What Judges Can Do About Implicit Bias

Jerry Kang

"Implicit bias" was not well known in legal communities twenty years ago. But now, the idea of implicit bias circulates widely in both popular and academic discussions. Even the casually interested judge knows a great deal about the topic. Still, even as the problem of implicit bias has grown familiar, potential solutions remain out of focus. Specifically, what can judges do about implicit bias, in their capacities as managers of a workplace, as well as vessels of state power?

In 2009 I wrote a Primer for the National Center for State Courts, which described the challenge of implicit bias to judicial audiences.[1] In 2012, I was the lead author of a more systematic examination titled *Implicit Bias in the Courtroom*.[2] That author team included not only legal scholars but also psychology professors such as the inventor of the Implicit Association Test (IAT), as well as a sitting federal judge. Together, we described the then-state-of-the-art and recommended potential countermeasures.

The goal of this article, nearly a decade later, is to update the scientific understanding since 2012. It also revises, reorganizes, and streamlines recommendations for judges who believe that implicit bias is a genuine problem but don't know what to do about it.[3] To keep length manageable, it focuses on the challenge of implicit biases held by judges themselves and does not directly address the biases held by others, such as police officers, probation officers, prosecutors, and jurors. It also focuses mostly on individual-level responses that judges can take themselves although institutional-level reforms may be what's most important.

Before jumping to recommendations, let's begin with clear definitions and a scientific update.

I. WHAT IS IMPLICIT BIAS? THE IDEA

BIAS AS ATTITUDE OR STEREOTYPE

Let's start by defining "implicit bias." Focus first on the noun: "bias." Bias just means deviation from some baseline of comparison, which is presumably neutral or fair. Because we are thinking about human beings, bias here means some deviation in an attitude or stereotype about a social category, such as Black women, immigrants, or the elderly.

The words "attitude" and "stereotype" are terms-of-art in social psychology. An "attitude" is an overall evaluative valence toward a category, which ranges from positive to negative. To take an uncontroversial example, some people prefer dogs to cats. Their attitude toward dogs is positive whereas their attitude toward cats is less so and sometimes even negative.

More narrow and particular than a global attitude is a "stereotype," which is a specific trait that is probabilistically associated with a category. Consider the traits of "loyal" or "finicky" and how they might be more strongly associated with dogs versus cats, especially for dog lovers. Of course, we know that not all dogs are loyal, and not all cats are finicky—however, those traits might be scientifically measured. But almost all of us have stereotypes about these pet categories. We tend to "profile" animals and don't feel especially embarrassed in doing so.

To sum up (and return to human categories), *a bias is an attitude or stereotype about a social category that departs from some designated baseline*. To measure racial bias against non-Whites, we might select that baseline to be the attitude toward Whites. On gender bias against women, we might designate the baseline to be stereotypes about men. And so on.

EXPLICIT V. IMPLICIT BIAS

What does the adjective "implicit" add to the term? To understand implicit, it's easier to start with its opposite "explicit." Although understandable, it's a mistake to think of "explicit" in the way that that word is used in terms like "explicit lyrics" or "explicit violence." Explicit bias need not be graphic, extreme, or large in magnitude although it sometimes is. Instead, it's better to understand "explicit" as being *subject to direct introspection*.

Let's return to cats and dogs. Suppose I ask you what you think about cats. This is not a hard question. Suppose you adore cats. Indeed, you have a ragdoll purring on your lap right now. The fact that you're able to ask yourself and get a clear, immediate answer back through direct introspection means that you have accessed and reported an explicit attitude. And because there isn't much stigma about loving cats (at least in the United States), there's little pressure for you to conceal that explicit attitude from others.

By contrast, an "implicit" bias is an attitude or stereotype that

Author's Note: Thanks to Judge Mark W. Bennett (retired), Devon Carbado, Magistrate Judge Sarah Cave, Rachel Godsil, Sung Hui Kim, Calvin Lai, Judge David Prince, Judge Lorna Schofield, and Stephen Yeazell.

Footnotes

1. Jerry Kang, Implicit Bias: A Primer for Courts (Aug. 2009) (prepared for the National Center for State Courts) (available at http://jerrykang.net/research/2009-implicit-bias-primer-for-courts/).
2. Jerry Kang et al., *Implicit Bias in the Courtroom*, 59 UCLA L. Rev. 1124 (2012). The author team included Anthony G. Greenwald, who invented the IAT, and then District Judge Mark W. Bennett for the Northern District of Iowa.
3. Thoughtful advice has, of course, already been given to judges throughout the years. *See, e.g.*, Bernice B. Donald & Sarah E. Redfield, *Implicit Bias: Should the Legal Community Be Bothered?*, 2 PLI Current 615 (20818); Pamela M. Casey et al., *Addressing Implicit Bias in the Courts*, 49 Ct. Rev. 442 (2013) (Casey was also on the author team of the 2012 UCLA article).

is *not* subject to direct introspection, or at least not easily.[4] In other words, we cannot easily or accurately measure implicit social cognitions by asking ourselves direct questions about our attitudes and stereotypes. How can this be? Suppose you were born in a country with a culture that despised cats. That preference suffused childhood bedtime stories, holidays (think some version of Halloween), and the fact that all the rich, powerful, and beautiful people on television had dogs, not cats. But then suppose as a teenager you emigrated to a new country that espoused equal treatment for dogs and cats. In this new land, you learned not to dislike cats and stopped generalizing about them. After all, they weren't *all* dirty and diseased, roaming the alleys for vermin, incessantly screeching for food. As you enter middle age, after significant cultural assimilation and personal growth, when asked directly whether you prefer dogs to cats, you pause, mull it over briefly, and honestly report that you have no preferences either way. Your explicit attitudes have changed. Terrific. Nevertheless, is it possible that you still retain traces of that negative feline attitude?

Our understanding of human memories suggests that it is indeed possible. It's this plausible hypothesis—that we retain attitudes and stereotypes that we cannot readily access—that prompted scientists to devise novel instruments with which to measure implicit associations. To repeat, truthfully answering an anonymous survey will not suffice. Instead, we need some external instrument to measure implicit biases. One category of such instruments measures reaction times to differing stimuli flashed quickly on a computer screen. A prominent example is the well-known Implicit Association Test (IAT) invented by Anthony Greenwald based on theoretical work done together with Mahzarin Banaji.[5]

Current research suggests that the ideas of explicit bias and implicit bias are overlapping but independent constructs. Neither one is more authentic or real than the other. Each construct does its own work and must be measured in its own way. Because explicit bias is subject to direct introspection, it is typically measured by scientists through a survey or questionnaire, with the hope that participants answer honestly. As judges know, that hope is not always well-founded, especially on socially sensitive or inculpatory topics. By contrast, because implicit bias is not readily subject to direct introspection, direct questioning is largely pointless. It must be measured some other way.

II. HOW DO WE MEASURE IMPLICIT BIAS? THE MANY INSTRUMENTS, INCLUDING THE IAT

Experimental social psychologists have developed multiple instruments. Recently, Calvin Lai and Megan Wilson compiled an inventory of 18 different implicit bias instruments (or tasks) organized into three categories: (1) the Implicit Association Test (IAT) and its variants; (2) priming tasks (where brief exposure to priming stimuli facilitates or inhibits subsequent reactions); and (3) miscellaneous other tasks, including linguistic or writing exercises.[6] The length of this list reminds us to disentangle the *idea* of implicit bias from any particular *instrument* by which it is measured.

"The IAT is the most used and best validated instrument for measuring implicit bias."

The pandemic, which is top of mind, provides useful analogies. We have learned that there are multiple tests (using blood, spit, swabs, etc.) with different sensitivities, specificities, and reliabilities to determine whether anyone has or had a COVID-19 infection. We also roughly understand what false positives and false negatives mean for such tests. Few of us would confuse the *infection* for the *instrument* by which infection is detected. We should do the same with implicit bias and remember that the *idea* of implicit bias is separate from any specific *instrument* to detect that bias, including the exhaustively studied Implicit Association Test (IAT).

The IAT is the most used and best validated instrument for measuring implicit bias. I think of it as a sort of "videogame" requiring fast sorting of stimuli representing two social categories (e.g., White faces versus Black faces) and two sets of words representing, for example, a positive versus negative attitude. Sometimes the stimuli require keyboard responses that are consistent with our implicit social cognitions, and sometimes inconsistent. The former responses come faster than the latter. The raw reaction time delta, which is typically a few hundred milliseconds, is mathematically processed and transformed into statistical units that crudely signal the strength of the underlying implicit association. Since this test has been described extensively elsewhere,[7] I won't do so here. But if you're unfamiliar, try taking one for free, anonymously, at Project Implicit.[8]

Millions of people have already done so, in the United States and around the globe. The first systematic analysis of the pervasiveness and correlates of implicit attitudes and stereotypes, as measured by the IAT, was conducted by Brian Nosek and colleagues in 2007 (describing data collected on 17 different tests at Project Implicit during 2000-2006).[9] They found that implicit bias—as measured by the IAT—was pervasive. I have it. You have it. Not in precisely the same amounts, toward the same categories, but we all have it. Implicit bias was also found to be larger in magnitude than self-reported explicit bias.

Recently, Kate Ratliff and colleagues compiled an update with Project Implicit data from 2007-2015. They again found that implicit bias "favoring culturally dominant or societally valued

4. I add the qualifier because of recent work suggesting that implicit social cognitions may be preconscious, subject to some forms of introspection when it is guided by concrete stimuli and directions to pay attention to immediate affective responses. *See generally* Adam Hahn & Alexandra Goedderz, *Trait-Unconsciousness, State-Unconsciousness, Pre-Consciousness, and Social Miscalibration in the Context of the Implicit Evaluation*, 38 Soc. Cognition S115 (2020) (supplement).
5. *See* Anthony G. Greenwald et al., *Measuring Individual Differences in Implicit Cognition: The Implicit Association Test*, 74 J. Personality & Soc. Psychol. 1464, 1464–66 (1998) (introducing the IAT); Anthony G. Greenwald & Mahzarin R. Banaji, *Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes*, 102 Psychol. Rev. 4 (1995).
6. Calvin K. Lai & Megan E. Wilson, *Measuring Implicit Intergroup Biases*, 15 Soc. & Personality Psychol. Compass 1 (2021).
7. *See, e.g.*, Kristin A. Lane, Jerry Kang, and Mahzarin R. Banaji, *Implicit Social Cognition and the Law*, 3 Annu. Rev. Law Soc. Sci. 19.1, .2-.3 (2007); Jerry Kang & Kristin Lane, *Seeing through Colorblindness: Implicit Bias and the Law*, 58 UCLA L. Rev. 465, 472-73 (2010).
8. *See* <http://projectimplicit.org>.
9. Brian A. Nosek et al., *Pervasiveness and Correlates of Implicit Attitudes and Stereotypes*, 18 Eur. Rev. Soc. Psychol. 1 (2007).

"So, what should skeptical judges ... make of all this?"

groups" remains pervasive and stronger in magnitude than explicit bias. They also found that "ingroups are evaluated more positively than outgroups."[10] This finding underscores the importance of being on guard not only against outgroup derogation but also ingroup favoritism, which some scholars believe to be the dominant source of discrimination in modern times.[11] Overall, these large data set analyses are consistent with IAT data generated from experiments conducted in hundreds of laboratories around the world over the past two decades.

To sum up: When asked if we are colorblind (or genderblind, etc.), we may scratch our heads, then with all sincerity reply that we judge people only by the content of their character not the color of their skin (or gender, etc.). But at least on the IAT sorting game, it isn't so. Most of us tend to respond faster when White faces (as compared to Black faces) are on the same response key as Good words. Most of us tend to respond faster when Black faces are on the same response key as weapons (as compared to harmless objects). And so on. Lawyers, judges, and professors regardless of fancy degrees are no exception.[12]

III. WHY DOES IMPLICIT BIAS MATTER? THE IMPACT

By now, implicit bias enthusiasts may be losing patience. They are thankful for the refresher but want to cut to the chase and find out how to solve the problem. (For those who can't wait any longer, please skip to Part IV and the Appendix.) But skeptical readers still have questions, including whether these sorting asymmetries predict real-world behavior, like worse treatment? Judges are sophisticated enough to know that simply because something is *statistically* significant (and not likely due to chance) does not mean it is *socially* significant (and worthy of individual or institutional reform).

The topline answer is that implicit bias does predict discriminatory behavior, but to a low degree. The best way to avoid cherry-picking studies is to review meta-analyses. (A meta-analysis takes all studies that can be found in the relevant domain and stitches together their findings into a single composite number, usually the "effect size." In this literature, the effect size is the degree of correlation between an implicit bias measure and discriminatory behavior. The correlation is indicated by Pearson's r, which runs from 0, which means no relationship between bias and behavior, to ±1, which means a perfectly linear positive or negative relationship.) Multiple meta-analyses have been conducted specifically on IAT scores. Although differing in important ways, they all tend to show that IAT scores predict intergroup discriminatory behavior at a very low level. (The range of r values goes from 0.24 down to 0.10.[13] By convention, r values greater than or equal to 0.1, 0.3, and 0.5 are called small, medium, and large, respectively.) The small effect size that has been found should not be surprising given how crude an instrument the IAT is and how hard it is to measure discriminatory behavior, especially in realistic contexts. And imprecise measures of any two variables (in this case bias and behavior) make it difficult to discern the strength of any relationship between those two variables.

So, what should skeptical judges (who are unlikely also to be professional statisticians) make of all this? First, consider a direct comparison that comes from every jury trial you've heard. The meta-analyses generally confirm that implicit measures of bias predict intergroup discriminatory behavior better than explicit measures of bias.[14] Ponder this the next time you or the attorneys ask potential jurors about their explicit biases during voir dire. What information do you think their self-reports really reveal? Whatever that is, implicit bias measures probably tell you more.

Second, consider a stylized BigLaw hypothetical, which demonstrates how even slight differences in treatment caused by implicit bias can create headwinds and tailwinds that accumulate powerfully over time. Greg and Brandie have just started as associates at an elite firm and are initiated into the partnership "hunger games." Each month they must survive an up-or-out decision based on that month's performance. If they can survive 8 years of these monthly cuts, they are elected to equity partnership and win life tenure filled with esteem, repose, and high remuneration. (I did warn you that this was stylized.) To make this simulation more realistic, suppose that the base rate of survival for all associates is a generous 98.5%.

10. *See* Kate A. Ratliff et al, Documenting Bias from 2007-2015: Pervasiveness and Correlates of Implicit Attitudes and Stereotypes II (unpublished preprint) at 2. The meta-analytic effect size for implicit bias was Cohen's d =.80 (by convention called "large") as compared to explicit bias of d = .51 (by convention called "medium"). *Id*. at 21.
11. *See* Anthony G. Greenwald & Thomas F. Pettigrew, *With Malice toward None and Charity for Some: Ingroup Favoritism Enables Discrimination*, 69 Am. Psychologist 669 (2014). Even if every ingroup favors itself equally, the population and resource advantage of certain groups will lock in a net advantage indefinitely.
12. *See, e.g.,* Mark W. Bennett, *The Implicit Racial Bias in Sentencing: the Next Frontier*, Yale L. J. Forum (January 31, 2017), at 396-397; Jeffrey J. Rachlinski et al., *Does Unconscious Racial Bias Affect Trial Judges?,* 84 Notre Dame L. Rev. 1195, 1210 (2009);Theodore Eisenberg & Sheri Lynn Johnson, *Implicit Racial Attitudes of Death Penalty Lawyers*, 53 DePaul L. Rev. 1539, 1545–55 (2004). *See also* Ratliff et al., *supra* note 10, at 18 (finding a correlation between implicit attitudes/stereotypes and education to be r = .005).
13. *See* Anthony G. Greenwald et al., *Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity*, 97 J. Personality & Soc. Psychol. 17, 19–20 (2009) (r=.024 for Black/White bias); Frederick Oswald et al., *Predicting Ethnic and Racial Discrimination: A Meta-Analysis of IAT Research*, 105 J. Personality & Soc. Psychol. 171 (2013) (r=.15 on Black/White implicit bias); Benedikt Kurdi et al., *Relationship between the Implicit Association Test and Intergroup Behavior: A Meta-Analysis*, 74 Am. Psychologist 569 (2019) (personal communication to Anthony Greenwald that r = .10 for behavior in the *combined* domains of Black/White, gender, sexual orientation, weight, and disabilities). The highest estimate (r=0.24) would mean that an IAT score predicts approximately 5.6% of the statistical variance in the discriminatory behavior variable.
14. *See, e.g.,* Greenwald et al., *supra* note 13, at 73, Table 3 (finding that implicit attitude scores predicted behavior in the Black/White domain at an average correlation of r=0.24, whereas explicit attitude scores had correlations of average r=0.12); Kurdi et al, *supra* note 13 (finding that implicit biases provide a unique contribution to predicting behavior (B = .14) and does so more than explicit measures (B = .11)).

But now let's superimpose implicit bias, which produces a slight tailwind for Greg, who happens to be a White man. He gets a half percent boost so that his monthly survival chance goes up to 99%. By contrast, Brandie, who happens to be Black woman, suffers a slight headwind, which means that her monthly survival chance goes down to 98%.[15] In other words, the invisible winds of implicit bias create a mere 1% delta on the monthly survival rate between these two identically talented associates.

Under these assumptions, what are the chances that Greg and Brandie make partner? Assuming that each month's probability is independent, we would simply multiply the probability for each month. Greg's survival chance would thus be 0.99 (for month 1) x 0.99 (for month 2) x … 0.99 (all the way up to month 96). Similarly, Brandie's survival chance would be 0.98 (for month 1) x (0.98 for month 2) . . . x 0.98 (all the way up to month 96). After eight years (or 96 cuts), it turns out that Greg's partnership chance is 38.1% (0.99^96 = .381). Brandie's is only 14.4% (0.98^96 = .144).

That's a stunning disparity driven by a tiny difference. How can this be? It's the power of compound interest. It's why we should start investing in our retirement accounts early. Little differences compounded over time have huge consequences on the trajectory of an individual (not to mention a stock portfolio). And if we aggregate this across an entire population of individuals (e.g., all men versus all women), little differences can generate huge societal impacts. In emphasizing the impact of implicit bias, I am not suggesting that explicit bias or "structural" bias (however that term is defined) are irrelevant or matter less in the real world. They all matter.[16] But I have one unique reason to focus on implicit bias. It's the one strain of bias that cannot be easily relegated to a few "bad apples," or extremists, or the history books. Implicit bias is here, right now, in your own courtroom, in your own mind, and in mine.

IV. WHAT TO DO ABOUT IMPLICIT BIAS? SOME EVIDENCE-BASED INTERVENTIONS

Given these inconvenient truths, most judges will want to do something about implicit bias if the interventions are practical and not too costly. What might judges do? I offer four strategies: deflate, debias, defend, and data. A list of specific tactics organized by strategy appears in the Appendix. Finally, to reiterate, these recommendations focus on challenges caused by implicit biases held by judges themselves and not by others, such as prosecutors or jurors. And they focus mostly on individual-level responses that judges can adopt by themselves as compared to institutional-level ones that would require substantial coordination.

"Justice Anthony Kennedy was right … that 'bias is easy to attribute to others and difficult to discern in oneself.'"

A. DEFLATE (YOUR EGO) AND EMBRACE FALLIBILITY

First, we must *deflate* our egos. We must recognize that we are not as objective, as fair, as virtuous as we view ourselves to be.[17] Justice Anthony Kennedy was right to observe that "bias is easy to attribute to others and difficult to discern in oneself."[18] Worse, thinking ourselves to be fair and objective leads us to perform worse on audit or tester studies.[19] When we confidently assume that we already get things right, we pay less attention and take less care in decision making. Paradoxically, only by assuming that we will be biased will we have any chance of being truly fair.

I want to highlight the related danger of "moral credentialing."[20] One danger of implicit bias education, which includes reading this article, is assuming that education has directly cured the malady. That, of course, is not the case. Education does not directly change behavior. For example, learning about mRNA and virus-replication doesn't directly generate antibodies or alter long-standing habits of touching our faces with our hands. Frankly, education is not even training—as you likely recall the difference between a law school doctrinal class versus a clinic with live clients. So, it behooves us to avoid the pride, confidence, and moral credentialing that creeps in with greater expertise.

Having deflated our egos, we should simultaneously cultivate our *internal motivation* to be fair.[21] Social psychology distinguishes between internal and external motivations for behavior. Internal motivation to be fair means that we are striving to achieve our personal values, consistent with our genuine ethical commitments. It's how we behave even when no one is watching, as we strive toward our ideal selves. By contrast, external motivation means that we feel more coerced than persuaded into the behavior. We are driven by fear that we will be shunned, punished, or canceled. As compared to internal motivation, external motivation to avoid appearing prejudiced is less helpful and may even backfire.[22] It's generally correlated with larger explicit biases that are concealed but eventually leak out in expression or behavior.

15. The 1% difference in this hypothetical is mathematically equivalent to r = .041, which is far smaller than the effect sizes found by the three meta-analyses. *See* Anthony G. Greenwald et al., *Importance of Small-to-Moderate IAT Effects*, 108 J. Personality & Soc. Psychol. 553, 558 (2015).
16. For discussion on the various layers of bias and their interactions, see Jerry Kang, *The Realities of Race*. 358 Sci. 1137 (2017) (book review); Jerry Kang, *Implicit Bias and Pushback from the Left*, 54 St. Louis L. Rev. 1139 (2010).
17. For slightly embarrassing evidence, consider the fact that 87% of (non-senior) federal district judges and 92% of senior federal district judges view themselves as in the top 25% of their colleagues in their ability to make decisions free from racial bias. This is not mathematically possible. *See* Bennett, *supra* note 12, at 396-97.
18. Williams v. Pennsylvania, 136 S. Ct. 1899, 1905 (2016).
19. *See* Eric Luis Uhlmann & Geoffrey L. Cohen, *"I Think It, Therefore It's True": Effects of Self-perceived Objectivity on Hiring Discrimination*, 104 Org. Behavior & Hum. Decision Processes 207, 210 (2007).
20. Benoît Monin & Dale T. Miller, *Moral Credentials and the Expression of Prejudice,* 81 J. Personality & Soc. Psychol. 33 (2001).
21. *See* Margo J. Monteith et al., *Schooling the Cognitive Monster: The Role of Motivation in the Regulation and Control of Prejudice*, 3 Soc. & Person. Psychol. Compass 211 (2009).
22. *See generally,* George V. Gushue & Kimberly A. Hinman, *Promoting Justice or Preventing Prejudice? Interrupting External Motivation in Multicultural Training*, 88 Am. J. Orthopsychiatry 142 (2018); Lisa Legault et al., *Ironic Effects of Anti-Prejudice Messages: How Motivational Interventions Can Reduce (but Also Increase) Prejudice*, 22 Psychol. Sci. 1472 (2011).

"Social contact generally decreases biases."

B. DEBIAS (WITH SHORT-TERM "SPOT CLEANING" AND LONG-TERM INTERACTIONS)

With this humble mindset, what else might we do? For example, can we simply delete the embarrassing or unwanted implicit biases from our brains so that our social cognitions line up with our explicit commitments? This is the *debiasing* strategy. Early research into implicit bias suggested that implicit social cognitions were highly malleable and could be changed by brief imagination exercises or exposures to people who defied stereotypes (think Black woman surgeon, male nurse, or Asian leading man).[23] But in the past ten years, that initial optimism has waned.

A useful place to start is the 2014 paper by Calvin Lai and colleagues, who created a tournament and invited scientists to submit quick interventions (no longer than five minutes) that could decrease implicit bias as measured by the Black/White IAT.[24] Here's the good news. Three categories of interventions, including exposure to counterstereotypical exemplars,[25] successfully decreased implicit bias scores. Now for the bad news. As Lai reported in 2016, none of these successes persisted for even a few days.[26] Put another way, there seems to be no quick fix that creates long-lasting or durable changes in implicit bias, as measured by the IAT. In retrospect, we should not be surprised. Our implicit associations were not created overnight. They are like old stains on a well-trodden carpet. Why should they magically disappear after a five-minute scrub?[27]

Given what we've learned, we should distinguish short-term and long-term debiasing tactics. In the *short term*, we might investigate ways to deploy "spot cleaning," even if the debiasing lasts only a few hours. To take an extreme example, in the tournament, Lai and colleagues found that imagining a vivid scenario of being beaten unconscious by a White sadist and saved by a Black hero produced a significant (although temporary) reduction in implicit bias. But I can't in good conscience recommend that judges start their day with a cappuccino and a five-minute contemplation of being tortured by White people. That would be awkward.

But it's not at all awkward to have pictures and other reminders of admired figures from racial minority communities within one's office, bookshelf, courtroom, and building.[28] Who are the "firsts" in your jurisdiction (first lawyer, first judge, first prosecutor, first law professor)? Are they celebrated on your walls and websites? Why not feature the new Americans, beaming with pride, who have recently been naturalized in your courthouse?[29] These techniques always risk being window dressing, but there may be some value in "spot cleaning" the built environment that surrounds you and thus constantly reminds you. Even if their value is ephemeral, they also serve an important expressive function that signals belonging to the diverse community members who enter the courthouse, often with anxiety and reservations.

The *long-term* debiasing tactics look different. If quick-and-dirty doesn't create lasting change, slow-and-steady wins the race. Social contact generally decreases biases, and the longer and greater the amount and depth of contact with members of other groups, including those who defy stereotypes, the greater the improvement.[30] Nilanjana Dasgupta has conducted field studies that suggest that repeat exposure, in the real world, to people who defy stereotypes and expectations decreases implicit biases. In one study, she and Shaki Asgari studied women who attended either an all-women's college or a comparable coed institution.[31] For the women who attended the coed institution, their implicit stereotypes (that Men = Leaders and Women = Supporters) surprisingly *increased* after freshman year of college. By contrast, the implicit stereotypes of women who attended the all-women's college decreased to an average of zero. After examining multiple variables, such as courses taken, extracurricular activities, and other campus variables, the one variable that mattered most was the number of women professors and deans they had. And students in the all-women's college were simply exposed to more women professors and deans.

23. *See generally* Irene V. Blair, *The Malleability of Automatic Stereotypes and Prejudice*, 6 PERSONALITY & SOC. PSYCHOL. REV. 242 (2002) (literature review).
24. Calvin K. Lai et al., *Reducing Implicit Racial Preferences: I. A Comparative Investigation of 17 Interventions*, 143 J. EXPERIMENTAL PSYCHOL. GEN. 1765 (2014).
25. The other two categories were called "intentional strategies to overcome biases" and "evaluative conditioning." These categories included techniques that will seem "Pavlovian" in the lay sense. It involved, for example, setting an intention of thinking good whenever one saw a Black face, or repeated exposures of Black faces with good words, and White faces with bad words. *See id.* at 1773-74.
26. Calvin K. Lai et al., *Reducing Implicit Preferences: II. Intervention Effectiveness across Time*, 145 J. EXPERIMENTAL PSYCHOL. GEN. 1001 (2016).
27. If my tongue-in-cheek use of "scrubbing" raises autonomy concerns, see my discussion of the "autonomy objection." *See* Jerry Kang, *Trojan Horses of Race*, 118 HARV. L. REV. 1489, 1584-89 (2005).
28. *See, e.g.,* Nilanjana Dasgupta & Anthony G. Greenwald, *On the Malleability of Automatic Attitudes: Combating Automatic Prejudice With Images of Admired and Disliked Individuals,* 81 J. PERSONALITY & SOC. PSYCHOL. 800, 807 (2001); See Bernd Wittenbrink et al., *Spontaneous Prejudice in Context: Variability in Automatically Activated Attitudes,* 81 J. PERSONALITY & SOC. PSYCHOL. 815, 818–19 (2001) (finding that situating African Americans in a positive versus negative setting, e.g., outdoor barbecue versus gang-related incident, produced lower implicit bias scores).
29. I am reliably informed that one federal district judge has replaced the portraits of district judges with professional portraits of a more demographically diverse group of citizens who recently underwent naturalization ceremonies at the courthouse.
30. *See* Thomas F. Pettigrew & Linda R. Tropp, *A Meta-Analytic Test and Reformulation of Intergroup Contact Theory*, J. PERSONALITY & SOC. PSYCHOL. (2006) (explicit measures). For examples regarding implicit measures, see, e.g., Natalie J. Shook & Russell H. Fazio, *Interracial Roommate Relationships: An Experimental Field Test of the Contact Hypothesis*, 19 PSYCHOL. SCI. 717 (2008); Nilanjana Dasgupta & Luis M. Rivera, *From Automatic Antigay Prejudice to Behavior: The Moderating Role of Conscious Beliefs About Gender and Behavioral Control*, 91 J. PERSONALITY & SOC. PSYCHOL. 268, 270 (2006).
31. *See* Nilanjana Dasgupta & Shaki Asgari, *Seeing Is Believing: Exposure to Counterstereotypic Women Leaders and Its Effect on the Malleability of Automatic Gender Stereotyping*, 40 J. EXPERIMENTAL SOC. PSYCHOL. 642, 649–54 (2004).

Dasgupta and her co-authors have produced two other studies with consistent findings. For example, they examined students who were randomly assigned to male or female professors for the same calculus course. Women students assigned to the female professor improved their implicit attitudes toward mathematics and how much they identified with mathematics as a discipline. Importantly, this difference persisted up to three months later.[32] In another study, women engineering students who were assigned randomly a female (versus male) senior engineering student mentor experienced changes in implicit associations, which persisted up to a year after mentoring had completed.[33]

Recall the nutritional adage "you are what you eat." Taking this statement seriously encourages more mindful eating—the what, when, why, and how we stuff food into our mouths. The same might be true with our minds. *You are what you see.* And if you proactively cultivate an environment that involves seeing and meeting people in their full complexity and diversity, these interactions may slowly alter the negative attitudes and stereotypes we hold. This is valuable, difficult, long-term work that all Americans should engage in, including judges. We would all be better served if we affirmatively cultivated colleagues, friendships, social relations, civic participation, and even media consumption[34] that expand our horizons and comfort levels with anyone marked as "other." Indeed, we could go beyond passive and casual social integration and seek out civic, community, and charitable projects that require us to cooperate actively, deeply, and repeatedly with fellow Americans that we would not otherwise interact with, except as objects within a hierarchical judicial system.

In addition, leverage your market power the next time you are invited to speak on a panel or keynote a conference, or are given an award, to see if organizers are lazily inviting and recognizing the usual suspects. This does not mean insisting rigidly that, for example, every single panel must have maximum demographic diversity. That's difficult to achieve and breeds tokenism. Instead, take a more gracious longitudinal view, and examine their long-term practices and trends. On that view, you may still have good grounds to nudge organizers to do better than reprogramming with the same-old-same-old. Even better, provide a referral. This both creates opportunity for the speaker who's featured and increases the audience's exposure to someone who varies from what's expected and thus functions as a "debiasing agent."[35]

All these strategies are long-term investments in life and country that will not show immediate or easily quantifiable returns. And we should recall the Kantian injunction to treat human beings as ends in and of themselves and not just the means for some self-improvement makeover project. But overall, I see much to admire in embracing such a life strategy, especially for those who have chosen the honorable profession of judge.

C. DEFEND (AGAINST THE BIAS THAT PERSISTS)

"I discourage ... obsessing over ... individual IAT scores."

When people brainstorm countermeasures to implicit bias, their natural inclination is to focus on debiasing. But I discourage people from obsessing over reducing their individual IAT scores. Far more valuable will be creating *defenses* against the implicit biases that will persist or soon return. Here's a (non-coronavirus) virus analogy. Suppose you have an irreplaceable computer (your brain). Suppose that it has been infected with a Trojan Horse virus (implicit bias), and none of the antivirus software packages work.[36] Even when the problem seems fixed, the infection returns within 24 hours. Maybe that's because the virus has burrowed deeply into the operating system or boot sector. Or maybe it's because surfing the Internet guarantees daily re-infection. Thankfully, the virus is not a game stopper; it doesn't crash your machine, steal your passwords, encrypt your storage and ask for ransom, or randomly transpose digits on budget spreadsheets. In fact, most users don't even realize that their machines are infected. But after careful study, you believe that this Trojan Horse virus influences your computer's work in small but consequential ways. Even if the virus cannot be removed, can you nevertheless *defend* against its impact? And might those defenses have the added benefit of countering other variants of bias, beyond just the implicit?

1. Carefully consider blinding, dimming, or temporary cloaking of social category information

One logical approach to consider is *blinding*. If we are entirely unaware of (and do not assume and cannot infer) the social category of a person, implicit biases regarding that category cannot directly impact our decision making. In this sense, even though the implicit bias persists, it can't easily be activated because we have been *blinded* to the triggering datum. This is the rationale behind blind grading of examinations, as well as orchestra auditions behind curtains.[37]

In judicial practice, there may be situations in which removing social category identity, for example, from paper files, may successfully defend against the activation of implicit bias. Some examples include tasks that judges might do as managers of a workplace, such as initial sorting of clerkship and employment applications. But as attractive a solution as blinding may seem, this tactic suffers numerous limitations.

First, the identity of the person of interest will often be known or assumed, for example, after an in-person or video interaction. We read identity off of faces and names. Removing that information will often be difficult, expensive, or impractical.

32. Jane G. Stout et al., *STEMming the Tide: Using Ingroup Experts to Inoculate Women's Self-Concept in Science, Technology, Engineering, and Mathematics (STEM)*, 100 J. Personality & Soc. Psychol. 255 (2011).
33. *See* Tara C. Dennehy & Nilanjana Dasgupta, *Female Peer Mentors Early in College Increase Women's Positive Academic Experiences and Retention in Engineering*, 114 Proc. Nat'l. Acad. Sci. 5964 (2017).
34. Jerry Kang, *Bits of Bias*, in Implicit Bias Across the Law 132-45 (Justin Levinson & Robert Smith, eds. 2012).
35. For legal analysis of role models and debiasing agents, see Jerry Kang & Mahzarin Banaji, *Fair Measures: A Behavioral Realist Revision of Affirmative Action*, 94 Calif. L. Rev. 1063, 1109-15 (2006).
36. I introduced this analogy in the first systematic exploration of implicit bias, including the Implicit Association Test, in the law reviews. *See* Kang, *supra* note 27.
37. *See* Claudia Goldin & Cecilia Rouse, *Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians*, 90 Am. Econ. Rev. 715, 717, 725 (2000) (explaining how the number of female new hires increased anywhere between 25 to 46% once auditions were conducted behind screens).

"Like most actors in the judicial system, judges are stressed, overworked, and starved for time."

Second, blinding may not be appropriate if social identity is partially relevant to the decision to be made. Consider, for example, some equitable decision on parole, punishment, or child custody. Part of that decision may require appreciation of a person's biography to gauge "distance traveled," trajectory, or cultural context. By deleting certain social category information, such as race, ethnicity, religion, or language spoken, we may be deleting specific streams of relevant information.[38]

Third, blinding risks "pass-through" discrimination. Let's revisit the orchestra audition. Suppose that a high school orchestra program gave male students preferential equipment, training, and encouragement over eight weeks, then conducted a blind audition for some first chair. A blind decision-making process at summer's end would simply pass through the gender-based tailwind enjoyed by men and headwind suffered by women. Worse, it could morally "launder" the prior discrimination because any male winner could proudly assert that he was chosen behind a curtain, entirely on the merits.

Notwithstanding these limitations, blinding still can be useful in the selection or judging process when identity should be entirely irrelevant. But because blinding may have unintended consequences, any implementation of this tactic ought to be carefully analyzed. In addition, consider the following variations to blinding, which I call "dimming" and "temporary cloaking."

Dimming. There are multiple ways in which we can know social category information, such as the race of someone about to be sentenced. We could see the race listed on some demographic form, we could infer it from the name, or we could see a picture (black-and-white, color, high-resolution, small- or large-size, etc.) There is troubling evidence that darkness of skin hue and the Afrocentricity of a defendant's facial features may drive severity in punishment.[39] Given this concern, one could reasonably decide that a presentencing report need not have a prominent photograph of the defendant on the first page.[40] Race information will be available throughout, and it may be a hassle to remove. Moreover, it may actually be relevant depending on the context. But it's hard to see any need to observe specific facial features. By declining to see them, we are not blinding ourselves to race per se, but we would be *dimming* the intensity of that information, including the potential impact of implicit stereotypes associated with Afrocentric features.

Temporary cloaking. Consider a two-stage process of temporary cloaking. In the first stage, blinding can remove social category information, for example, in the initial sort of clerkship applications. After making a tentative decision (e.g., to produce a rough shortlist), in the second stage, the cloak is lifted to check for other factors, such as possible pass-through discrimination and unintended consequences.[41] Of course, this second stage of analysis can raise hard questions about race and gender consciousness, the social construction of merit, and corrective justice—all of which require careful explication.[42]

2. Give yourself ample time, emotional calm, and mental energy

Like most actors in the judicial system, judges are stressed, overworked, and starved for time.[43] Unfortunately, there's general evidence that stress leads us to scan alternatives less systematically and completely.[44] Intense emotions, including happiness[45] and disgust,[46] are also linked to less systematic thinking. Finally, time pressures are correlated with less accurate decisions.[47]

The above findings are not specifically or uniquely connected to the problem of implicit bias. However, recent work by Jordan Axt and Calvin Lai demonstrates how accuracy can be increased by providing more time on two tasks connected with implicit measures of bias. One task was the First-Person Shooter Task (FPST) created by Joshua Correll,[48] which requires people to respond quickly and "shoot" if they see a gun and "not shoot" if they see something harmless held by either White or Black men in photorealistic settings. They found that "[m]ore time pressure meant more errors."[49] Because the distribution of errors was biased—favoring White lives (erring by not shooting Whites

38. *See* Devon W. Carbado & Cheryl I. Harris, *The New Racial Preferences*, 96 CALIF. L. REV. 1139 (2008).
39. *See* Mark W. Bennett & Victoria C. Plaut, *Looking Criminal and the Presumption of Dangerousness: Afrocentric Facial Features, Skin Tone, and Criminal Justice*, 51 U.C. DAVIS L. REV. 745, 773-85 (2018) (summarizing studies).
40. This is Judge Bennett's practice. *See id.* at 801.
41. *See, e.g.,* Annabelle Krause et al., *Anonymous Job Applications of Fresh PhD Economists*, 117 ECON. LETTERS 441, 443 (2012) (showing that women had a higher probability to receive an interview invitation on standard application processes, but that higher probability disappeared with anonymous applications).
42. For further discussion of the social and psychological construction merit, see Kang & Banaji, *supra* note 35, at 1081-82.
43. *See* L. Song Richardson, *Systemic Triage: Implicit Racial Bias in the Criminal Courtroom*, 126 YALE L. J. 862 (2017).
44. *See* Giora Keinan, *Decision-Making under Stress: Scanning of Alternatives under Controllable and Uncontrollable Threats*, 52 J. PERSONALITY & SOC. PSYCHOL. 639 (1987) ("psychological stress, in and of itself, has a significant effect on the manner in which the decision-makers scanned the alternatives available to them").
45. *See* Bernd Wittenbrink et al., *Spontaneous Prejudice in Context: Variability in Automatically Activated Attitudes*, 81 J. PERSONALITY & SOC. PSYCHOL. 815, 818–19 (2001).
46. *See* Nilanjana Dasgupta et al., *Fanning the Flames of Prejudice: The Influence of Specific Incidental Emotions on Implicit Prejudice*, 9 EMOTION 585 (2009).
47. Various studies with accountants show that decreased time and mental resources produce less thorough results. *See, e.g.,* Robert L Braun, *The Effective Time Pressure on Auditor Attention to Qualitative Aspects of Misstatements Indicative of Potential Fraudulent Financial Reporting*, 25 ACCT., ORG. & SOC'Y 243, 255 (2000) ("Lack of attention to qualitative aspects of misstatements indicative of potential fraudulent financial reporting may be a manifestation of a lack of professional skepticism. The data appear to indicate that those under time pressure may not have maintained a questioning mind and may not have critically examined audit evidence to the same extent as those under less time pressure.").
48. *See* Joshua Correll et al., *The Police Officer's Dilemma: Using Ethnicity to Disambiguate Potentially Threatening Individuals,* 83 J. PERSONALITY & SOC. PSYCHOL. 1314, 1315-17 (2002) (describing FPST) (available at < http://psych.colorado.edu/~jclab/FPST.html>).
49. Jordan R. Axt & Calvin K. Lai, *Reducing Discrimination: A Bias Versus Noise Perspective*, 117 J. PERSONALITY & SOC. PSYCHOL. 26, 34 (2019).

even when they held guns) and devaluing Black lives (erring by shooting Blacks even when they lacked guns)—the increase in total number of errors produced an increase in overall race-based discrimination.

The other task they tested was an academic version of the recently created Judgment Bias Task (JBT).[50] It requires participants to decide whether students should be admitted into an honor society. Each candidate's profile includes only bare-bones information: a photograph (to signal social category), science GPA (on a 4 point scale), humanities GPA (on a 4 point scale), recommendation quality (either "excellent" versus "good"), and interview score (on a 100 point scale). Half of the profiles were objectively better than the other half, and participants were instructed to admit about half of the students they reviewed. In addition, half of the pictures were clearly more attractive than the other half (thus testing for attractiveness bias).

In making selections, participants were told either to take all the time they needed or were forced to evaluate each profile for less than two seconds. Again, time pressure produced more errors. Because the distribution of errors was biased in favor of attractive people (erring by admitting unqualified attractive people) and disfavored unattractive people (erring by rejecting qualified unattractive people), the increase in total number of errors increased overall attractiveness-based discrimination. The general upshot, confirmed in these experiments, is that time matters.

3. Instruct yourself to deliberate carefully

To promote accuracy, we must have not only the *ability* but also the *willingness* to be careful. The prior suggestion focused on ability, supported by ample time and cognitive resources. What about willingness? One way to increase willingness is to give ourselves an instruction to slow down and take care. In another study, Axt and Lai had participants read the following general instruction to take care:

> Prior research suggest that people may do a better job on this task if they put in more time to deliberate and think over their decisions. As a result, it is important that you think hard and slow down when making your decisions.[51]

In contrast, another group heard the opposite instruction that said that it was "important that you go with your gut and make your decisions more quickly." The "be careful" group demonstrated higher accuracy than the "trust your gut" group on the academic JBT.

There are almost never one-size-fits-all recommendations. And in certain contexts, such as picking an ice cream flavor, "going with one's gut" might produce more accurate or more satisfying answers. That said, if the goal is to avoid social category bias, we should all be skeptical of our guts. We should be wary of intuitive responses and remind ourselves to deliberate and reason carefully.

4. Cabin discretion by using checklists and rubrics

"[W]e should be skeptical of our guts."

What I hate most about being a professor is grading exams. I don't mind giving feedback and lots of it; I just dread scoring exams and assigning grades. Over my decades in teaching, I've vacillated between the "gestalt" and "spreadsheet" methods of grading. On one extreme, I've just read the exam, jotted down some reactions on the margins, come to an overall reaction, and gave a gestalt grade. On the other extreme, I've created elaborate spreadsheets with a hundred entries grouped by issues and sub-issues, with weighting factors and bonus points for novel thinking or cogent writing. The raw scores are then converted into standard units (Z-scores), weighted, aggregated, and fit into a curve.

The gestalt is easy and enjoyable. It allows me freedom to credit originality and brilliance and to penalize catastrophic errors. But I worry about consistency. Would it matter if I were grading the same exam in the morning instead of evening, after a snack or a beer, after exercising or arguing with a family member? Would that "B" move up or down by half a grade, or more?

By contrast, the spreadsheet method feels like a mechanistic grind. It's as if I've given an essay exam but am now perversely trying to grade it as if it were multiple-choice. At times I'll fill out the spreadsheet and be surprised that some mediocre exam grazed enough of the issues to register a high total. Or an insightful and beautifully written exam dropped one important matter and therefore scores below average. In these moments, the spreadsheet method feels off. Still, it produces more consistent results.

I offer this digression about grading exams for two reasons. First, it highlights the pervasiveness of the problem that all experts face when making highly subjective decisions that rely on professional "judgment." Faculty, managers, judges all struggle with the basic choice between some version of the gestalt and spreadsheet methods. Second, it empathizes with judges who chafe at the idea of being forced to adopt some spreadsheet when they prefer the gestalt. I get it. No professional wants her expert judgment to be constrained by forms, checklists, rubrics, and algorithms especially if they are created by bureaucratic others.

Still, there's one crucial difference between exam grading and judging. In most of my classes, I have the luxury of grading blind. This is one of the rarefied environments in which blinding prevents implicit biases from activation, with few if any unintended consequences. Accordingly, I'm not worried about implicit bias influencing my grading even when I go gestalt. But you, as judges, generally do not have that option. Accordingly, I encourage you to find ways to move, at least incrementally, toward the spreadsheet model.

The justification is that checklists and rubrics help cabin discretion in ways that increase overall accuracy.[52] Much of that evidence was presented in the 2012 *Implicit Bias in the Courtroom* article, which discussed how phenomena such as "constructed criteria,"

50. *See* Jordan R. Axt et al., *The Judgment Bias Task: A Flexible Method for Assessing Individual Differences in Social Judgment Biases*, 76 J. Experimental Soc. Psychol. 337 (2018).

51. *Id.* at 38.

52. *See, e.g.,* Robert H. Ashton, *Effective Justification and a Mechanical Aid on Judgment Performance*, 52 Org. Behav. & Hum. Decision Processes 292 (1992).

"Recent work suggests, for example, the value of a specific 'countersteering' instruction."

"shifting standards," and "casuistry" lead decision makers to alter their decision criteria subtly and unconsciously, in real time, to justify an underlying intuition or preference. In other words, we often go with our gut, which often means preferring people we like (warmth)[53] or seem to be like us (ingroup favoritism), then rationalize a post hoc explanation to justify that decision.[54]

But when we constrain our decision making, by adopting some features of a spreadsheet-like approach, our decisions tend to be more accurate and consistent. This recommendation jibes with the structured interview literature,[55] which suggests that asking a similar set of validated questions across candidates makes it easier to conduct more accurate interpersonal comparisons. It's also consistent with the grading literature.[56] Frankly, it's consistent with "thinking like a lawyer," which features element-by-element analysis of a larger legal doctrine. It's reflected in the careful way that we design jury instructions on each claim or cause of action.

The responsibility for building the checklists, rubrics, and algorithms falls on judges themselves, working together and with relevant stakeholders toward consolidating best practices.[57] In designing these decision aids, we should take care not to bake in biases into the "spreadsheet" (think about federal sentencing guidelines treatment of powder versus crack cocaine) or formalistically pass through prior acts of discrimination.[58]

5. Give yourself specific countersteering instructions

In many cases, race (or some other salient social category) is not directly at issue. Nevertheless, race looms in the air. This presents the judge a choice. On the one hand, you could embrace colorblindness and reason that because race is not directly relevant, you shouldn't think about it. It could be a distraction, or worse activate racialized thinking when it's unnecessary. On the other hand, you could embrace race-consciousness. After all, from an implicit social cognition perspective, you can't really be colorblind.

Back in the 2012 *Implicit Bias in the Courtroom* article, my coauthors and I argued in favor of the race consciousness approach in the context of instructing jurors.[59] This recommendation relied on mock juror research that found that White jurors were less likely to be biased when they were specifically put "on guard" about the potential of racial bias when evaluating ambiguous facts regarding an interracial dispute.[60] I still stand by this recommendation and thoughtful commentators, such as Cynthia Lee, have elaborated further, in the context of instructing juries.[61]

I also recommend this approach for judges themselves, who are the focus of this article. Recent work suggests, for example, the value of a specific "countersteering" instruction. I call this a countersteering instruction for two reasons. First, it is more *particular* than the general injunction to "drive carefully." When you learned how to drive (especially if you lived in a snowy climate), you may recall learning to countersteer in response to a skid: if the rear of your car starts skidding left, turn the steering wheel to the left. If it skids right, then turn the steering wheel to the right. Second, for many drivers, the countersteering instruction is *counterintuitive*: If your car is drifting left, why wouldn't you steer towards the right? By rough analogy, if you're worried about noticing race (implicitly), why wouldn't you try extra hard to push it (explicitly) out of your mind? The answer is that explicitly noticing the potential for bias is the best way to counter it.

In the series of studies we've already discussed, Axt and Lai measured the impact of a very specific instruction to notice and avoid the attractiveness bias when selecting students for the honor society. Instead of just being told generically to "be careful," participants were more particularly instructed:

> In addition to differing on their qualifications, applicants will differ in physical attractiveness. Prior research suggests that decision makers are easier on more physically attractive applicants and tougher on less physically attractive applicants.

In the prior interventions, we saw that more time and the general instruction to "be careful" improved accuracy and decreased

53. *See* Erik J. Girvan, *Wise Restraints?: Learning Legal Rules, Not Standards, Reduces the Effects of Stereotypes and Legal Decision-Making*, 22 PSYCHOL., PUB. POL'Y & L. 31 (2016).
54. *See* Kang et al., *supra* note 2, at 1156-59, 1164-66.
55. *See, e.g.,* Julie M. McCarthy et al., A*re Highly Structured Job Interviews Resistant to Demographic Similarity Effects?*, 63 PERSONNEL PSYCHOL. 325 (2010); Julia Levashina et al., *The Structured Employment Interview: Narrative and Quantitative Review of the Research Literature*, 67 PERSONNEL PSYCHOL. 241, 274 (2014).
56. *See, e.g.,* David M. Quinn, *Experimental Evidence on Teachers Racial Bias in Student Evaluation: The Role of Grading Scales*, 42 EDUC. EVALUATION & POL'Y ANALYSIS 375 (2020); John M. Malouff & Einar Thorsteinsson, *Bias in Grading: A Meta-Analysis of Experimental Research Findings*, 60 AUSTL. J. EDUC. 245 (2016).
57. *See, e.g.,* Crystal Soderman Duarte & Alicia Summers, *A Three-Pronged Approach to Addressing Racial Disproportionality and Disparities in Child Welfare: The Santa Clara County Example of Leadership, Collaboration, and Data-Driven Decisions*, 30 CHILD & ADOLESCENT SOC. WORK J. 1, 14 (2013) (discussing implementation of the CCC bench card in the Santa Clara County dependency court); Bennett & Plaut, *supra* note 39, at 801 (describing sentencing range algorithm); National Council of Juvenile and Family Court Judges, Addressing Bias in Delinquency and Child Welfare Systems (bench card), available at <https://www.ncjfcj.org/publications/addressing-bias-in-delinquency-and-child-welfare-systems>.
58. As you may know, these are the same challenges facing Artificial Intelligence (AI) systems that suffer from what computer scientists call a "garbage in-garbage out" problem.
59. *See* Kang et al., *supra* note 2, *at* 1184.
60. *See, e.g.*, Samuel R. Sommers & Phoebe C. Ellsworth, *"Race Salience" in Juror Decision-Making: Misconceptions, Clarifications, and Unanswered Questions,* 27 BEHAV. SCI. & L. 599 (2009).
61. Cynthia Lee, *Making Race Salient: Trayvon Martin and Implicit Bias in a Not Yet Post-Racial Society*, 91 N.C. L. REV. 1555, 1597-1600 (2013). In one study, which examined whether a specific implicit bias jury instruction might mitigate against racial bias, the researchers could not replicate the racial bias under the control condition that had previously been found by Sommers and Ellsworth. *See* Jennifer K. Elek & Paula Hannaford-Agor, *Implicit Bias and the American Juror*, 51 CT. REV. 116, 120 (2015).

the overall number of errors. Interestingly, that's not the effect that this countersteering instruction had. It didn't decrease the total number of errors—in other words, the same total number of unqualified students were elected and the same total number of qualified students were rejected. But it did change the biased distribution of those errors such that attractive and unattractive candidates were now equally likely to receive leniency (admitted to the honor society when they were unqualified) and harshness (rejected from the honor society when they were qualified). By removing the bias in the distribution of errors, this instruction decreased the total amount of discrimination suffered by the disfavored group even though the absolute number of errors remained constant.

In sum, it appears that both the general "be careful" instruction (Part IV.C.3) and the more specific countersteering instruction (do not try to suppress and instead notice and respond to a particular bias) reduce discrimination but through different causal pathways. The former reduces the *absolute number* of errors, whereas the latter changes the unfair *distribution* of those errors.

Here's one specific application of a countersteering instruction especially useful for judges (and your staff). As judges, you are constantly interacting with members of the community, who are nervous at being in the courthouse. For example, it is a site filled with with what Rachel Godsil has extensively elaborated as "racial anxiety."[62] On their side, this anxiety is likely to manifest in awkward body language, which can come off as nervousness, unresponsiveness, unfriendliness, untrustworthiness, and even hostility. To make matters worse, on your side, implicit biases alter the way we read nonverbal behavior. For example, it may take longer for us to recognize a smile on a Black face compared to a White one, even though the smiles are identical.[63] Numerous field studies in medicine have found that implicit bias predicts awkwardness in doctor-patient communication patterns,[64] and it's not a stretch to think the same might happen with judges interacting with parties or witnesses.

Accordingly, give yourself a very specific countersteering instruction on friendliness. Whenever you interact with someone who belongs to some outgroup (someone who is not of your race, ethnicity, religion, sexual orientation, eliteness of educational credentials, etc.) or group with marginalized status (non-native speaker, immigrant, lower socioeconomic status, etc.), make sure to countersteer and err on the side of warmth, respect, and welcome. Doing so can trigger recursive benefits.[65] Your hospitality may decrease environmental threat, which may relax their behavior, which may alter their body language in a way that you and your staff will respond to positively, which can further decrease threat, and so on in a virtuous cycle.

"To make matters worse, ... implicit biases alter the way we read nonverbal behavior."

6. Engage in perspective shifting and category switching

In the 2012 *Implicit Bias in the Courtroom* paper, we encouraged judges to recommend to jurors that they engage in perspective-taking.[66] Perspective-taking roughly means putting oneself in the shoes of another. We pointed out that actively contemplating the feelings and experiences of others, especially outgroups, could weaken automatic expression of bias, including implicit bias measured by the IAT.[67] Since that time, slightly greater evidence has accumulated in favor of perspective-taking.[68]

For example, certain studies have demonstrated that perspective-taking improved implicit measures of bias regarding various social groups, such as Turks, elderly,[69] and Asians.[70] Unfortunately, the evidence is mixed with some researchers finding no changes in implicit bias, at least as measured by the IAT, from one of the perspective-taking interventions.[71] We find ourselves again in a position with imperfect scientific knowledge. But this is an opportune moment to remind ourselves that the goal is not to reduce IAT scores *per se*. Instead, we should keep our eyes on the prize, which is to decrease discriminatory behavior. And if perspective-taking might incrementally nudge us toward that goal, we should pursue it regardless of whether our implicit bias scores change.

Perspective-taking interventions have correlated with changes in behavior, including subtle choices such as seating distance and helping behaviors (such as helping to pick up dropped keys). In the medical context, perspective-taking has decreased the racial gap in empathizing with the pain experienced by White and Black patients.[72] Based on such evidence, I

62. For a discussion of the concept, see Rachel D. Godsil & L. Song Richardson, *Racial Anxiety*, 102 Iowa L. Rev. 2235, 2239 (2017) (identifying "concerns that often arise both before and during interracial interactions" even when the interacting parties seek a positive experience); Rachel D. Godsil et al., The Science of Equality, Volume 1: Addressing Implicit Bias, Racial Anxiety, and Stereotype Threat in Education and Health Care (Perception Institute 2014).
63. *See* Kurt Hugenberg & Galen V. Bodenhausen, *Facing Prejudice: Implicit Prejudice and the Perception of Facial Threat*, 14 Psychol. Sci. 640 (2003).
64. *See* Ivy W. Maina et al., *A Decade of Studying Implicit Racial/Ethnic Bias in Healthcare Providers Using the Implicit Association Test,* 199 Soc. Sci. & Med. 219, 223 (2017).
65. For discussion of how small interventions can produce substantial changes through recursive phenomena, see Gregory M. Walton & Timothy D Wilson, *Wise Interventions: Psychological Remedies for Social and Personal Problems*, 125 Psychol. Rev. 617 (2018).
66. *See* Kang et al., *supra* note 2, at 1185-86.
67. *See, e.g.,* Andrew R. Todd et al., *Perspective Taking Combats Automatic Expressions of Racial Bias*, 100 J. Personality & Soc. Psychol. 1027, 1031-33 (2011).
68. For a useful review of perspective taking, see Andrew R. Todd & Adam D. Galinsky, *Perspective-Taking as a Strategy for Improving Intergroup Relations: Evidence, Mechanisms, and Qualifications*, 8 Soc. & Personality Psychol. Compass 374 (2014).
69. *See* Andrew R. Todd & Pascal Burgmer, *Perspective Taking and Automatic Intergroup Evaluation Change: Testing An Associative Self-Anchoring Account*, 104 J. Personality & Soc. Psychol. 786 (2013).
70. *See* Margaret J. Shih et al., *Perspective-Taking and Empathy: Generalizing the Reduction of Group Bias toward Asian Americans to General Outgroups*, 4 J. Abnormal Psychol. 79 (2013).
71. This is what Lai found in his tournament approach. See Lai et al., *supra* note 24, at 1770.
72. *See* Adam T. Hirsh et al., *A Randomized Controlled Trial Testing a Virtual Perspective-Taking Intervention to Reduce Race and SES Disparities in Pain Care*, 160 Pain 2229 (2019); Brian B. Drwecki et al., *Reducing Racial Disparities in Pain Treatment: The Role of Empathy and Perspective-Taking,* 152 Pain 1001 (2011).

"[T]ry to stand still in that perspective, and see if your judgment moves at all."

recommend that judges experiment with *perspective-taking*. More specifically, before exercising discretion or making a judgment call (e.g., granting a motion to dismiss or a motion for summary judgment on an employment discrimination claim) against an outgroup member or target of implicit bias, put yourself briefly in the shoes of a member of that group.[73] While doing so, try to resist any immediate impulse to say something like "I would have never done that!" Instead, try to stand still in that perspective, and see if your judgment moves at all. In addition, I encourage you to experiment with the tactic of *counterfactual category switching*. For example, if you are about to depart upward from sentencing guidelines, ask yourself whether you would do the same if the defendant were of a different race or member of your ingroup.

7. Prefer diverse decision-making teams

There is a rich literature examining whether diverse teams—according to various metrics—deliberate differently and produce better answers.[74] In some cases, they clearly do deliberate differently, often by canvassing a larger solution space. And in some cases, they clearly do generate better answers. But here, I focus narrowly on how the diversity of teams might counter implicit bias.

One way to mitigate a headwind is to combine it with a tailwind. So, if most members of a decision-making body lean implicitly in one direction, it could be useful to have another member of that body who leans implicitly in another direction. The goal cannot be anything like precise calibration so that the vector sum of all possible implicit associations equals zero. That is infeasible. That said, it's reasonable to assume that a more heterogeneous group is likely to have a more heterogeneous set of implicit (and explicit) biases, with the inevitable result of some members' biases dampening out the impact of others'.

One final way that diversity could help counter implicit bias is that the very existence of a member of another social category can function as a physical reminder to be mindful about how to think and talk about that category.[75] This may be most important in constructing a diverse jury,[76] which is outside the scope of this article. But even if we stay focused on judges themselves, we know that judges form and participate in various panels, teams, committees, and task forces. As they do so, they should be mindful of the kinds of diversity that might decrease the vector sum of implicit biases within the group. We should, as always, not be overconfident given the possibility that "token" representation could produce moral credentialing, and an unwarranted confidence that the group itself couldn't possibly be biased, which would then backfire.[77]

D. DATA (TO CREATE EARLY WARNING SYSTEMS)

Scientific advancements allow us to see what was previously invisible, from the microscopic to the galactic. Arguably that's what instruments such as the IAT give us, a blurry window into an otherwise opaque mental domain. Collecting and visualizing data often allow us to do the same. As individual judges of goodwill exercise their daily discretion, it will often be impossible to spot in any specific case whether an implicit or other variant of bias played a causal role. However, if similar decisions are logged across time and/or multiple decision makers, the data may reveal interesting patterns.

For instance, would it surprise you to find out that regional IAT scores (which average over a large population of people, and thus wash out the noise in individual measurement errors) correlate with regional differences in racially disproportionate lethal force[78] and school discipline?[79] Of course, correlation does not mean causation. As such, the data often cannot definitively answer whether discrimination is taking place. But they do plant red flags and identify areas of concern that warrant deeper examination.

Judges should initiate data collection on decisions that involve substantial discretion. At the individual level, it could involve ordinary human resources processes within your chambers such as hiring law clerks and staff. Or it can involve your individual

73. *See* Stefanie Simon et al., *Pick Your Perspective: Racial Group Membership and Judgments of Intent, Harm, and Discrimination*, 22 GROUP PROCESSES & INTERGROUP REL. 215, 229 (2019) (showing that perspective-taking alters assessments of intent and harm for White participants).
74. *See, e.g.,* SCOTT E. PAGE, THE DIFFERENCE: HOW THE POWER OF DIVERSITY CREATES BETTER GROUPS, FIRMS, SCHOOLS, AND SOCIETIES (2007).
75. *See* Kang et al., *supra* note 2, at 1180; Samuel R. Sommers, *On Racial Diversity and Group Decision Making: Identifying Multiple Effects of Racial Composition on Jury Deliberations*, 90 J. PERSONALITY & SOC. PSYCHOL. 597 (2006).
76. Janet Bond Arterton, *Unconscious Bias and the Impartial Jury*, 40 CONN. L. REV. 1023, 1033 (2008) (quoting letter from anonymous juror).
77. *See* Manuel Bagues et al., *Does the Gender Composition of Scientific Committees Matter?,* 107 AM. ECON. REV. 1207, 1227 (2017) (finding that "increasing the proportion of women and scientific committees does not increase the success rate of female candidates" in Italian and Spanish promotion decisions to full professorships partly because female evaluators do not vote more in favor of female candidates in a statistically significant manner and "the presence of women in the committee decreases the probability that female candidates receive a positive vote from male evaluators").
78. Eric Hehman et al., *Disproportionate Use of Lethal Force in Policing Is Associated with Regional Racial Biases of Residents*, 9 SOC. PSYCHOL. & PERSONALITY SCI. 393, 397 (2018) (finding that "the implicit racial biases [both attitudes and weapon stereotypes] of White residents predict disproportionate regional use of lethal force with Blacks by police. This association is robust, reliably emerging across two conceptually distinct measures of racial bias, multiple imputations, three different transformations of the outcome measure, traditional and bootstrapped distributions, and above and beyond 14 sociodemographic covariates."). By contrast, explicit measures had no statistically significant effect. *Id.* at 396.
79. *See* Travis Riddle & Stacy Sinclair, *Racial Disparities in School-based Disciplinary Actions Are Associated with County-Level Rates of Racial Bias*, 116 PROC. NAT.'L ACAD. SCI. 8255 (2019). They found that explicit bias scores were more predictive but also found that implicit bias and disciplinary disparities were correlated. They checked for confounds that typically occur, including socioeconomic status and population demographics. *See id* at 8258.

patterns in exercising judicial power. For instance, on federal sentencing matters, it would not be difficult to keep a running record of the computed "guideline range," your final sentencing recommendation within that range, and the key social category variables of the defendant (e.g., race and gender). By computing averages and standard deviations, you could easily alert yourself to disparities that warrant a closer examination.

At the institutional level, judges could call for broader counting of the exercise of sovereign power in areas such as prosecutorial charging decisions,[80] plea bargains, setting bail,[81] sentencing recommendations made by probation officers, and sentencing.[82] Anywhere judges believe that implicit bias might be infecting the decision-making process is a good place to start counting.

The first cut of the data would examine whether the exercise of discretion seems correlated with salient demographic categories, such as race. The second cut would examine whether that relationship persists after controlling for confounding factors. If the data reveal, for example, racial disparities that cannot easily be explained by other relevant factors, then we should plant a red flag. If these disparities appear at the institutional level, judges should call for the convening of (diverse) task forces to analyze their causes and examine whether checklists, rubrics, and other algorithmic guardrails might improve accuracy and decrease biased results.

Another benefit of data collection is that it generates soft accountability pressures. If you are accountable to explain and justify publicly your decisions, for example, in a published opinion with precedential value, you will make them more carefully and more accurately. Similarly, if you know that your exercise of discretion, which historically has been invisible, will now suddenly become more visible through individual and institutional counting practices, you will start taking greater care.

Supporting evidence comes from economists studying referees and judges in professional sports. For example, large data analyses found referees and umpires making calls in a race-based way, under certain conditions. Interestingly, these racially biased decisions stopped when the judges were subject to greater scrutiny, either in the form of video data collection (through Questec cameras installed in ballparks that measured human umpire accuracy in calling balls and strikes)[83] or increased media coverage after news outlets such as ESPN popularized the research findings.[84]

CONCLUSION

"Judges ... can organize their response according to the four 'D's': deflate, debias, defend, and data."

Over the past twenty years, we have come to accept the *idea* of implicit bias. It no longer seems odd to believe that we have attitudes and stereotypes that we largely lack access to. Scientists continue to innovate and improve the *instruments* that can measure this idea. The most popular instrument remains the Implicit Association Test (IAT). As good as it is, frankly, it's just a videogame, and we should not be shocked that it lacks the measurement precision necessary for responsible individual diagnostics.[85] Nevertheless, it speaks volumes about society.

Implicit biases about social categories are pervasive, stronger than explicit biases, and show low-level correlation with discriminatory behavior. The correlations are small, partly due to the difficulties in getting precise measurements of either bias or behavior. Nevertheless, when we aggregate these effects over time and across entire populations, implicit bias can produce tailwinds and headwinds that profoundly perturb our commitment to giving everyone a fair shot and equal justice under law.

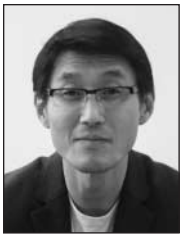
So, what can judges do? Unfortunately, there is no silver bullet or panacea, and the scientific evidence remains sometimes frustratingly limited. In this article, which is already too long and complex, I addressed only the problem of implicit bias held by judges themselves. To be explicit, I did not directly discuss how judges might confront the implicit biases of jurors or other players within the judicial system, such as prosecutors or lawyers. I also have also not repeated my call for a "behavioral realism" in legal doctrine and jurisprudence since I've discussed those matters extensively elsewhere.[86]

Judges who believe that implicit bias is a genuine problem can organize their response according to the four "D's": deflate, debias, defend, and data. Specific and concrete tactics under these strategies appear in the Appendix. I confess that it's hard to know whether the strategies will have great impact. And implementation will take hard, persistent work, driven by your internal motivation to be fair, not only as individuals but also as parts of a larger system of justice. Still, I have curated these evidence-based recommendations not to be especially costly, impractical, or objectionable. In addition, they are unlikely to backfire or produce ironic consequences that make matters worse. Finally, many

80. *See* Kang et al., *supra* note 2, at 1140 (collecting evidence of disparities).
81. See Ian Ayres & Joel Waldfogel, *A Market Test for Race Discrimination in Bail Setting*, 46 STAN. L. REV. 987, 992 (1994) (finding 35 percent higher bail amounts for Black defendants after controlling for eleven other variables).
82. *See* Irene V. Blair et al., *The Influence of Afrocentric Facial Features in Criminal Sentencing*, 15 PSYCHOL. SCI. 674, 675 (2004).
83. *See* Christopher A. Parsons et al., *Strike Three: Discrimination, Incentives, and Evaluation*, 101 AM. ECON. REV. 1410, 1433 (2011).
84. *See* Joseph Price & Justin Wolfers, *Racial Discrimination Among NBA Referees*, 125 Q. J. ECON. 1859, 1885 (2010).
85. *See, e.g.,* Kang et al., *supra* note 2, at 1179. For an updated discussion of the measurement precision of various implicit bias instruments, including the IAT, see Anthony G. Greenwald & Calvin K. Lai, *Implicit Social Cognition*, 71 ANN. REV. PSYCHOL. 419, 425-26 (2020) (elaborating the relationships between internal consistency and test-retest reliability).
86. *See, e.g.,* Jerry Kang, *Rethinking Intent and Impact: Some Behavioral Realism about Equal Protection*, 66 ALABAMA L. REV. 627 -51 (2015) (Meador Endowed Lecture); Jerry Kang, *The Missing Quadrants of Anti-discrimination: Going Beyond the "Prejudice Polygraph,"* 68 J. SOC. ISSUES 314-27 (2012); Jerry Kang & Kristin Lane, *Seeing through Colorblindness: Implicit Bias and the Law*, 58 UCLA L. REV. 465–520 (2010).

of these recommendations will improve decision making regardless of the precise variant of bias.

Much work remains to be done. At the individual level, it will require judges to work methodically and consistently toward deeper scientific understanding and personal introspection, improved habits, and increased experimentation with procedures and practices. At the institutional level, it will require convening judges, legal scholars, and social scientists to sit together on blue-ribbon committees with the charge, resources, and access to data to generate scientifically sophisticated and evidence-based guidance. It will be hard work.



*Jerry Kang is Distinguished Professor of Law and (by courtesy) Asian American Studies at UCLA. He was also the inaugural Korea Times –*Hankook Ilbo *Endowed Chair for Law and Korean American Studies (2010-20) and the university's Founding Vice Chancellor for Equity, Diversity and Inclusion. A leading scholar on implicit bias and critical race studies, Professor Kang collaborates broadly across disciplines and industries on scholarly, educational, and advocacy projects.*

APPENDIX
24 THINGS JUDGES CAN DO ABOUT IMPLICIT BIAS

I. DEFLATE (YOUR EGO) AND EMBRACE FALLIBILITY

1. Recognize that **you are fallible**.
2. **Avoid "moral credentialing"** simply because you have studied implicit bias.
3. Don't fret over *external* motivations for political correctness. Instead, **cultivate your *internal* motivation** to be fair.
4. Continue to **learn more** about all kinds of biases and decision-making errors [87] not because education directly decreases those errors but because deeper awareness will support your internal motivation to improve continuously both individually and institutionally.[88]

II. DEBIAS (WITH SHORT-TERM "SPOT CLEANING" AND LONG-TERM INTERACTIONS)

A. SHORT-TERM TACTICS

5. **Change the built environment** (e.g., photographs, art, posters, statues, books) to include regular, consistent exposure to admired figures from diverse groups and countertypical exemplars ("debiasing agents").

B. LONG-TERM TACTICS

6. **Expand social contact** with other, less familiar social groups directly and vicariously.[89] In so doing, always **curate complexity, not caricature**.
7. **Leverage your market power** to nudge others to be mindful of whom they feature as speakers or experts because "we are what we see."

III. DEFEND (AGAINST THE BIAS THAT PERSISTS)

A. CAREFULLY CONSIDER BLINDING, DIMMING, OR TEMPORARY CLOAKING SOCIAL CATEGORY INFORMATION

8. **Consider** whether ***blinding*** may improve fairness and not simply pass through prior acts of discrimination by the judicial system and others.
9. **Consider *dimming*** by decreasing the intensity, salience, or completeness of social category information. For example, you can keep the race field in documents but remove the photograph.
10. **Consider** using the two-stage process of ***temporary cloaking*** to first cloak identity and make a tentative decision, then uncloak to check for unintended consequences.

B. GIVE YOURSELF AMPLE TIME, EMOTIONAL CALM, AND MENTAL ENERGY

11. **Give** yourself **ample time** to improve accuracy in making complex, subjective, multifaceted decisions.
12. If you are in an especially high or low emotional state or feel especially stressed or cognitively depleted, try to delay making complex, subjective, multifaceted decisions until you **return closer to your baseline**.

C. REMIND YOURSELF TO DELIBERATE CAREFULLY

13. **Remind yourself to be careful** instead of jumping to conclusions or relying on intuitions or gut feelings.

87. *See, e.g.,* Pamela Casey et al., Minding the Court: Enhancing the Decision-Making Process (American Judges Association 2012) (white paper); Chris Guthrie, Jeffrey J. Rachlinski, & Andrew J. Wistrich, *Inside the Judicial Mind*, 86 CORNELL L. REV. 777 (2001).
88. On implicit bias, here are some resources I periodically update: http://jerrykang.net/2011/03/13/getting-up-to-speed-on-implicit-bias/. For evidence that education can drive awareness and internal motivation, see Patrick Forscher et al., *Breaking the Prejudice Habit: Mechanisms, Time Course, and Longevity*, 72 J. EXPERIMENTAL SOC. PSYCHOL. 133 (2017) (showing that intervention produced changes in knowledge and belief about race-related issues, which correlated with behavior measured years later); Molly Carnes et al., *Effect of an Intervention to Break the Gender Bias Habit for Faculty at One Institution: A Cluster Randomized, Controlled Trial*, 90 ACAD. MED. 221 (2015) (finding changes in self-efficacy, self-reported action to promote gender equity).
89. *See* Jerry Kang, *Cyber-Race,* 113 HARV. L. REV. 1130, 1166–67 (2000) (comparing vicarious with direct experiences).

D. CABIN DISCRETION BY USING CHECKLISTS AND RUBRICS

14. For important subjective decisions you regularly make, **use a checklist, rubric, or algorithm** to help guide your decision making.

15. If no such checklist or rubric exists, work with your institution to **develop collaboratively** the substantive content of such **decision aids**. In doing so, watch out for algorithms that bake in biases or pass through prior acts of discrimination.

E. GIVE YOURSELF SPECIFIC COUNTERSTEERING INSTRUCTIONS

16. Identify the social categories[90] that might trigger either explicit or implicit bias in any interaction or matter. Consciously **surface such dangers instead of trying to suppress** them.

17. Go beyond a general reminder to take care and **give yourself specific countersteering instructions** about the specific relevant bias.[91]

18. Because implicit bias can influence how we read body language, countersteer and explicitly **signal respect and warmth to people who are members of an outgroup** or have marginalized status.

F. ENGAGE IN PERSPECTIVE SHIFTING AND CATEGORY SWITCHING

19. Before exercising discretion or making a judgment call against an outgroup member or a target of implicit bias, briefly **put yourself in the shoes of that person**.

20. Also, **try switching categories counterfactually** (for example, from Black to White) and see if your judgment changes.

G. PREFER DIVERSE DECISION-MAKING TEAMS

21. **Assemble diverse** decision-making **teams** so that one member's implicit biases might dampen out another's.[92]

IV. DATA (TO CREATE EARLY WARNING SYSTEMS)

22. **Count your** own exercises of **discretion**, such as the clerks you've hired or the families you've had over for dinner. See if the data reveal a pattern that concerns you.

23. **Encourage institutional-level counting** of highly discretionary decision making to look for disparities by race and other social categories that cannot be explained away by confounding variables. Plant a red flag wherever you spot troubling disparities.

24. Call for teams and task forces to **examine those red flags** and study whether any of the tactics above, including cabining discretion, might improve accuracy and decrease bias.

90. Recent research suggests that a countersteering instruction targeted at one bias (e.g., attractiveness) may have no impact on another bias (e.g., ingroup favoritism for one's own university). So, specificity will be important. Jordan R. Axt et al., *Reducing Social Judgment Biases May Require Identifying the Potential Source of Bias*, 45 PERSONALITY & SOC. PSYCHOL. BULL. 1232 (2019).

91. An instruction for juries would look something like this: "Research suggests that decision makers generally have an implicit attitude in favor of [X over Y], or an implicit stereotype that associates [X with Z]. These implicit biases may operate without our conscious awareness and change how we react to people and their stories. Our jobs are to treat everyone fairly and not allow that to happen."

92. As applied to juries (which again goes beyond the scope of this article), I recommend curtailing selection procedures that undermine achieving such diversity, such as peremptory challenges. *See* Kang et al., *supra* note 2, at 1181.

Answers to Crossword

from page 127