

Race & Ethnic Fairness in the Courts

Implicit Bias

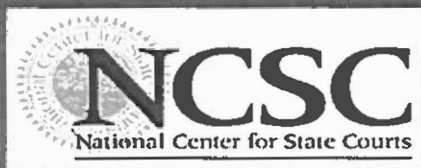
A Primer for Courts

Jerry Kang

Prepared for the National Campaign to Ensure the Racial
and Ethnic Fairness of America's State Courts

August 2009

©National Center for State Courts, 2009



ABOUT THE PRIMER

This Primer was produced as part of the National Campaign to Ensure the Racial and Ethnic Fairness of America's State Courts. The Campaign seeks to mobilize the significant expertise, experience, and commitment of state court judges and court officers to ensure both the perception and reality of racial and ethnic fairness across the nation's state courts. The Campaign is funded by the Open Society Institute, the State Justice Institute, and the National Center for State Courts. Points of view or opinions expressed in the Primer are those of the author and do not represent the official position of the funding agencies. To learn more about the Campaign, visit www.nesconline.org/ref.

ABOUT THE AUTHOR & REVIEWERS

Jerry Kang is Professor of Law at UCLA School of Law. He has written and lectured extensively on the role of implicit bias in the law. For more information on Professor Kang, please visit jerrykang.net. The Primer benefited from the review and comments of several individuals working with the National Campaign, including Dr. Pamela Casey, Dr. Fred Cheesman, Hon. Ken M. Kawaichi, Hon. Robert Lowenbach, Dr. Shawn Marsh, Hon. Patricia M. Martin, Ms. Kimberly Papillon, Hon. Louis Trosch, and Hon. Roger K. Warren.

Table of Contents

Implicit Bias: A Primer.....	1
Schemas and Implicit Cognitions.....	1
Implicit Social Cognitions.....	1
Asking about Bias.....	2
Implicit measurement devices.....	2
Pervasive implicit bias.....	3
Real-world consequences.....	4
Malleability	4
The big picture	5
Glossary.....	7
Bibliography	10

Implicit Bias: A Primer

Schemas and Implicit Cognitions (or “mental shortcuts”)

Stop for a moment and consider what bombards your senses every day. Think about everything you see, both still and moving, with all their color, detail, and depth. Think about what you hear in the background, perhaps a song on the radio, as you decode lyrics and musical notes. Think about touch, smell, and even taste. And while all that’s happening, you might be walking or driving down the street, avoiding pedestrians and cars, chewing gum, digesting your breakfast, flipping through email on your smartphone. How does your brain do all this simultaneously?

It does so by processing through schemas, which are templates of knowledge that help us organize specific examples into broader categories. When we see, for example, something with a flat seat, a back, and some legs, we recognize it as a “chair.” Regardless of whether it is plush or wooden, with wheels or bolted down, we know what to do with an object that fits into the category “chair.” Without spending a lot of mental energy, we simply sit. Of course, if for some reason we have to study the chair carefully--because we like the style or think it might collapse--we can and will do so. But typically, we just sit down.

We have schemas not only for objects, but also processes, such as how to order food at a restaurant. Without much explanation, we know what it means when a smiling person hands us laminated paper with detailed

descriptions of food and prices. Even when we land in a foreign airport, we know how to follow the crazy mess of arrows and baggage icons toward ground transportation.

These schemas are helpful because they allow us to operate without expending valuable mental resources. In fact, unless something goes wrong, these thoughts take place automatically without our awareness or conscious direction. In this way, most cognitions are implicit.

Implicit Social Cognitions (or “thoughts about people you didn’t know you had”)

What is interesting is that schemas apply not only to objects (e.g., “chairs”) or behaviors (e.g., “ordering food”) but also to human beings (e.g., “the elderly”). We naturally assign people into various social categories divided by salient and chronically accessible traits, such as age, gender, race, and role. And just as we might have implicit cognitions that help us walk and drive, we have implicit social cognitions that guide our thinking about social categories. Where do these schemas come from? They come from our experiences with other people, some of them direct (i.e., real-world encounters) but most of them vicarious (i.e., relayed to us through stories, books, movies, media, and culture).

If we unpack these schemas further, we see that some of the underlying cognitions include stereotypes, which are simply traits that we associate with a category. For instance, if we think that a particular category of human beings is frail--such as

the elderly--we will not raise our guard. If we think that another category is foreign--such as Asians--we will be surprised by their fluent English. These cognitions also include attitudes, which are overall, evaluative feelings that are positive or negative. For instance, if we identify someone as having graduated from our beloved alma mater, we will feel more at ease. The term "implicit bias" includes both implicit stereotypes and implicit attitudes.

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?

Or have we already internalized the lessons of Martin Luther King, Jr. and navigate life in a perfectly "colorblind" (or gender-blind, ethnicity-blind, class-blind, etc.) way?

Asking about Bias (or "it's murky in here")

One way to find out about implicit bias is simply to ask people. However, in a post-civil rights environment, it has become much less useful to ask explicit questions on sensitive topics. We run into a "willing and able" problem.

First, people may not be willing to tell pollsters and researchers what they really feel. They may be chilled by an air of political correctness.

Second, and more important, people may not know what is inside their heads. Indeed, a wealth of cognitive psychology has demonstrated that we are lousy at introspection. For example, slight environmental changes alter our judgments and behavior without our realizing. If the room smells of Lysol, people eat more neatly. People holding a warm cup of coffee (versus a cold cup) ascribe warmer (versus cooler) personality traits to a stranger described in a vignette. The experiments go on and on. And recall that by definition, implicit biases are those that we carry without awareness or conscious direction. So how do we know whether we are being biased or fair-and-square?

Implicit measurement devices (or "don't tell me how much you weigh, just get on the scale")

In response, social and cognitive psychologists with neuroscientists have tried to develop instruments that measure stereotypes and attitudes, without having to rely on potentially untrustworthy self-reports. Some instruments have been linguistic, asking folks to write out sentences to describe a certain scene from a newspaper article. It turns out that if someone engages in stereotypical behavior, we just describe what happened. If it is counter-typical, we feel a need to explain what happened. (Von Hippel 1997; Sekaquaptewa 2003).

Others are physiological, measuring how much we sweat, how our blood pressure changes, or even which regions of our brain light up on an fMRI (functional magnetic resonance imaging) scan. (Phelps 2000).

Still other techniques borrow from marketers. For instance, conjoint analysis asks people to give an overall evaluation to slightly different product bundles (e.g., how do you compare a 17" screen laptop with 2GB memory and 3 USB ports, versus a 15" laptop with 3 GB of memory and 2 USB ports). By offering multiple rounds of choices, one can get a measure of how important each feature is to a person even if she had no clue to the question "How much would you pay for an extra USB port?" Recently, social cognitionists have adapted this methodology by creating "bundles" that include demographic attributes. For instance, how would you rank a job with the title Assistant Manager that paid \$160,000 in Miami working for Ms. Smith, as compared to another job with the title Vice President that paid \$150,000 in Chicago for Mr. Jones? (Caruso 2009).

Scientists have been endlessly creative, but so far, the most widely accepted instruments have used reaction times--some variant of which has been used for over a century to study psychological phenomena. These instruments draw on the basic insight that any two concepts that are closely associated in our minds should be easier to sort together. If you hear the word "moon," and I then ask you to think of a laundry detergent, then "Tide" might come more quickly to

mind. If the word "RED" is painted in the color red, we will be faster in stating its color than the case when the word "GREEN" is painted in red.

Although there are various reaction time measures, the most thoroughly tested one is the Implicit Association Test (IAT). It is a sort of video game you play, typically on a computer, where you are asked to sort categories of pictures and words. For example, in the Black- White race attitude test, you sort pictures of European American faces and African American faces, Good words and Bad words in front of a computer. It turns out that most of us respond more quickly when the European American face and Good words are assigned to the same key (and African American face and Bad words are assigned to the other key), as compared to when the European American face and Bad words are assigned to the same key (and African American face and Good words are assigned to the other key). This average time differential is the measure of implicit bias. [If the description is hard to follow, try an IAT yourself at Project Implicit.]

Pervasive implicit bias (or "it ain't no accident")

It may seem silly to measure bias by playing a sorting game (i.e. the IAT). But, a decade of research using the IAT reveals pervasive reaction time differences in every country tested, in the direction consistent with the general social hierarchies: German over Turk (in Germany), Japanese over Korean (for Japanese), White over Black, men over women (on the stereotype of "career" versus "family"), light- skinned over dark skin,

youth over elderly, straight over gay, etc. These time differentials, which are taken to be a measure of implicit bias, are systematic and pervasive. They are statistically significant and not due to random chance variations in measurements.

These pervasive results do not mean that everyone has the exact same bias scores. Instead, there is wide variability among individuals. Further, the social category you belong to can influence what sorts of biases you are likely to have. For example, although most Whites (and Asians, Latinos, and American Indians) show an implicit attitude in favor of Whites over Blacks, African Americans show no such preference on average. (This means, of course, that about half of African Americans do prefer Whites, but the other half prefer Blacks.)

Interestingly, implicit biases are dissociated from explicit biases. In other words, they are related to but differ sometimes substantially from explicit biases--those stereotypes and attitudes that we expressly self-report on surveys. The best understanding is that implicit and explicit biases are related but different mental constructs. Neither kind should be viewed as the solely "accurate" or "authentic" measure of bias. Both measures tell us something important.

Real-world consequences (or "why should we care?")

All these scientific measures are intellectually interesting, but lawyers care most about real- world consequences. Do these measures of implicit bias predict an

individual's behaviors or decisions? Do milliseconds really matter? (Chugh 2004). If, for example, well-intentioned people committed to being "fair and square" are not influenced by these implicit biases, then who cares about silly video game results?

There is increasing evidence that implicit biases, as measured by the IAT, do predict behavior in the real world--in ways that can have real effects on real lives. Prof. John Jost (NYU, psychology) and colleagues have provided a recent literature review (in press) of ten studies that managers should not ignore. Among the findings from various laboratories are:

- implicit bias predicts the rate of callback interviews (Rooth 2007, based on implicit stereotype in Sweden that Arabs are lazy);
- implicit bias predicts awkward body language (McConnell & Leibold 2001), which could influence whether folks feel that they are being treated fairly or courteously;
- implicit bias predicts how we read the friendliness of facial expressions (Hugenberg & Bodenhausen 2003);
- implicit bias predicts more negative evaluations of ambiguous actions by an African American (Rudman & Lee 2002), which could influence decisionmaking in hard cases;
- implicit bias predicts more negative evaluations of agentic (i.e. confident, aggressive, ambitious) women in certain hiring conditions (Rudman & Glick 2001);
-

implicit bias predicts the amount of shooter bias--how much easier it is to shoot African Americans compared to Whites in a videogame simulation (Glaser & Knowles 2008);

- implicit bias predicts voting behavior in Italy (Arcari 2008);
- implicit bias predicts binge-drinking (Ostafin & Palfai 2006), suicide ideation (Nock & Banaji 2007), and sexual attraction to children (Gray 2005).

With any new scientific field, there remain questions and criticisms--sometimes strident. (Arkes & Tetlock 2004; Mitchell & Tetlock 2006). And on-the-merits skepticism should be encouraged as the hallmark of good, rigorous science. But most scientists studying implicit bias find the accumulating evidence persuasive. For instance, a recent meta-analysis of 122 research reports, involving a total of 14,900 subjects, revealed that in the sensitive domains of stereotyping and prejudice, implicit bias IAT scores better predict behavior than explicit self-reports. (Greenwald et al. 2009).

And again, even though much of the recent research focus is on the IAT, other instruments and experimental methods have corroborated the existence of implicit biases with real world consequences. For example, a few studies have demonstrated that criminal defendants with more Afro-centric facial features receive in certain contexts more severe criminal punishment (Banks et al. 2006; Blair 2004).

Malleability (or "is there any good news?")

The findings of real-world consequence are disturbing for all of us who sincerely believe that we do not let biases prevalent in our culture infect our individual decisionmaking. Even a little bit.

Fortunately, there is evidence that implicit biases are malleable and can be changed.

- An individual's motivation to be fair does matter. But we must first believe that there's a potential problem before we try to fix it.
- The environment seems to matter. Social contact across social groups seems to have a positive effect not only on explicit attitudes but also implicit ones.
- Third, environmental exposure to countertypical exemplars who function as "debiasing agents" seems to decrease our bias.
 - In one study, a mental imagery exercise of imagining a professional business woman (versus a Caribbean vacation) decreased implicit stereotypes of women. (Blair et al. 2001).
 - Exposure to "positive" exemplars, such as Tiger Woods and Martin Luther King in a history questionnaire, decreased implicit bias against Blacks. (Dasgupta & Greenwald 2001).
 - Contact with female professors and deans decreased implicit bias against women for college-aged women. (Dasgupta & Asgari 2004).
- Fourth, various procedural changes can disrupt the link between implicit bias and discriminatory behavior.
 - In a simple example, orchestras started using a blind screen in auditioning new musicians; afterwards women had much

greater success. (Goldin & Rouse 2000).

- In another example, by committing beforehand to merit criteria (is book smarts or street smarts more important?), there was less gender discrimination in hiring a police chief. (Uhlmann & Cohen 2005).
- In order to check against bias in any particular situation, we must often recognize that race, gender, sexual orientation, and other social categories may be influencing decisionmaking. This recognition is the opposite of various forms of “blindness” (e.g., color- blindness).

In outlining these findings of malleability, we do not mean to be Pollyanish. For example, mere social contact is not a panacea since psychologists have emphasized that certain conditions are important to decreasing prejudice (e.g., interaction on equal terms; repeated, non-trivial cooperation). Also, fleeting exposure to countertypical exemplars may be drowned out by repeated exposure to more typical stereotypes from the media (Kang 2005).

Even if we are skeptical, the bottom line is that there’s no justification for throwing our hands up in resignation. Certainly the science doesn’t require us to. Although the task is challenging, we can make real improvements in our goal toward justice and fairness.

The big picture (or “what it means to be a faithful steward of the judicial system”)

It’s important to keep an eye on the big picture. The focus on implicit bias does not address the existence and impact of explicit bias--the stereotypes and attitudes that folks recognize and embrace. Also, the past has an inertia that has not dissipated. Even if all explicit and implicit biases were wiped away through some magical wand, life today would still bear the burdens of an unjust yesterday. That said, as careful stewards of the justice system, we should still strive to take all forms of bias seriously, including implicit bias.

After all, Americans view the court system as the single institution that is most unbiased, impartial, fair, and just. Yet, a typical trial courtroom setting mixes together many people, often strangers, from different social backgrounds, in intense, stressful, emotional, and sometimes hostile contexts. In such environments, a complex jumble of implicit and explicit biases will inevitably be at play. It is the primary responsibility of the judge and other court staff to manage this complex and bias-rich social situation to the end that fairness and justice be done--and be seen to be done.

Glossary

Note: Many of these definitions draw from Jerry Kang & Kristin Lane, A Future History of Law and Implicit Social Cognition (unpublished manuscript 2009)

Attitude

An attitude is “an association between a given object and a given evaluative category.” R.H. Fazio, et al., Attitude accessibility, attitude- behavior consistency, and the strength of the object-evaluation association, 18 J.

EXPERIMENTAL SOCIAL PSYCHOLOGY 339, 341(1982).

Evaluative categories are either positive or negative, and as such, attitudes reflect what we like and dislike, favor and disfavor, approach and avoid. See also stereotype.

Behavioral realism

A school of thought within legal scholarship that calls for more accurate and realistic models of human decision-making and behavior to be incorporated into law and policy. It involves a three step process:

First, identify advances in the mind and behavioral sciences that provide a more accurate model of human cognition and behavior.

Second, compare that new model with the latent theories of human behavior and decision- making embedded within the law. These latent theories typically reflect “common sense” based on naïve psychological theories.

Third, when the new model and the latent theories are discrepant, ask lawmakers and legal institutions to account for this disparity. An accounting requires either altering the law to comport with more accurate models of thinking and behavior or providing a transparent explanation of “the prudential, economic, political, or religious reasons for retaining a less accurate and outdated view.” Kristin Lane, Jerry Kang, & Mahzarin Banaji, Implicit Social Cognition and the Law, 3 ANNU. REV. LAW SOC. SCI. 19.1-19.25 (2007)

Dissociation

Dissociation is the gap between explicit and implicit biases. Typically, implicit biases are larger, as measured in standardized units, than explicit biases. Often, our explicit biases may be close to zero even though our implicit biases are larger.

There seems to be some moderate-strength relation between explicit and implicit biases. See Wilhelm Hofmann, A Meta-Analysis on the Correlation Between the Implicit Association Test and Explicit Self-Report Measures, 31 PERSONALITY & SOC. PSYCH. BULL. 1369 (2005) (reporting mean population correlation $r=0.24$ after analyzing 126 correlations). Most scientists reject the idea that implicit biases are the only “true” or “authentic” measure; both explicit and implicit biases contribute to a full understanding of bias.

Explicit

Explicit means that we are aware that we have a particular thought or feeling. The term sometimes also connotes that we have

an accurate understanding of the source of that thought or feeling. Finally, the term often connotes conscious endorsement of the thought or feeling. For example, if one has an explicitly positive attitude toward chocolate, then one has a positive attitude, knows that one has a positive attitude, and consciously endorses and celebrates that preference. See also implicit.

Implicit

Implicit means that we are either unaware of or mistaken about the source of the thought or feeling. R. Zajonc, *Feeling and thinking: Preferences need no inferences*, 35 *AMERICAN PSYCHOLOGIST* 151 (1980). If we are unaware of a thought or feeling, then we cannot report it when asked. See also explicit.

Implicit Association Test

The IAT requires participants to classify rapidly individual stimuli into one of four distinct categories using only two responses (for example, in a the traditional computerized IAT, participants might respond using only the “E” key on the left side of the keyboard, or “I” on the right side). For instance, in an age attitude IAT, there are two social categories, YOUNG and OLD, and two attitudinal categories, GOOD and BAD. YOUNG and OLD might be represented by black-and-white photographs of the faces of young and old people. GOOD and BAD could be represented by words that are easily identified as being linked to positive or negative affect, such as “joy” or “agony”. A person with a negative implicit attitude toward OLD would be expected to

go more quickly when OLD and BAD share one key, and YOUNG and GOOD the other, than when the pairings of good and bad are switched.

The IAT was invented by Anthony Greenwald and colleagues in the mid 1990s. Project Implicit, which allows individuals to take these tests online, is maintained by Anthony Greenwald (Washington), Mahzarin Banaji (Harvard), and Brian Nosek (Virginia).

Implicit Attitudes

“Implicit attitudes are introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects.” Anthony Greenwald & Mahzarin Banaji, *Implicit social cognition: attitudes, self-esteem, and stereotypes*, 102 *Psychol. Rev.* 4, 8 (1995). Generally, we are unaware of our implicit attitudes and may not endorse them upon self-reflection. See also attitude; implicit.

Implicit Biases

A bias is a departure from some point that has been marked as “neutral.” Biases in implicit stereotypes and implicit attitudes are called “implicit biases.”

Implicit Stereotypes

“Implicit stereotypes are the introspectively unidentified (or inaccurately identified) traces of past experience that mediate attributions of qualities to members of a social category” Anthony Greenwald & Mahzarin Banaji, *Implicit social cognition:*

attitudes, self-esteem, and stereotypes, 102
Psychol. Rev. 4, 8 (1995).

Generally, we are unaware of our implicit stereotypes and may not endorse them upon self-reflection. See also stereotype; implicit.

Implicit Social Cognitions

Social cognitions are stereotypes and attitudes about social categories (e.g., Whites, youths, women). Implicit social cognitions are implicit stereotypes and implicit attitudes about social categories.

Stereotype

A stereotype is an association between a given object and a specific attribute. An example is “Norwegians are tall.” Stereotypes may support an overall attitude. For instance, if one likes tall people and Norwegians are tall, it is likely that this attribute will contribute toward a positive orientation toward Norwegians. See also attitude.

Validities

To decide whether some new instrument and findings are valid, scientists often look for various validities, such as statistical conclusion validity, internal validity, construct validity, and predictive validity.

- Statistical conclusion validity asks whether the correlation is found between independent and dependent variables have been correctly computed.
- Internal validity examines whether in addition to correlation, there has been a demonstration of causation. In

particular, could there be potential confounds that produced the correlation?

- Construct validity examines whether the concrete observables (the scores registered by some instrument) actually represent the abstract mental construct that we are interested in. As applied to the IAT, one could ask whether the test actually measures the strength of mental associations held by an individual between the social category and an attitude or stereotype
- Predictive validity examines whether some test predicts behavior, for example, in the form of evaluation, judgment, physical movement or response. If predictive validity is demonstrated in realistic settings, there is greater reason to take the measures seriously.

APPENDIX III: BIBLIOGRAPHY OF MATERIALS ON IMPLICIT BIAS

Mahzarin R. Banaji & Anthony G. Greenwald, *Blindspot: Hidden Biases of Good People*, Bantam, 2016.

Adam Benforado, *Unfair: The New Science of Criminal Injustice*, Crown, New York, 2016.
Jennifer L. Eberhardt PhD, *Biased: Uncovering the Hidden Prejudice That Shapes What We See, Think, and Do*, Viking, 2020.

Elizabeth Hinton et al., Vera Inst., *An Unjust Burden: The Disparate Treatment of Black Americans in the Criminal Justice System 2* (2018),
<https://www.vera.org/downloads/publications/for-the-record-unjust-burden-racial-disparities.pdf>.

Jerry Kang & Kristin Lane, *Seeing Through Colorblindness: Implicit Bias and the Law*, 58 UCLA L. Rev. 465, 471 (2010).

Reginald, King, et al., *A Punishing Look: Skin Tone and Afrocentric Features in the Halls of Justice*, *American Journal of Sociology* 122(1):90-124 (2016).

Praatika Prasad, *Implicit Racial Biases in Prosecutorial Summations: Proposing an Integrated Response*, 86 Fordham L. Rev. 667 (2018).

