: The role of temporal contiguity in embodied evaluative conditioning of social attitudes. *Social Psychological and Personality Science*, 3, 455-461. Kawakami, K., Dovidio, J. F., Moll, J., Hermesen, 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, 871-888. Gawronski, B., Deutsch, R., Mbirkou, S., Seibt, B., & Strack, F. (2008). When "just say no" is not enough: Affirmation versus negation training and the reduction of automatic stereotype activation. *Journal of Experimental Social Psychology*, 44, 370-377.

- 97 E.g., Columb, C. & Plant, A. (2016). The Obama Effect Six Years Later: The Effect of Exposure to Obama on Implicit Anti-Black Evaluative Bias and Implicit Racial Stereotyping. *Social Cognition*, 34(6), 523-543. Murrar, S. & Brauer, M. (2018). Entertainment-education effectively reduces prejudice. *Group Processes and Intergroup Relations*, 21(7), 1053-1077. But see Lybarger, J. & Monteith, M. (2011). The effect of Obama saliency on individual-level racial bias: Silver bullet or smokescreen? *Journal of Experimental Social Psychology*, 47, 647-652. Schmidt, K. & Axt, J. (2016). Implicit and Explicit Attitudes Toward African Americans and Barack Obama Did Not Substantively Change During Obama's Presidency. *Social Cognition*, 34(6), 559-588. Schmidt, K. & Nosek, B. (2010). Implicit (and explicit) racial attitudes barely changed during Barack Obama's presidential campaign and early presidency. *Journal of Experimental Social Psychology*, 46, 308-314.
- 98 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. 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. Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J.-E. L., Joy-Gaba, J. A., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E. E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., . . . Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143(4), 1765-1785. Edwards, D.J., McEnteggart, C., Barnes-Holmes, Y., Lowe, R., Evans, N., Vilardaga, R. (2017). The impact of mindfulness and perspective-taking on implicit associations toward the elderly: A relational frame theory account. *Mindfulness*, 8, 1615-1622.
- 99 Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J.-E. L., Joy-Gaba, J. A., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E. E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., . . . Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143(4), 1765-1785.
- 100 FitzGerald, C., Martin, A., Berner, D., & Hurst, S. (2019). Interventions designed to reduce implicit prejudices and implicit stereotypes in real world contexts: a systematic review. *BMC Psychology*, 7, 29. Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*, 117(3), 522-559. Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J.-E. L., Joy-Gaba, J. A., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E. E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., . . . Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143(4), 1765-1785. Calanchini, J., Lai, C. K., & Klauer, K. C. (2020). Reducing implicit racial preferences: III. A process-level examination of changes in implicit preferences. *Journal of Personality and Social Psychology*. Advance online publication. <https://doi.org/10.1037/pspi0000339>
- 101 Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., Marshburn, C. K., Simon, S., Blanchar, J. C., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhaas, F. M. H., Burns, M., . . . Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, 145(8), 1001-1016. Greenwald, A. & Lai, C. (2020). Implicit social cognition. *Annual Review of Psychology*, 71, 419-445. Lai, C., Hoffman, K., & Nosek, B. (2013). Reducing implicit prejudice. *Social and Personality Psychology Compass*, 7(5), 315-330.
- 102 McNulty, J.K., Olson, M.A., Jones, R.E., & Acosta, L.M. (2017). Automatic associations between one's partner and one's affect as the proximal mechanism of change in relationship satisfaction: Evidence from evaluative conditioning. *Psychological Science*, 28(8), 1031-40 (participants took part in 13 training sessions over 6 weeks). Shook, N.J. & Fazio, R.H. (2008). Interracial roommate relationships: an experimental field test of the contact hypothesis. *Psychological Science*, 19(7), 717-23 (participants lived with a roommate of another race for a full semester). Dasgupta, N. & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology*, 40, 642-658 (participants had ongoing contact with female instructors over the course of an academic year).
- 103 Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*, 117(3), 522-559. Vuletich, H. & Payne, K. (2019). Stability and change in implicit bias. *Psychological Science*, 30, 854-862.
- 104 Vuletich, H. & Payne, K. (2019). Stability and change in implicit bias. *Psychological Science*, 30, 854-862.
- 105 Lai, C., Hoffman, K., & Nosek, B. (2013). Reducing implicit prejudice. *Social and Personality Psychology Compass*, 7(5), 315-330. Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J.-E. L., Joy-Gaba, J. A., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E. E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., . . . Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143(4), 1765-1785.
- 106 E.g., see Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review*, 90, 715-741. Psaty, B., & Prentice, R. (2010). Minimizing bias in randomized trials: The importance of blinding. *Journal of the American Medical Association*, 304, 793-794. Snodgrass, R. (2006). Single- versus double-blind reviewing: An analysis of the literature. *ACM SIGMOD Record*, 35(3), 8-21.

- 107 National Research Council. (2014). *Identifying the culprit: Assessing eyewitness identification*. Washington, DC: National Academies Press. See also Dysart, J. E., Lawson, V. Z., & Rainey, A. (2012). Blind lineup administration as a prophylactic against the post-identification feedback effect. *Law and Human Behavior*, 36, 312–319. Garrett, B. L. (2014). Eyewitness identifications and police practices: A Virginia case study. *Virginia Journal of Criminal Law*, 2, 2013–2026. Wells, G. L. (1988). *Eyewitness identification: A system handbook*. Toronto, Ontario, Canada: Carswell Legal.
- 108 Queally, J. (2019, June 12). *San Francisco D.A. unveils program aimed at removing implicit bias from prosecutions*. Available at <https://www.latimes.com/local/lanow/la-me-san-francisco-da-prosecutions-implicit-bias-software-20190612-story.html>. Stanford Computational Policy Lab (2021). Blind charging: mitigating bias in charging decisions with automated race redaction. Available at <https://policylab.stanford.edu/projects/blind-charging.html>. Chohlas-Wood, A., Nudell, J., Yao, K., Lin, Z., Nyarko, J., & Goel, S. (2021). Blind justice: Algorithmically masking race in charging decisions. Association for the Advancement of Artificial Intelligence. See also Sah, S., Robertson, C., & Baughman, S. (2015). Blinding prosecutors to defendants' race: A policy proposal to reduce unconscious bias in the criminal justice system. *Behavioral Science & Policy*, 1, 69–76.
- 109 Allen, T. J., Sherman, J.W., & Klauer, K. C. (2010). Social context and the self-regulation of implicit bias. *Group Processes and Intergroup Relations*, 13, 137–150. Legault, L., Green-Demers, I., & Eadie, A. L. (2009). When internalization leads to automatization: The role of self-determination in automatic stereotype suppression and implicit prejudice regulation. *Motivation and Emotion*, 33, 10–24. Maddux, W. W., Barden, J., Brewer, M. B., & Petty, R. E. (2005). Saying no to negativity: The effects of context and motivation to control prejudice on automatic evaluative responses. *Journal of Experimental Social Psychology*, 41, 19–35. 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, 167–184. Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of antiprejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science*, 22, 1472–1477.
- 110 Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of antiprejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science*, 22, 1472–1477.
- 111 Castelli, L., & Tomelleri, S. (2008). Contextual effects on prejudiced attitudes: When the presence of others leads to more egalitarian responses. *Journal of Experimental Social Psychology*, 44, 679–686. Richeson, J. A., & Ambady, N. (2001). Who's in charge? Effects of situational roles on automatic gender bias. *Sex Roles*, 44, 493–512. Richeson, J. A. & Ambady, N. (2003). Effects of situational power on automatic racial prejudice. *Journal of Experimental Social Psychology*, 39, 177–183. Sechrist, G. B., & Stangor, C. (2001). Perceived consensus influences intergroup behavior and stereotype accessibility. *Journal of Personality and Social Psychology*, 80, 645–654. Sinclair, S., Lowery, B. S., Hardin, C. D., & Colangelo, A. (2005). Social tuning of automatic racial attitudes: The role of affiliative motivation. *Journal of Personality and Social Psychology*, 89, 583–592. Williams, J. (2018). Accountability as debiasing strategy: Testing the effect of racial diversity in employment committees. *Iowa Law Review*, 103(4), 1593–1638.
- 112 Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influences on automatic racial prejudice. *Journal of Personality and Social Psychology*, 81, 842–855.
- 113 Sommers, S. (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.
- 114 Devine, P. G., & Monteith, M. J. (1999). Automaticity and control in stereotyping. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (p. 339–360). The Guilford Press. Monteith, M. J., Parker, L. R., & Burns, M. D. (2016). The self-regulation of prejudice. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (p. 409–432). Psychology Press.
- 115 Carnes, M., Devine, P.G., Baier Manwell, L., Byars-Winston, A., Fine, E., Ford, C.E., Forscher, P., Isaac, C., Kaatz, A., Magua, W., Palta, M., & Sheridan, J. (2015). The effect of an intervention to break the gender bias habit for faculty at one institution: A cluster randomized, controlled trial. *Academic Medicine*, 90, 221–230. Devine, P.G., Forscher, P.S., Cox, W.T.L., Kaatz, A., Sheridan, J., & Carnes, M. (2017). A gender bias habit-breaking intervention led to increased hiring of female faculty in STEM departments. *Journal of Experimental Social Psychology*, 73, 211–215. Forscher, P., Mitamura, C., Dix, E., Cox, W., & Devine, P. (2017). Breaking the prejudice habit: Mechanisms, timecourse, and longevity. *Journal of Experimental Psychology*, 72, 133–146.
- 116 Monteith, M. J., Parker, L. R., & Burns, M. D. (2016). The self-regulation of prejudice. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (p. 409–432). Psychology Press.
- 117 Mendoza, S. A., Gollwitzer, P. M., & Amodio, D. M. (2010). Reducing the expression of implicit stereotypes: Reflexive control through implementation intentions. *Personality and Social Psychology Bulletin*, 36, 512–523. Stewart, B. D., & Payne, B. K. (2008). Bringing automatic stereotyping under control: Implementation intentions as efficient means of thought control. *Personality and Social Psychology Bulletin*, 34, 1332–1345. Webb, T. L., Sheeran, P., & Pepper, J. (2012). Gaining control over responses to implicit attitude tests: Implementation intentions engender fast responses on attitude-incongruent trials. *British Journal of Social Psychology*, 51, 13–32. Rees, H., Rivers, A., Sherman, J. (2019). Implementation intentions reduce implicit stereotype activation and application. *Personality and Social Psychology Bulletin*, 45(1), 37–53.

- 118 Monteith, M. J., Parker, L. R., & Burns, M. D. (2016). The self-regulation of prejudice. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (p. 409–432). Psychology Press.
- 119 Aarts, H., Dijksterhuis, A., & Midden, C. (1999). To plan or not to plan? Goal achievement or interrupting the performance of mundane behaviors. *European Journal of Social Psychology*, 29, 971–979.
- 120 Bayer, U. C., Achtziger, A., Gollwitzer, P. M., & Moskowitz, G. (2009). Responding to subliminal cues: Do if-then plans facilitate action preparation and initiation without conscious intent? *Social Cognition*, 27, 183–201. Gollwitzer, P. M., & Brandstätter, V. (1997). Implementation intentions and effective goal pursuit. *Journal of Personality and Social Psychology*, 73, 186–199. Martijn, C., Alberts, H., Sheeran, P., Peters, G. Y., Mikolajczak, J., & de Vries, N. K. (2008). Blocked goals, persistent action: Implementation intentions engender tenacious goal striving. *Journal of Experimental Social Psychology*, 44, 1137–1143.
- 121 Achtziger, A., Gollwitzer, P. M., & Sheeran, P. (2008). Implementation intentions and shielding goal striving from unwanted thoughts and feelings. *Personality and Social Psychology Bulletin*, 34, 381–393.
- 122 Calanchini, J., Lai, C. K., & Klauer, K. C. (2020). Reducing implicit racial preferences: III. A process-level examination of changes in implicit preferences. *Journal of Personality and Social Psychology*. Advance online publication. <https://doi.org/10.1037/pspi0000339>. Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J.-E. L., Joy-Gaba, J. A., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E. E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., . . . Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143(4), 1765–1785.
- 123 Trevino, L., den Nieuwenboer, N., & Kish-Gephart, J. (2014). (Un)ethical behavior in organizations. *Annual Review of Psychology*, 65, 635–660.
- 124 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.
- 125 National Center for State Courts, *State court statements on racial justice*. Available at: <https://www.ncsc.org/newsroom/state-court-statements-on-racial-justice>.
- 126 For information about recent activities, see <https://www.ncsc.org/information-and-resources/racial-justice/principles-and-policy>
- 127 See the National Consortium on Racial and Ethnic Fairness in the Courts' website at <https://www.national-consortium.org/state-efforts> and the National Center for State Courts website at <https://www.ncsc.org/information-and-resources/racial-justice>.
- 128 See <https://www.ncsc.org/information-and-resources/racial-justice>
- 129 Oregon Supreme Court (2020, June 5). Letter re: George Floyd. Available at <https://www.osbar.org/docs/resources/OSCMessages/reGeorgeFloyd.pdf>
- 130 Roberson, Q. (2006). Disentangling the Meanings of Diversity and Inclusion in Organizations. *Group and Organization Management*, 31, 212–236.
- 131 Roberson, Q. (2006). Disentangling the Meanings of Diversity and Inclusion in Organizations. *Group and Organization Management*, 31, 212–236.
- 132 Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: a new model for analysis of the educational process. *Educational Research*, 49, 343–36. Morgan, W. & Sawyer, J. (1967). Bargaining, expectations, and the preference for equality over equity. *Journal of Personality and Social Psychology*, 6, 139–149.
- 133 Shriver Center on Poverty Law, <https://www.povertylaw.org/>
- 134 Cook, K. & Hegtvedt, K. (1983). Distributive justice, equity, and equality. *Annual Review of Sociology*, 9, 217–241. Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: a new model for analysis of the educational process. *Educational Research*, 49, 343–36. Morgan, W. & Sawyer, J. (1967). Bargaining, expectations, and the preference for equality over equity. *Journal of Personality and Social Psychology*, 6, 139–149.
- 135 E.g., National Center for State Courts, *State court statements on racial justice*, available at: <https://www.ncsc.org/newsroom/state-court-statements-on-racial-justice>. JustLead Washington, *Washington Race Equity & Justice Initiative Organizational Race Equity Toolkit*, <https://justleadwa.org/learn/rejitoolkit/>.

- 136 Onyeador, I., Hudson, S., & Lewis, N. (2021). Moving beyond implicit bias training: policy insights for increasing organizational diversity. *Policy Insights from the Behavioral and Brain Sciences*, 8, 19-26.
- 137 Dobbin, F., & Kalev, A. (2016). Why diversity programs fail – and what works better. *Harvard Business Review*, 94, 7. Available at <https://hbr.org/2016/07/why-diversity-programs-fail>. Greenwald, A. & Lai, C. (2020). Implicit social cognition. *Annual Review of Psychology*, 71, 419-445. Onyeador, I., Wittlin, N., Burke, S., Dovidio, J., Perry, S., Hardeman, R., Dyrbye, L., Herrin, J., Phelan, S., & van Ryn, M., (2020). The value of interracial contact for reducing anti-Black bias among non-Black physicians: A cognitive habits and growth evaluation (CHANGE) study report. *Psychological Science*, 31(1), 18-30.
- 138 Dobbin, F., & Kalev, A. (2016). Why diversity programs fail – and what works better. *Harvard Business Review*, 94, 7. Available at <https://hbr.org/2016/07/why-diversity-programs-fail>. Carter, E., Onyeador, I., & Lewis, N. (2020). Developing & delivering effective anti-bias training: Challenges & recommendations. *Behavioral Science & Policy*, 6, 57-70.
- 139 Carter, E., Onyeador, I., & Lewis, N. (2020). Developing & delivering effective anti-bias training: Challenges & recommendations. *Behavioral Science & Policy*, 6, 57-70.
- 140 E.g., Carnes, M., Devine, P. G., Manwell, L. B., Byars-Winston, A., Fine, E., Ford, C. E., ... & Palta, M. (2015). Effect of an intervention to break the gender bias habit for faculty at one institution: a cluster randomized, controlled trial. *Academic Medicine: Journal of the Association of American Medical Colleges*, 90(2), 221. Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, 48(6), 1267-1278. Devine, P., Forscher, P., Cox, W., Kaatz, A., Sheridan, J., & Carnes, M. (2017). A gender bias habit-breaking intervention led to increased hiring of female faculty in STEM departments. *Journal of Experimental Social Psychology*, 73, 211-215. Forscher, P. S., Mitamura, C., Dix, E. L., Cox, W. T., & Devine, P. G. (2017). Breaking the prejudice habit: Mechanisms, timecourse, and longevity. *Journal of experimental social psychology*, 72, 133-146. Moss-Racusin, Pietri, E., Hennes, E., Dovidio, J., Brescoll, V., Roussos, G., & Handelsman, J. (2018). Reducing STEM gender bias with VIDS (video interventions for diversity in STEM). *Journal of Experimental Psychology: Applied*, 24, 236-260.
- 141 E.g., see Livingston, R. (2020). How to promote racial equity in the workplace: a five-step plan. *Harvard Business Review*. Available at <https://hbr.org/2020/09/how-to-promote-racial-equity-in-the-workplace>.
- 142 Biden, J. (2021, January 20). Executive order on advancing racial equity and support for underserved communities through the federal government. Available at <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/>.
- 143 National Center for State Courts (2021, February 8). Court Statistics Project data governance special topic: Collecting race & ethnicity data. Available at https://www.ncsc.org/data/assets/pdf_file/0016/60343/Race_special_topic_v2.pdf.
- 144 E.g., Bishop, E., Hopkins, B., Obiofuma, C., & Owusu, F. (2020). Racial disparities in the Massachusetts criminal system. Available at <https://hls.harvard.edu/content/uploads/2020/11/Massachusetts-Racial-Disparity-Report-FINAL.pdf?text=Racial%20Disparities%20in%20the%20Massachusetts%20Criminal%20System%20,Hopkins%2C%20Chijindu%20Obiofuma%2C%20Felix%20Owusu%20,%20September%202020>.
- 145 Equal Justice in the Courts Task Force (2020, October 1). Report from the Special Adviser on Equal Justice in the New York State Courts. Available at <http://www.nycourts.gov/whatsnew/pdf/SpecialAdviserEqualJusticeReport.pdf>.
- 146 Gaddis, S. M. (2021, March 23). Invited panelist on the prevalence and pervasiveness of implicit bias. *National Academy of Sciences Virtual Workshop on the Science of Implicit Bias: Implications for Law and Policy*. See <https://www.nationalacademies.org/event/01-12-2021/the-science-of-implicit-bias-implications-for-law-and-policy-a-workshop>. Gaddis, S.M. (2018). *Audit studies: Behind the scenes with theory, method, and nuance*. Cham, Switzerland: Springer International.
- 147 E.g., see the National Center for State Courts' Community Engagement and the State Courts Initiative at <https://www.ncsc.org/information-and-resources/racial-justice/community-engagement-initiative>. For a local example, see <https://www.safetyandjusticechallenge.org/challenge-site/multnomah-county/>, <https://multco.us/lpscc/macarthursafetyandjusticechallenge>, and <https://multco.us/lpscc/documents-0> for information about Multnomah County, OR's work as a John D. and Catherine T. MacArthur Foundation Safety and Justice Challenge site, which includes efforts to examine racial and ethnic disparities across seven key criminal justice decision points and a variety of community engagement activities.
- 148 pp. 129-131 in Stephens-Davidowitz, S. (2017) *Everybody lies: Big data, new data, and what the internet can tell us about who we really are*. New York, NY: Harper Collins.
- 149 E.g., see van Knippenberg, D., Nishii, L., & Dwertmann, D. (2020). Synergy from diversity: Managing team diversity to enhance performance. *Behavioral Science & Policy*, 6, 75-92.

¹⁵⁰ For more information about the types of engagement and the characteristics of effective community engagement activities, see Advancing Pretrial Policy and Research (2020). *Strengthening and sustaining public engagement*. Available at <https://cdn.filestackcontent>.

¹⁵¹ Iris Bohnet suggests two other techniques for improved decision-making. In the *consider-the-opposite* approach, decision-makers play devil's advocate with themselves to identify arguments for why their thinking, including their conclusions, may be wrong before settling on a decision. In the *crowd-within* approach, decision-makers formulate multiple answers based on reconsidered evidence and takes the average of the answers to produce a more accurate result. First, the decision-maker writes down an initial, intuitive answer. Then, the decision-maker reviews the evidence used to formulate the initial answer, possibly remembering different pieces of evidence, to produce a more deliberative second answer. Third, the decision-maker repeats this process, but this time looks to identify and consider other information that was disregarded the first two times – such as by changing the focus of attention or asking new questions – to produce a third answer. Finally, the decision-maker averages the three guesses to produce a final judgment. This technique is recommended when estimating likelihoods. Bohnet, I. (2016). *What works: gender equality by design*. Cambridge, MA: Harvard University Press.

¹⁵² pp. 162-163 in Stephens-Davidowitz, S. (2017) *Everybody lies: Big data, new data, and what the internet can tell us about who we really are*. New York, NY: Harper Collins.

¹⁵³ pp. 248-249 in Barrett, L. F. (2017). *How emotions are made: the secret life of the brain*. Boston: Houghton Mifflin Harcourt.

¹⁵⁴ Robinson, R. (2020, June 9). Current events letter. Available at <https://jud.ct.gov/Homepdfs/Currenteventsletter.pdf>.

Appendix A: Glossary of Terms

Automatic: one end of a spectrum that captures the nature of a psychological process. Fully automatic processes do not require intention, effort, or conscious awareness in order to be enacted.

Bias: the unintended influence of factors that are not meant to be considered on a final decision or result. Bias can occur either when relevant information does not influence the decision or when irrelevant information influences the decision. The particular situation or legal context surrounding a decision determines which factors are considered relevant or irrelevant.

Conscious: mental processes involving both awareness and volition.

Controlled: one end of a spectrum that captures the nature of a psychological process. Controlled processes require at least some intention, effort, or conscious awareness in order to be enacted.

Cultural inequality: the inequality that is built into our literature, art, music, language, morals, customs, beliefs, and ideology to such an extent that it defines a generally agreed-upon way of life.

Discrimination: differential treatment of, or outcomes for, different people, based on their membership in a particular social group.

Diversity: the presence of individuals who represent a variety of groups or perspectives.

Equality: equal treatment or distribution of resources, regardless of people's needs or starting positions.

Equity: the state that exists when we cannot predict outcomes based on a person's group membership, and outcomes for all groups are improved. Equity often involves the differential treatment of different individuals, based on their needs and starting positions, with the goal that everyone will arrive at the same outcome.

Explicit bias: bias that is measured using an explicit, or direct, measure.

Explicit measures: measures of cognition, affect, and behavior that require participants to self-report their responses. These rely on the assumption that individuals are aware of their responses and are willing to express them.

Implicit bias: bias that is measured using an implicit, or indirect, measure. This technical definition is used by many research scientists today, but it differs from how the term is used in common vernacular. In common vernacular, *implicit bias* and *unconscious bias* are often used synonymously to refer to an attitude, stereotype, or prejudice that a person is unaware of possessing but which may operate automatically to influence thinking or behavior.

Implicit measures: measures of cognition, affect, and behavior that capture participants' responses in ways that do not rely on individuals' awareness or willingness to respond, such as by measuring reaction time to different groups of stimuli.

Implicit social cognition: the scientific field of study that uses implicit or indirect measures in research on attitudes, stereotypes, and self-esteem.

Inclusion: the meaningful involvement of people from different groups, or the extent to which diverse perspectives are incorporated into systems, processes, and decisions.

Institutional inequality: the network of institutional structures, policies and practices that create advantages and benefits for some groups over others. Institutions can include the justice system, schools, media, banks, business, health care, governmental bodies, family units, religious organizations, and civic groups.

Interpersonal inequality: a situation in which inequality manifests at the individual, person-to-person level.

Microaggressions: brief, everyday exchanges that send denigrating messages to certain individual because of their membership in certain groups.

Organizational inequality: the practices, rules and policies of formal organizations (such as corporations or government agencies) that result in different outcomes for different groups.

Prejudice: the emotion, attitude, or evaluation that a person feels about members of a particular social group.

Privilege: an unearned favored state conferred simply because of one's group membership.

Stereotype: beliefs and opinions about the characteristics, attributes, and behaviors of members of a group. When one engages in the act of stereotyping, one assumes that because an individual belongs to a particular social group, the individual must share the characteristics of the group.

Systemic inequality: the combination of a diverse array of discriminatory and inequitable practices in society, including the unjustly gained economic and political power of some groups over others, ongoing resource inequalities, ideologies and attitudes that regard some groups as superior to others, and the set of institutions that preserve the advantages of some groups over others.

Unconscious: mental processes that lack either full awareness or full volition.

Appendix B: Implicit Measures

Researchers use several scientific methods to measure implicit bias.ⁱ Some of these measures are used as demonstrations of implicit bias in educational settings. There is no implicit measure that is appropriate for use as a diagnostic assessment in professional settings.

Implicit measures rely on the assumption that automatic associations between two concepts (e.g., race and valence) will influence behavior in a measurable way (e.g., reaction time, sweat). Two of the most common classes of implicit measures are (1) the Implicit Association Test (IAT) and other response competition procedures and (2) sequential priming procedures. Each are described in turn.

1. The Implicit Association Test (IAT) and Other Response Competition Procedures

The Implicit Association Test (IAT) is one of the best known implicit measures.ⁱⁱ The IAT is considered a *response competition procedure*, an implicit measure that emerged from research on interference effects. Specifically, when a stimulus has multiple different interpretations (e.g., the word “red” is written in green font), the different meanings can lead to competing reactions in a given task (e.g., identifying the color of the word) that can interfere with the respondent’s performance on the task.ⁱⁱⁱ Response competition procedures take advantage of interference effects by presenting two competing categorization tasks in a single procedure and measuring the disparities in reaction time.^{iv}

In the classic IAT, respondents are asked to categorize a sequence of images (e.g., a Black or White face) and words (e.g., good or bad) by pressing one of two pre-labeled buttons on a computer. For example, respondents may be

instructed to press the left button whenever they see a Black face or whenever a negative word appears on the computer screen, and to press the right button whenever they see a White face or a positive word. The program systematically varies how the faces, words, and buttons are matched. Because of interference effects, individuals who associate “Black” with “bad,” for example, will respond more slowly when “Black” and “good” share the same response button.

The structure of the IAT has also been modified to accommodate a variety of research needs, including a short version (Brief IAT) and comparing a single category of stimuli (Single Category IAT).^v Other response competition procedures have also gained popularity, including the Go/No-Go Association Task (GNAT) and the Extrinsic Affective Simon Task (EAST).^{vi}

Despite its many positive attributes, the creators of the IAT have emphasized that the IAT should not be used as a diagnostic assessment.^{vii} This means that an IAT score should not be used to make predictions about an individual’s behavior, for example, to inform hiring decisions about that individual or as part of a jury selection process. Part of the reason for this is because the IAT does not meet the scientific standards for predictive validity and test-retest reliability that is required of diagnostic assessments.^{viii}

The IAT is primarily used in research studies, but the measure is often completed in educational settings as an interactive exercise intended to illustrate implicit bias. The most common implementation of the IAT requires the use of a computer and a strong internet connection. A variety of IATs (e.g., on gender, sexuality, race, religion, weight, skin tone, age, disability, and more) are available for free at

<https://implicit.harvard.edu/implicit/selectatest.html>. If participants do not have access to computers, there are several other options for practitioners. Manual adaptations of the IAT can be printed out ahead of time and only require the use of a pencil.^{ix} For presentation purposes, the IAT can also be

demonstrated as a group exercise, with participants raising either their left or right hand in response to words or photographs presented to the class.

A Note on Educational Uses of the IAT

The Implicit Association Test (IAT) is popular among researchers, educators, and the media, with over 17 million IATs taken on the Project Implicit website between its launch in 1998 and 2015.^x Because of the popularity of the IAT and free public access to an array of computerized versions of the test, many educators have used it as an interactive demonstration of implicit bias, and many educational participants may be curious about the instrument.

When used in educational sessions, course organizers may ask participants to take the IAT before facilitating a discussion about their experiences taking the test, including what they found surprising and whether the results changed their perspective about the nature of bias.^{xi} In this way, the IAT can act as an icebreaker to get participants comfortable thinking and talking about implicit bias and provides them with an example of how automatic associations can alter measurable behavior.

Although most people who take the IAT report having a positive learning experience, some of those who demonstrate implicit bias on the test may not respond as favorably.^{xii} Facilitators should frame the taking of an IAT within a broader conversation about individual and systemic biases and discrimination, emphasizing how a person may express implicit bias even if they endorse egalitarian values. Clarifying the science of implicit bias can help overcome some of the barriers caused by mistaken beliefs that one's values and behaviors are always aligned, that the lack of intent absolves a person of responsibility for their actions, and about the role of individual free will vs. the power of the situation.

2. Sequential Priming Procedures

Sequential priming procedures are based on evidence demonstrating that when two concepts are associated in a person's memory, the presentation of one of those concepts facilitates the recall or recognition of the other.^{xiii} For example, when people are presented with one concept (e.g., a picture of an apple), they are faster at identifying the next concept (e.g., a picture of a banana) when they associate the two concepts in memory (e.g., as fruits). Evidence suggests priming procedures work even if

the primes are flashed on a screen so quickly that they are not consciously detected by the respondent.

One popular procedure for measuring this phenomenon is the evaluative priming task (also referred to as the *bona-fide pipeline*).^{xiv} In this task, respondents are briefly presented with a Black or White face immediately before a positive or negative target word appears on the screen. They must then identify, as quickly as possible, the meaning of the presented word as "good" or "bad." In the standard paradigm,

respondents with racial bias more quickly identify negative words as “bad” and more slowly identify positive words as “good” when that word appears immediately after the presentation of a Black face.^{xv}

A similar priming procedure, called the Affect Misattribution Procedure (AMP), briefly presents respondents with the prime of a Black or White face before viewing a neutral Chinese character.^{xvi} The respondents are asked to evaluate the character as more or less visually pleasant than the average Chinese character. Researchers found that individuals’ racial attitudes affected their evaluations of the Chinese characters, with White respondents reporting more favorable ratings for the characters that appeared after White primes compared to Black primes. This effect emerged even when respondents received a forewarning about the influence of the racial primes on subsequent evaluations.

Another priming procedure is the Weapon Identification Task.^{xvii} In this task, participants are primed with a photograph of either a Black or White face and then asked to determine if a photographed object is either a weapon or a tool. In an updated version of this procedure referred to as a “shooter task,” respondents are shown images of Black and White targets either holding a weapon or a non-weapon (e.g., a cellphone, a drill). The respondent must quickly decide whether to “shoot” the target by pressing a computer key. Implicit bias is measured by disparities in reaction time and the number of mistakes made by respondents. In comparison to White targets, responders are faster to shoot an armed Black target and slower to not shoot an unarmed Black target.^{xviii}

Like the IAT, sequential priming tasks have relatively low predictive validity and test-retest reliability and should not be used as diagnostic assessments.^{xix} Many priming procedures require the use of a computer and an internet connection. Some sequential priming procedures like the AMP may be more easily adapted to a paper and pencil format than the IAT because the procedure does not involve measurement

of response time.^{xx} There are a host of simple priming demonstrations that can be done as icebreakers in educational sessions, such as the Word Fragment Completion (WFC) task, in which people are presented with fragments of words (e.g., POLI_E) and are asked to fill in the missing letters. These word fragments, however, can be completed in stereotypic or non-stereotypic ways (e.g., POLITE, POLICE), and the number of stereotypic word completions in the WFC task has been used as a measure of implicit bias.^{xxi} Other priming procedures are more complex, such as experiments that use virtual reality headsets that place the participant in the body of someone of a different race and measure how the participants respond to the virtual world differently in different bodies.^{xxii}

A Note on the Use of Technology

As noted above, standard versions of both the IAT and sequential priming procedures require modest resources (e.g., a computer and an internet connection). However, implementation options range from high-tech (e.g., using physiological measurements or a virtual reality headset) to low-tech (using only a paper and pencil).

High-tech implicit measures include physiological and neuropsychological techniques. Physiological measures have been used by researchers to better understand the underpinnings of implicit bias.^{xxiii} The physiological study of implicit bias has focused



on autonomic nervous system responses such as the amount of sweat produced, heart rate, and even small facial muscle movements that are nearly imperceptible to the untrained human eye.^{xxiv} Some of these physiological measures (e.g., sweat production and heart rate) indicate only when a stimulus (e.g., a photograph of a person) provokes a heightened response; they do not differentiate between positive and negative responses. Other techniques, like facial electromyography, provide a measure of the valence of a reaction because certain bodily responses (e.g., cheek activity associated with smiling or frowning movements) are associated with positive and negative emotions. More recently, neuroscientists have attempted to understand the neural underpinnings of implicit bias by using fMRI machines to measure blood flow in the brain. There is a correlation between the degree of activation in the amygdala region of the brain (which is linked to affective information processing), as measured by fMRI, and scores on the IAT.^{xxvi} These techniques are useful for advancing scientific understanding of the physiology of implicit bias, but are not intended for diagnostic assessment.

More recently, researchers have begun using virtual reality technology to measure implicit bias.^{xxvii} In these tasks, participants are given virtual avatars from different races and asked to interact with one another in a virtual world. Participants tend to demonstrate reduced bias for outgroup members when using a virtual avatar from another race (as measured by how much the participant imitated the behaviors of outgroup and ingroup members in the virtual reality world). White respondents who take the IAT after playing a virtual reality game with a Black avatar show reduced bias.^{xxviii} This suggests there may be some promise in perspective-shifting experiences that allow one to literally (or virtually) walk in another's shoes.

Low-tech implicit measures are often useful as demonstrations of implicit bias in educational settings. While the IAT and sequential priming tasks have paper and pencil versions (as described in the preceding section), there are

other implicit measures that are specifically designed to only require paper and pencil. Several of these measures assess attribution processing styles, which show how a respondent uses associations in memory to infer a cause for an observed behavior. One such example is the Stereotypic Explanatory Bias (SEB), which is the tendency to ascribe the stereotype-consistent behavior of minorities to factors intrinsic to the individual (e.g., trait or dispositional attributions like hard work or talent), but stereotype-inconsistent behavior to extrinsic, situational factors (e.g., the weather, luck).^{xxix} Similarly, the Linguistic Intergroup Bias is the tendency to describe stereotypic behavior using abstract language (e.g., by ascribing the behavior to a global trait) but non-stereotypic behavior using concrete language (e.g., by describing the behavior as a specific event).^{xxx} By carefully examining the respondent's choice of language or agreement with particular summaries of a behavioral event, researchers have used these tendencies as indicators of implicit bias.^{xxxi}

Endnotes

- i. For reviews, see Hahn, A., & Gawronski, B. (2018). Implicit social cognition. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, 4, 1-33. 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. Gawronski, B. (2009). Ten frequently asked questions about implicit measures and their frequently supposed, but not entirely correct answers. *Canadian Psychology/Psychologie canadienne*, 50(3), 141. Wittenbrink, B., & Schwarz, N. (Eds.). (2007). *Implicit measures of attitudes*. Guilford Press.
- ii. Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464.
- iii. The classic Stroop effect is one example of interference effects at work. See Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18(6), 643. MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological bulletin*, 109(2), 163.
- iv. Wittenbrink, B., & Schwarz, N. (Eds.). (2007). *Implicit measures of attitudes*. Guilford Press.
- v. See Sriram, N., & Greenwald, A. G. (2009). The brief implicit association test. *Experimental psychology*, 56(4), 283-294. Karpinski, A., & Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. *Journal of personality and social psychology*, 91(1), 16.
- vi. E.g., Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social cognition*, 19(6), 625-666. E.g., De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental psychology*, 50(2), 77.
- vii. Project Implicit website, available at <https://implicit.harvard.edu>.
- viii. See Forscher, P., Lai, C., Axt, J., Ebersole, C., Herman, M., Devine, P., & Nosek, B. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*, 117, 522-559. Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the Implicit Association Test can have societally large effects. Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, 105, 171-192. 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.), *Automatic processes in social thinking and behavior* (pp. 265-292). New York: Psychological Press.
- ix. Kitayama, S., & Uchida, Y. (2003). Explicit self-criticism and implicit self-regard: Evaluating self and friend in two cultures. *Journal of Experimental Social Psychology*, 39(5), 476-482. Lemm, K., Sattler, D., Khan, S., Mitchell, R., & Dahl, J. (2002). *Reliability and validity of a paper-based implicit association test*. Paper presented at the fourth annual convention of the Society for Personality and Social Psychology, Savannah, GA.
- x. Kassin, S., Fein, S., & Markus, H. (2016) *Social psychology*, 10th ed. Belmont, CA: Cengage.
- xi. E.g., see Estes, M., & Smith, N. (2015, November 12). #IAmFruitvale#: An approach to teaching court staff about racism, prejudice, and implicit bias. Available at <https://nasje.org/iamfruitvale-an-approach-to-teaching-court-staff-about-racism-prejudice-and-implicit-bias/>.
- xii. Hillard, V. H., Devos, T., Rivera, L. M., Smith, H., & Vega, L. A. (2014). Teaching about implicit prejudices and stereotypes: a pedagogical demonstration. *Society for the Teaching of Psychology*, 41(3), 204-212.
- xiii. see Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner, & G. W. Humphreys (Eds.), *Basic processes in reading* (pp. 264-336). Hillsdale, NJ: Lawrence Erlbaum Associates.
- xiv. Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of personality and social psychology*, 50(2), 229.
- xv. Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of personality and social psychology*, 69(6), 1013.
- xvi. Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: affect misattribution as implicit measurement. *Journal of personality and social psychology*, 89(3), 277.
- xvii. 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.
- xviii. Mekawi, Y., & Bresin, K. (2015). Is the evidence from racial bias shooting task studies a smoking gun? Results from a meta-analysis. *Journal of Experimental Social Psychology*, 61, 120-130.

- xix. see Cameron, C. D., Brown-Iannuzzi, J. L., & Payne, B. K. (2012). Sequential priming measures of implicit social cognition: A meta-analysis of associations with behavior and explicit attitudes. *Personality and Social Psychology Review*, 16(4), 330-350. Gawronski, B., & De Houwer, J. (2014). Implicit measures in social and personality psychology. *Handbook of research methods in social and personality psychology*, 2, 283-310.
- xx. Vargas, P. T., Sakaquaptewa, D., & von Hippel, W. (2007). Armed Only with Paper and Pencil: "Low-Tech" Measures of Implicit Attitudes.
- xxi. e.g., Son Hing, L. S., Li, W., & Zanna, M. P. (2002). Inducing hypocrisy to reduce prejudicial responses among aversive racists. *Journal of Experimental Social Psychology*, 38(1), 71-78. 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.
- xxii. Hasler, B. S., Spanlang, B., & Slater, M. (2017). Virtual race transformation reverses racial in-group bias. *PLoS one*, 12(4), e0174965.
- xxiii. For reviews, see Banaji, M. R., & Heiphetz, L. (2010). Attitudes. In: S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (pp. 348-388). New York: John Wiley & Sons. Blascovich, J., & Mendes, W. (2010). Social psychophysiology and embodiment. In S. Fiske, D. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology*, 5th edition. New York: Wiley. Ito, T., & Cacioppo, J. (2007). Attitudes as mental and neural states of readiness. In B. Wittenbrink and N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 125-158). New York: Guilford.
- xxiv. e.g., Rankin, R., & Campbell, D. (1955). Galvanic skin response to Negro and White experimenters. *Journal of Abnormal and Social Psychology*, 51, 30-33. Shields, S., & Harriman, R. (1984). Fear of male homosexuality: Cardiac responses of low and high homonegative males. *Journal of Homosexuality*, 10, 53-67. Vanman, E., Saltz, J., Nathan, L., & Warren, J. (2004). Racial discrimination by low-prejudiced Whites: Facial movements as implicit measures of attitudes related to behavior. *Psychological Science*, 15, 711-714. Mahaffey, A., Bryan, A., & Hutchison, K. (2005). Using startle eye blink to measure the affective component of antigay bias. *Basic and Applied Social Psychology*, 35, 537-545.
- xxv. e.g., Banaji, M. R., & Heiphetz, L. (2010). Attitudes. In: S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (pp. 348-388). New York: John Wiley & Sons. Cunningham, W., Johnson, M., Gatenby, J., Gore, J., & Banaji, M. (2003). Neural components of social evaluation. *Journal of Personality and Social Psychology*, 85, 639-649. Stanley, D., Phelps, E., & Banaji, M. (2008). The neural basis of implicit attitudes. *Current Directions in Psychological Science*, 17, 164-170.
- xxvi. Phelps, E., O'Connor, K., Cunningham, W., Funayama, E., Gatenby, J., Gore, J., & Banaji, M. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, 12, 729-738. Cunningham, W., Johnson, M., Gatenby, J., Gore, J., & Banaji, M. (2003). Neural components of social evaluation. *Journal of Personality and Social Psychology*, 85, 639-649.
- xxvii. e.g., Hasler, B. S., Spanlang, B., & Slater, M. (2017). Virtual race transformation reverses racial in-group bias. *PLoS one*, 12(4), e0174965.
- xxviii. E.g., Salmanowitz, N. (2018). The impact of virtual reality on implicit racial bias and mock legal decisions. *Journal of Law and the Biosciences*, 5(1), 174-203. Peck, T., Seinfeld, S., Aglioti, S., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, 22, 779-787. Schutte, N., & Stilianovic, E. (2017). Facilitating empathy through virtual reality. *Motivation & Emotion*, 41, 708-712.
- xxix. Sakaquaptewa, D., Espinoza, P., Thompson, M., Vargas, P., & von Hippel, W. (2003). Stereotypic explanatory bias: Implicit stereotyping as a predictor of discrimination. *Journal of Experimental Social Psychology*, 39, 75-82.
- xxx. Maass, A., Salvi, D., Arcuri, L., & Semin, G. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology*, 57, 981-993.
- xxxi. see von Hippel, W., Sakaquaptewa, D., & Vargas, P. (1997). The linguistic intergroup bias as an implicit indicator of prejudice. *Journal of Experimental Social Psychology*, 33, 490-509. Sakaquaptewa, D., Espinoza, P., Thompson, M., Vargas, P., & von Hippel, W. (2003). Stereotypic explanatory bias: Implicit stereotyping as a predictor of discrimination. *Journal of Experimental Social Psychology*, 39, 75-82.



Racial Justice,
Equity and Inclusion



State Justice Institute