

# First, Do No Harm:

## On Addressing the Problem of Implicit Bias in Juror Decision Making

By Jennifer K. Elek & Paula Hannaford-Agor

Over the past three decades, court leaders across the country have taken aggressive steps to confront racial bias in the courts. Recent efforts include in-depth judicial education and training about how an individual's unconscious attitudes (including culturally learned associations or generalizations that we tend to think of as stereotypes) introduce unjustified assumptions about other people and related evidence that can distort a person's judgment and behavior. This phenomenon is now referred to as *implicit bias* to differentiate it from explicit or intentional bias. Judicial-education programs focus on raising judicial awareness about implicit bias and introducing techniques that judges may use to help minimize the impact of implicit bias on judicial decision making.

Despite high levels of interest and genuine commitment to racial fairness in the justice system, disparate treatment of racial minorities persists and pervades all stages of the criminal justice process. Jury trials are a particularly troubling component of the justice system with regard to the potential for racial bias. Courts have extremely limited opportunities to educate jurors about the pernicious effects of complex psychological phenomena like implicit bias and how these implicit forms of bias may distort jurors' interpretation of trial evidence. Jurors are randomly selected from the local community. Other than statutory qualifications such as U.S. citizenship, age (adults 18 or older), and the ability to speak and understand English, state courts have no educational, occupational, or personal experience requirements to be eligible for jury service. Most jurors in this country serve only for the duration of the trial (typically two to three days) and then are released from service. No time is available during this short period to provide the type of in-depth education on implicit bias that judges and court staff may receive. Instead, judges and lawyers are increasingly looking to existing opportunities within the jury-selection and trial period (e.g., juror orientation, voir dire, jury instructions) in which to inform jurors about the propensity of implicit bias to affect decision making and to provide concrete strategies to minimize the impact of implicit bias on jury verdicts.

This article focuses on several of these interventions and the factors that may increase or undermine their effectiveness. Most Americans are aware of the existence of explicit bias and its effects on decision making generally, but implicit bias is still a relatively new concept about which many people in the justice system are unaware. The first section of this article discusses the difference between explicit bias and implicit bias and why

contemporary researchers have become more convinced that much of the disparity in legal outcomes for African-Americans compared to whites is likely due to implicit bias. We then describe different interventions that have been proposed to reduce the impact of implicit bias, and findings from empirical research about their effectiveness. One complication of these interventions is that some otherwise well-intentioned approaches can provoke a backlash effect in which the individuals exposed to the intervention are actually more likely to make judgments or behave in ways that manifest prejudice. In the context of administering these interventions with trial jurors, there are a number of pros and cons, many of which involve purely logistical concerns. We conclude with an update about interventions that are currently being tried, including a pilot test of an implicit-bias jury instruction.

### THE IMPACT OF IMPLICIT BIAS IN THE JUSTICE SYSTEM

Judges, lawyers, and court staff have long recognized that explicit, or consciously endorsed, racial prejudices have no place in the American justice system. The Code of Judicial Conduct in most states expressly prohibits judges from engaging in bias, prejudice, or harassment on the basis of race, gender, ethnicity, or other factors, and the code even extends the prohibition to court employees over which the judges have control and to lawyers appearing in cases before them.<sup>1</sup> In fact, most judicial-performance evaluations include measures of judicial impartiality as a major focus. The underlying justification for this prohibition is that discriminatory speech or behavior undermines public perceptions of judicial impartiality and competence. In contemporary society, most people recognize that explicit racial bias is normatively bad, and they make efforts to suppress biased behaviors or speech, even if they consciously recognize that they have those attitudes.

What often surprises members of the court community and other professionals is that more subtle biases or prejudices can operate automatically, without awareness, intent, or conscious control. Personal attitudes and acquired knowledge often help individuals function more efficiently by making it easier for the brain to recognize and respond quickly to new people or situations. But some attitudes, especially racial and cultural stereotypes, distort decision making by unfairly influencing judgments about others. Although explicit or consciously endorsed racial prejudices in contemporary American society may be on the decline, this subtler form of implicit racial bias persists.

Over the past few decades, a number of specialized tests have

### Footnotes

1. ABA MODEL CODE OF JUD. CONDUCT R. 2.3 (2011). Twenty-seven states have adopted the language of Rule 2.3 or substantially similar language.

been developed to help researchers identify, measure, and study implicit forms of bias. One of the most popular is the Implicit Association Test (IAT), developed by researchers in the mid-1990s at Yale University and the University of Washington. The IAT operates under a basic premise of human cognition that when an idea is consistent with a person's attitudes or cultural understanding, he or she will be able to mentally associate concepts related to that idea more quickly and easily than if the idea is inconsistent with a person's attitudes or cultural understanding. In an early version of the IAT, for example, researchers measured the amount of time it took people to associate pictures or words representing flowers or insects with positive attributes ("good") and with negative attributes ("bad") by hitting right or left keys on a computer keyboard. Because flowers are generally viewed as intrinsically good and insects as intrinsically bad (or at least significantly less good), most were able to hit the keys associating flowers and words indicating positive attributes, and insects and words indicating negative attributes, much faster and with fewer errors than when they asked to associate flowers with words indicating negative attributes or insects and words indicating positive attributes.<sup>2</sup> The difference in the amount of time and the number of errors reflects the strength of the person's preference for flowers over insects. Interestingly, young boys and entomologists tended to show weaker preferences for flowers over insects compared to young girls and people who do not study insects professionally.<sup>3</sup> This pattern of stronger preferences for more culturally familiar and socially and individually learned concepts has been found to be consistent for IAT tests measuring implicit biases based on race, gender, ethnicity, sexual orientation, age, religion, disability, body weight, and other characteristics developed and employed over the past 15 years. To try an Implicit Association Test, go to [www.implicit.harvard.edu](http://www.implicit.harvard.edu).

A large body of empirical literature now documents the existence of implicit biases and their behavioral implications. One recent meta-analysis of 122 research reports found implicit biases to be valuable, independent predictors of social behavior and judgment.<sup>4</sup> Implicit racial bias is the most studied type of implicit bias. Research shows that implicit racial bias can predict the quality of social interactions and decision-making outcomes in a variety of contexts, including voting, hiring, performance assessment, budget setting, policing, and medical treatment.<sup>5</sup> In the context of the American justice system, researchers now point to linkages between implicit racial bias

and disparities in detention decisions, jury verdicts, capital punishment, and other sentencing outcomes.<sup>6</sup>

Research on judicial decision making suggests that judges are affected by implicit bias in ways similar to the general population. In one study of actual trial judges in three jurisdictions, white judges showed strong implicit attitudes favoring whites over blacks (consistent with studies of implicit bias in the general population).<sup>7</sup> The judges were presented with three vignettes, two of which did not identify the defendant's race but some of which included words designed to trigger an association with African-Americans (e.g., Harlem, dreadlocks). The third vignette explicitly identified the defendant as white or black. The judges were asked to recommend a judgment on guilt, to share their confidence in that judgment, and to predict the defendant's likelihood of future recidivism. Interestingly, judges did not differ in their judgments in the first two vignettes based solely on whether the vignette included cues regarding race, but the judges' Race IAT was a marginally significant predictor of the harshness of the sentence. Judges whose Race IAT indicated a preference for whites over blacks were more likely to convict the defendant, had greater confidence in that judgment, and believed the likelihood of recidivism to be higher than judges whose Race IAT indicated a preference for blacks over whites. In the third vignette, judges with greater implicit bias against blacks convicted black defendants at the same rate as white defendants, but judges with greater implicit bias in favor of blacks convicted black defendants less frequently than they did white defendants.

Studies of juror decision making also demonstrate the impact of implicit bias on judgments. Levinson and Young, for example, conducted a mock-jury experiment in which mock jurors studied 20 pieces of trial evidence including photographs of a crime scene, one of which was a surveillance camera photograph featuring a masked gunman whose hand and forearm were visible. Half the cases showed light skin on the gunman and half showed dark skin.<sup>8</sup> This subtle manipulation of skin color produced significantly different results in jurors' confidence in their verdict. On a scale of 1 (not at all guilty) to 100 (absolutely guilty), jurors who viewed the dark-skinned gun-

**Studies of juror decision making also demonstrate the impact of implicit bias on judgments.**

2. Kristin A. Lane, Mahzarin R. Banaji, Brian A. Nosek & Anthony G. Greenwald, *Understanding and Using the Implicit Association Test: IV. What We Know (So Far) About the Method*, in *IMPLICIT MEASURES OF ATTITUDES* (Bernd Wittenbrink & Norbert Schwarz eds., 2007).

3. *Id.*

4. Anthony Greenwald et al., *Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity*, 97 *J. PERSONALITY & SOC. PSYCHOL.* 17 (2009).

5. For reviews, see John T. Jost et al., *The Existence of Implicit Bias Is Beyond Reasonable Doubt: A Refutation of Ideological and Methodological Objections and Executive Summary of Ten Studies That No Manager Should Ignore*, 29 *RESEARCH ORG. BEHAV.* 39 (2009); Jerry

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

6. See, e.g., Jennifer Eberhardt et al., *Looking Deathworthy: Perceived Stereotypicality of Black Defendants Predicts Capital-Sentencing Outcomes*, 17 *PSYCHOL. SCI.* 383 (2006); Jerry Kang et al., *Implicit Bias in the Courtroom*, 59 *UCLA L. REV.* 1124 (2012); Samuel R. Sommers, *Race and the Decision Making of Juries*, 12 *LEGAL & CRIMINOLOGICAL PSYCHOL.* 171 (2007).

7. Jeffrey J. Rachlinski et al., *Does Unconscious Racial Bias Affect Trial Judges?* 84 *NOTRE DAME L. REV.* 1195 (2009).

8. Justin D. Levinson & Danielle Young, *Different Shades of Bias: Skin Tone, Implicit Bias, and Judgments of Ambiguous Evidence*, 112 *W. VA. L. REV.* 307 (2010).

**In the discrete area of juror decision making, there is little evidence to suggest a straightforward, simple relationship between defendant race and juror verdict preferences.**

man's photograph rated the defendant's guilt at 66.97 on average compared to 56.37 for jurors who viewed the light-skinned gunman's photograph, suggesting that skin color affected how jurors perceived and interpreted the trial evidence. Other measures of explicit racial bias were unrelated to study findings.

Eberhardt and colleagues investigated the extent to which capital felony defendants with stereotypically black facial features are more likely to be sentenced to death

than defendants without such features.<sup>9</sup> Using a database of death-eligible cases in Philadelphia that advanced to the penalty phase between 1979 and 1999, the researchers identified 44 cases in which a black defendant was convicted of murdering a white victim. They then obtained photographs of these defendants and had neutral observers rate each defendant's looks on a scale of 1 (not at all stereotypically black) through 11 (extremely stereotypical). After controlling for nonracial factors known to influence sentencing,<sup>10</sup> they found that defendants who were rated as having less stereotypically black features were sentenced to death in 24.4% of the cases whereas defendants who were rated as appearing more stereotypically black were sentenced to death in 57.5% of the cases. These findings are consistent with previous research that people associate black physical traits with criminality.

Employing the same methodology, Eberhardt and colleagues also examined death-penalty rates in cases in which a black defendant was convicted of murdering a black victim, but they found no significant difference based on stereotypically black appearances. Noting that juries may view black-on-white crimes as intergroup conflict, rather than interpersonal conflict involved in black-on-black crimes, they concluded that juries may use physical appearance as a powerful cue to in determining whether a defendant deserves to die.

These studies and others<sup>11</sup> demonstrate that racial biases in legal decision making often arise in ways not easily or consistently explained by traditional factors such as trial participant demographic characteristics. In the discrete area of juror decision making, there is little evidence to suggest a straightfor-

ward, simple relationship between defendant race and juror verdict preferences. Mock-juror studies such as the ones discussed here show some evidence of in-group biases,<sup>12</sup> but studies that focus on the decision making of actual jurors in actual trials find that the relationship between juror and defendant race accounted for only a very small amount of the variance in jury verdicts. Strength of evidence is generally the overwhelming predictor. Garvey and colleagues, for example, examined decision making by more than 3,000 jurors in nearly 400 non-capital felony trials in four jurisdictions.<sup>13</sup> Only in the D.C. Superior Court did the defendant's race affect juror's first votes during deliberations, but even this relationship did not survive into the juries' final verdicts. Strength of the evidence, including the credibility of police testimony, was the strongest factor related to final verdicts.<sup>14</sup>

Similarly, Visher conducted 90-minute in-person interviews with 331 jurors from 38 forcible-sexual-assault trials to examine the effects of juror characteristics, defendant and victim characteristics, and evidentiary factors on juror decision making.<sup>15</sup> She found that juror characteristics accounted for only 2% of the variance in jury verdicts, and defendant and victim characteristics accounted for only 8% of the variance. In contrast, evidentiary factors, especially the use of force and physical evidence, accounted for 34% of the variance. These findings point away from strict demographic explanations for racial disparities in legal decision making and toward a more complex, nuanced alternative: one that explores how the decision maker's attitudes and cognitive schemas inform the perception and interpretation of a host of evidentiary factors critical to fair legal judgment.

## **PROMISING STRATEGIES FOR COMBATING IMPLICIT BIAS**

Social scientists have made great strides in recent years to identify effective (and ineffective) interventions for combating more insidious forms of racial bias. To reduce the effects of implicit forms of bias in judgment and behavior, several interventions have shown promise. We discuss a few of these in turn.

### **EDUCATION INTERVENTIONS**

In general, basic education about the existence of implicit forms of bias and how these can manifest in judgment is an important first step. Personal awareness of one's own potential for any type of cognitive bias is necessary before an individual is capable of engaging in efforts to correct for it. Although simply being aware of the potential for racial bias may prompt

9. Eberhardt et al., *supra* note 6.

10. The nonracial factors that researchers controlled for included (1) aggravating circumstances; (2) mitigating circumstances; (3) severity of the murder; (4) defendant's socioeconomic status; (5) victim's socioeconomic status; and (6) defendant's attractiveness. Eberhardt et al., *supra* note 6, at 384.

11. See, e.g., Samuel R. Sommers & Phoebe C. Ellsworth, *White Juror Bias: An Investigation of Prejudice Against Black Defendants in the American Courtroom*, 7 *PSYCHOL. PUB. POLY & L.* 201 (2001); Sommers, *supra* note 6.

12. M. Juliet Bonazzoli, *Jury Selection and Bias: Debunking Invidious Stereotypes Through Science*, 18 *QUINNIAC L. REV.* 775 (1998);

Derek Chadee, *Race, Trial Evidence, and Jury Decision Making*, 1 *CARIBBEAN J. CRIM. & SOC. PSYCHOL.* 59 (1996); Nancy J. King, *Postconviction Review of Jury Discrimination: Measuring the Effects of Juror Race on Jury Decisions*, 92 *MICH. L. REV.* 63 (1993).

13. Stephen P. Garvey et al., *Juror First Votes in Criminal Trials*, 1 *J. EMPIRICAL LEGAL STUD.* 371 (2004).

14. See also PAULA L. HANNAFORD-AGOR, VALERIE P. HANS, NICOLE L. MOTT & G. THOMAS MUNSTERMAN, *ARE HUNG JURIES A PROBLEM?* (2002).

15. Christy A. Visher, *Juror Decision Making: The Importance of Evidence*, 11 *LAW & HUM. BEHAV.* 1 (1987).

some individuals to pursue corrective action, it is not sufficient to ensure that debiasing efforts consistently reduce or prevent expressions of implicit bias.<sup>16</sup> Individuals must understand the nature of implicit bias—what it is and how it can affect judgment—to increase the likelihood that the corrective efforts they engage in are effective.<sup>17</sup> And they must also possess the motivation to fully implement such debiasing efforts.

In framing an educational message on implicit bias, however, the appeal used has important ramifications. For example, one set of studies has shown that some types of individuals are angered and feel threatened by external pressure to comply with mandatory nondiscrimination standards.<sup>18</sup> When away from the watchful eye of the authority figure setting the standards for compliance, these individuals are more likely to engage in biased decision making, presumably in attempts to “reassert their personal freedom.”<sup>19</sup> Thus if an authority designs the educational message to pressure individuals to comply with social or institutional standards for racial fairness, this *extrinsic* motivation to regulate prejudice can incite hostility and generate backlash that may increase expressions of racial prejudice. Legault and colleagues showed that exposing individuals to educational messages designed to compel adherence to racial fairness generated more explicit prejudice (in the form of self-reported racial attitudes) and implicit prejudice (in the form of reaction time measures like the IAT) than a no-intervention alternative. That approach shows that forced-compliance interventions can actually increase expressions of prejudice over doing nothing.<sup>20</sup> In contrast, Legault and colleagues also found that educational messages designed to inform and appeal to personal standards for egalitarianism (*i.e.*, to generate *intrinsic* motivation to regulate prejudice) reduced expressions of explicit and implicit prejudice compared to the no-intervention alternative. Thus interventions designed to educate individuals in an effort to encourage buy-in at a personal level and appeal to these personal egalitarian standards are more likely to reduce expressions of prejudice and avoid harmful backfire effects than programs in which authorities force individuals to comply with external anti-prejudice standards.

In addition, the effectiveness of an educational intervention can depend on the ideology underlying the approach. The traditional institutional approach to racial fairness, referred to in

relevant literature as the colorblindness approach, explicitly directs individuals to ignore race and other differences. This popular colorblindness strategy underlies the mandate that judicial decision makers disregard extralegal factors like race and gender when weighing the evidence. Given the mounting evidence that messages using intrinsic appeals are more effective at reducing prejudice than messages conveying an external pressure to comply, the colorblindness approach is not an optimal bias-reduction strategy. This approach has been shown to generate greater individual expressions of racial bias on both explicit and implicit measures compared to a multiculturalism approach that promotes the value of diversity and encourages individuals to appreciate group differences.<sup>21</sup> In addition to other research showing that a colorblindness approach to reducing expressions of racial prejudice often backfires,<sup>22</sup> a multiculturalism approach has been shown to improve the likelihood that a person will accurately detect instances of racial discrimination when observed, whereas a colorblindness approach produces a reduced likelihood of detection. This trend suggests that the colorblindness approach may appear to work to improve racial fairness but in actuality may result in an underreporting of incidents of racial discrimination.<sup>23</sup>

The counterproductive effects of particular strategies in the promotion of racial fairness can spread beyond individuals in the immediate educational environment. The mainstream popularity of the colorblindness approach can prompt white individuals, in response to implied (but unspoken) social cues to ignore race, to spontaneously adopt a colorblindness strategy to avoid the appearance of racial bias when interacting with a black partner. Strategic demands to ignore race as part of a colorblindness approach to reducing racial prejudice can produce unintentional, “ironic” effects: One study showed that as white individuals devoted mental resources to the task of ignoring race, they had fewer resources available to dedicate toward monitoring and controlling their nonverbal behavior in social interactions with black partners. As a result of this cognitive

**In framing an educational message on implicit bias, . . . the appeal used has important ramifications.**

16. See Alexander R. Green, Dana R. Carney, Daniel J. Pallin, Long H. Ngo, Krista L. Raymond, Lisa I. Iezzoni & Mahzarin R. Banaji, *Implicit Bias Among Physicians and Its Prediction of Thrombolysis Decisions for Black and White Patients*, 22 J. GEN. INTERNAL MED. 1231 (2007); Do-Yeong Kim, *Voluntary Controllability of the Implicit Association Test (IAT)*, 66 SOC. PSYCHOL. Q. 83 (2003).

17. See Timothy D. Wilson & Nancy Brekke, *Mental Contamination and Mental Correction: Unwanted Influences on Judgments and Evaluations*, 116 PSYCHOL. BULL. 117 (1994); Dwayne T. Wegener et al., *Flexible Correction of Juror Judgments: Implications for Jury Instructions*, 6 PSYCHOL. PUB. POLY & L. 629 (2000). For example, engaging in a strategy like perspective-taking (*i.e.*, imagining how one would feel and act if in the stigmatized person's shoes or if seeing the world from their point of view) has been shown to decrease expressions of implicit and explicit bias. Adam D. Galinsky & Gordon B. Moskowitz, *Perspective-Taking: Decreasing Stereotype Expression, Stereotype Accessibility, and Ingroup Favoritism*, 78 J.

PERSONALITY & SOC. PSYCHOL. 708 (2000).

18. E. Ashley Plant & Patricia G. Devine, *Responses to Other-Imposed Pro-Black Pressure: Acceptance or Backlash?* 37 J. EXPERIMENTAL SOC. PSYCHOL. 486 (2001).

19. *Id.* at 486.

20. Lisa Legault et al., *Ironic Effects of Antiprejudice Messages: How Motivational Interventions Can Reduce (But Also Increase) Prejudice*, 22 PSYCHOL. SCI. 1472 (2011).

21. Jennifer A. Richeson & Richard J. Nussbaum, *The Impact of Multiculturalism Versus Color-Blindness on Racial Bias*, 40 J. EXPERIMENTAL SOC. PSYCHOL. 417 (2004).

22. Evan P. Apfelbaum et al., *Seeing Race and Seeming Racist? Evaluating Strategic Colorblindness in Social Interaction*, 95 J. PERSONALITY & SOC. PSYCHOL. 918 (2008).

23. Evan P. Apfelbaum et al., *In Blind Pursuit of Racial Equality?* 21 PSYCHOL. SCI. 1587 (2010).

**More diverse juries tend to produce decisions less biased by the defendant's race . . . .**

resource strain, white participants actually exhibited less friendly non-verbal behavior toward the black partner than was observed when white participants interacted with a white partner.<sup>24</sup> Other research indicates that this kind of discriminatory non-verbal behavior can negatively influence the subsequent perceptions and responses of the stigmatized individual or individuals<sup>25</sup> and may hinder

future efforts to engage in interracial interaction.<sup>26</sup>

Thus, educational initiatives promoting racial fairness should focus not just on the individual, but also on the climate of the organization as a whole. When peers and leadership figures demonstrate behavior consistent with the multiculturalism approach, these expressions of egalitarian goals and beliefs will positively influence others in the community.<sup>27</sup> Other educational efforts to change racial attitudes through diversity-training courses have also helped to at least temporarily reduce individuals' expressions of implicit and explicit racial biases.<sup>28</sup>

**CONTACT AND EXPOSURE INTERVENTIONS**

Generally, increased interracial contact seems to have a positive effect on both implicit and explicit attitudes.<sup>29</sup> Exposure to individuals who contradict prevailing cultural or social stereotypes can, in particular, reduce the expression of implicit racial biases. This works when people have an opportunity to see or interact with stigmatized group members in respected leadership roles or as role models, or otherwise observe them behaving in a manner that contradicts prevailing social stereotypes.<sup>30</sup>

Simply imagining stigmatized group members in counter-stereotypic ways can also reduce the expression of implicit biases.<sup>31</sup> Even individuals who engage in extensive practice mentally countering or negating stereotypes appear to be able to successfully reduce implicit biases based on those stereotypes over time.<sup>32</sup>

More diverse juries tend to produce decisions less biased by the defendant's race, presumably because they force jurors to engage with one another on an equal basis in deliberations and to expressly confront different conclusions about the trial evidence that were reached based on the jurors' unique life experiences and attitudes, including implicit biases.<sup>33</sup> Although more research is needed on the precise mechanisms by which jury diversity affects juror decision making, it appears that the presence of minorities on a jury not only brings more diverse perspectives to the table, but it also increases white juror awareness of race-related concerns in a way that stimulates a more thorough and more factually accurate evaluation and discussion of trial evidence.

**INTERVENTIONS THAT CLARIFY STANDARDS FOR JUDGMENT**

Discrimination tends to emerge more in ambiguous decision-making contexts than straightforward ones.<sup>34</sup> White-majority juries more often convict and recommend harsher sentences for black defendants than white defendants when the prosecution presents weak or ambiguous evidence against them.<sup>35</sup> Other studies show that mock jurors are more likely to discriminate against black defendants than white defendants in verdict and sentencing decisions when presented with mixed or incriminating but inadmissible evidence.<sup>36</sup> To check for poten-

24. Apfelbaum et al., *supra* note 22.

25. See generally John F. Dovidio et al., *Why Can We Just Get Along? Interpersonal Biases and Interracial Distrust*, 8 CULTURAL DIVERSITY & ETHNIC MINORITY PSYCHOL. 88 (2002); Mark Snyder, Elizabeth D. Tanke & Ellen Berscheid, *Social Perception and Interpersonal Behavior: On the Self-fulfilling Nature of Social Stereotypes*, 35 J. PERSONALITY & SOC. PSYCHOL. 656 (1977); Mark Chen & John A. Bargh, *Nonconscious Behavioral Confirmation Processes: The Self-fulfilling Nature of Automatic Stereotype Activation*, 33 J. EXPERIMENTAL SOC. PSYCHOL. 541 (1997).

26. J. Nicole Shelton & Jennifer Richeson, *Intergroup Contact and Pluralistic Ignorance*, 88 J. PERSONALITY & SOC. PSYCHOL. 91 (2005).

27. Henk Aarts, Peter M. Gollwitzer & Ran R. Hassin, *Goal Contagion: Perceiving Is for Pursuing*, 87 J. PERSONALITY & SOC. PSYCHOL. 23 (2004); Gretchen B. Sechrist & Charles Stangor, *Perceived Consensus Influences Intergroup Behavior and Stereotype Accessibility*, 80 J. PERSONALITY & SOC. PSYCHOL. 645 (2001).

28. Lauri A. Rudman et al., "Unlearning" Automatic Biases: *The Malleability of Implicit Prejudice and Stereotypes*, 81 J. PERSONALITY & SOC. PSYCHOL. 856 (2001).

29. See Thomas F. Pettigrew & Linda R. Tropp, *A Meta-Analytic Test of Intergroup Contact Theory*, 90 J. PERSONALITY & SOC. PSYCHOL. 751 (2006).

30. 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 (2004); 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 (2001); Nilanjana Dasgupta & Luis M. Rivera, *When Social Context Matters: The Influence of Long-Term Contact and Short-Term Exposure to Admired Outgroup Members on Implicit Attitudes and Behavioral Intentions*, 26 SOC. COGNITION 112 (2008); Michael A. Olson & Russell H. Fazio, *Reducing Automatically Activated Racial Prejudice Through Implicit Evaluative Conditioning*, 32 PERSONALITY & SOC. PSYCHOL. BULL. 421 (2006).

31. Irene V. Blair, Jennifer E. Ma & Alison P. Lenton, *Imaging Stereotypes Away: The Moderation of Implicit Stereotypes Through Mental Imagery*, 81 J. PERSONALITY SOC. PSYCHOL. 828 (2001).

32. Kerry Kawakami et al., *Just Say No (to Stereotyping): Effects of Training in the Negation of Stereotypic Associations on Stereotype Activation*, 78 J. PERSONALITY & SOC. PSYCHOL. 871 (2000).

33. 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).

34. Eric Luis Uhlmann & Geoffrey L. Cohen, *Constructed Criteria: Redefining Merit to Justify Discrimination*, 16 PSYCHOL. SCI. 474 (2005); John F. Dovidio & Samuel L. Gaertner, *Aversive Racism & Selection Decisions: 1989 and 1999*, 4 PSYCHOL. SCI. 315 (2000).

35. Chadee, *supra* note 12.

36. Gordon Hodson et al., *Aversive Racism in Britain: Legal Decisions and the Use of Inadmissible Evidence*, 35 EUR. J. SOC. PSYCHOL. 437 (2005); James D. Johnson et al., *Justice Is Still Not Colorblind: Differential Racial Effects of Exposure to Inadmissible Evidence*, 21 PERSONALITY & SOC. PSYCHOL. BULL. 893 (1995).

tial bias, a decision maker may look to determine if he or she has reasonable justification for the decision based on legitimate decision-making factors. However, this research shows that it is difficult for decision makers to realize when their decisions are influenced by race, ethnicity, gender, or other extraneous factors if other selective information can be used to support their decision.

People may not be able to identify and correct for bias in ambiguous contexts because decision-making standards tend to change to rationalize judgments that are actually influenced by extraneous factors. In a seminal series of studies, Uhlmann and Cohen showed that when evaluating male and female job applicants for a gender-stereotypical job (e.g., a stereotypically masculine position as a police chief), people's perceptions about the credentials needed to be successful at the job tended to shift *post hoc* to justify gendered decision making.<sup>37</sup> That is, regardless of whether the male had "street smarts" or a strong educational background, people tended to justify their decision to select the male candidate over the female candidate by claiming that whichever credential the male had (but that the female did not) was more important to the job. Most telling is the fact that these decision makers thought they were providing an objective, rational, unbiased decision about the best candidate to hire.

If clear decision-making criteria are defined at the outset, however, the type of "shifting standards" effect that can unintentionally result in discrimination may be eliminated. In a follow-up study by Uhlmann and Cohen, people who committed to clear priorities about the criteria they would use in making the police-chief hiring decision showed no evidence of gendered decision making, but people who did not make such a commitment and relied more on discretionary, selective justification made decisions that were biased by gender. This shows that clarified decision-making standards can reduce stereotyping and discrimination in outcomes.

#### **POTENTIAL APPLICATIONS OF INTERVENTION STRATEGIES FOR USE WITH JURIES**

Although interventions have shown promise in reducing the effects of implicit bias on judgment and behavior more generally, not all of these strategies lend themselves well to practical application in jury decision making. We discuss some potential intervention strategies for use with juries below and consider the feasibility of each.

#### **EDUCATE JURORS ON IMPLICIT BIAS**

Education or training on the topic of implicit bias has been provided to judges and court staff in some states.<sup>38</sup> However,

even the most conservative of these educational initiatives take a significant amount of time to implement and require substantial resources and preparation. To introduce training during jury selection, courts would need to have trainers available to present educational material to prospective jurors, the time to implement a seminar that would likely last one to two hours at minimum due to the complexity of the subject

matter, and the resources to allow prospective jurors to explore the topic through feedback or the opportunity for practice. Moreover, most Americans now live in jurisdictions that employ a one-day/one-trial term of service.<sup>39</sup> This system substantially reduces the burden of jury service on individual jurors by distributing it across a much larger pool of prospective jurors. Courts that have implemented this system necessarily had to abandon the earlier practice of summoning prospective jurors for an "Orientation Day" and now conduct a brief juror orientation (typically 20-30 minutes) in the morning when jurors report and before they are sent to courtrooms for jury selection. The combination of resources required for an educational program on implicit bias plus the lack of time in which to present such a program makes it unlikely that any court would pursue this intervention option.

#### **RAISE AWARENESS OF IMPLICIT BIAS WITH IATS OR RACE-RELEVANT VOIR DIRE QUESTIONING**

A number of alternative interventions have been suggested for use in jury selection, but little is yet known about their efficacy in reducing bias. For example, the administration of IATs to jurors<sup>40</sup> and the addition of race-relevant voir dire questioning<sup>41</sup> have been proposed as means of raising juror awareness about implicit bias and alerting jurors to its potential influence on their decisions. Although such approaches may help to reduce expressions of bias, these options are impractical in many courts for many of the same reasons discussed with regard to the educational-seminar option above. Costs associated with these techniques (e.g., printing and processing of questionnaires at a time when states are facing new and significant budgetary challenges) and limited existing court resources (e.g., computer access for potential jurors to take the IAT, or trained staff to code and process a paper-and-pencil version of the test) preclude these options from consideration in many jurisdictions.

**A number of alternative interventions have been suggested for use in jury selection, but little is yet known about their efficacy . . . .**

37. Uhlmann & Cohen, *supra* note 34.

38. See PAMELA M. CASEY ET AL., HELPING COURTS ADDRESS IMPLICIT BIAS: RESOURCES FOR EDUCATION (2012); Pamela M. Casey, Roger K. Warren, Fred L. Cheesman, & Jennifer K. Elek, *Addressing Implicit Bias in the Courts*, 49 CT. REV. 64 (2013).

39. GREGORY E. MIZE, PAULA L. HANNAFORD-AGOR & NICOLE L. WATERS, THE STATE-OF-THE-STATES SURVEY OF JURY IMPROVEMENT EFFORTS: A COMPENDIUM REPORT (2007). Under a one-day/one-trial term of service, prospective jurors report to the courthouse on the sum-

mons date. If a juror is sworn as a trial juror or alternate, he or she serves for that trial and is released from further service at the completion of the trial. If a juror is not selected as a trial juror or alternate, he or she is released at the end of the reporting day.

40. Anna Roberts, *(Re)forming the Jury: Detection and Disinfection of Implicit Juror Bias*, 44 CONN. L. REV. 827 (2012).

41. Regina A. Schuller, Veronica Kazoleas & Kerry Kawakami, *The Impact of Prejudice Screening Procedures on Racial Bias in the Courtroom*, 33 LAW & HUM. BEHAV. 320 (2009); Sommers, *supra* note 33.

**The potential efficacy of . . . debiasing agents must be examined and demonstrated empirically . . . before courts are encouraged to use them . . .**

The National Center for State Courts' State-of-the-States Survey of Jury System Improvement Efforts further illustrates the limited viability of debiasing interventions during voir dire.<sup>42</sup> In nearly 12,000 jury trials, judges and lawyers reported that they spent two hours on average to select a jury, a task that also involves confirming jurors' qualifications and ability to serve for the length of the trial and investigating each juror's ability to be fair and impartial if selected as a

trial juror. Most judges and lawyers would not embrace new debiasing interventions during voir dire due to the additional time involved. States also vary in the extent to which voir dire is judge-dominated or attorney-dominated, but in either case, voir dire is perhaps the most individualistic stage of the trial. Judges have a great deal of discretion in how they conduct voir dire and are protective of that discretion as a matter of judicial independence. In a judge-dominated voir dire state, the likelihood of training the entire trial bench on how to use debiasing interventions effectively in all cases, and then ensuring that they actually apply that training, is very remote. Doing so in a lawyer-dominated voir dire state is even more remote given the complete absence of a unified bar. For consistent use in the majority of state courts, a realistic practical remedy for implicit bias in juror decision making must be not only effective but also expedient and economical.

#### **ASSEMBLE DIVERSE JURIES**

Convention assumes that deliberations among a demographically diverse group of jurors are more likely to facilitate a thorough and fair evaluation of the evidence because different perspectives are presumed to be represented in the discussion. Moreover, as indicated above, research shows that when white jurors expect to engage with a diverse jury, they tend to approach deliberations in a way that promotes a more thorough and factually accurate evaluation of the evidence.<sup>43</sup>

It is not always possible, however, to ensure a racially diverse jury. This may be of particular concern in jurisdictions with relatively homogeneous jury pools, which comprise the great majority of state courts. For example, the jury-eligible population of black/African-Americans comprises less than 10 percent

of the total jury-eligible population in more than three-quarters of counties in the United States. Those counties encompass more than half of the total U.S. population.<sup>44</sup> Unfortunately, even in more diverse communities, jury panels often fail to fully reflect community demographic characteristics due to non-systematic exclusion of minorities from jury pools,<sup>45</sup> reductions in the size of trial juries,<sup>46</sup> and the pervasive discriminatory use of peremptory challenges.<sup>47</sup>

#### **STRENGTHEN THE JURY DECISION-MAKING PROCESS**

The past two decades have seen a dramatic change in judges' management of jury trials. The traditional view that juror impartiality is best served when jurors maintain a strictly passive role has gradually given way to the view that jurors are active learners and perform best when given commonplace decision-making tools to better understand and remember trial evidence. These tools include permitting jurors to take notes, permitting jurors to submit written questions to witnesses, permitting jurors to discuss the evidence before final deliberations, instructing jurors on the basic elements of the law they will be told to apply before the evidentiary portion of the trial, and providing jurors with written copies of jury instructions.<sup>48</sup> Evaluations of each of these innovations have shown that they are effective decision-making aids in terms of improved comprehension of the evidence and law and increased retention of evidence presented at trial. By emphasizing the importance of juror comprehension of the evidence and law, these types of tools provide jurors with a stronger framework for decision making and lead to greater clarity about the basis for their collective verdict, which theoretically should reduce the potential for implicit bias to skew the verdict. No research has been conducted to explicitly examine the relationship between these jury-trial innovations and implicit bias. Although there is some reason to believe these tools may be helpful for this purpose, it is premature to conclude that these innovations will reduce the impact of implicit bias on jury verdicts. The potential efficacy of these tools as debiasing agents must be examined and demonstrated empirically, through rigorous scientific research, before courts are encouraged to use them as implicit-bias interventions.

#### **USE SPECIALIZED JURY INSTRUCTIONS ON IMPLICIT BIAS**

Historically, courts have relied extensively on jury instructions to guide juror decision making because this approach is relatively inexpensive, expedient, and easy to administer to

42. MIZE et al., *supra* note 39.

43. Sommers, *supra* note 33.

44. Paula Hannaford-Agor & Nicole L. Waters, *Safe Harbors from Fair Cross Section Challenges? The Practical Limitations of Measuring Representation in the Jury Pool*, 8 J. EMPIRICAL LEGAL STUD. 762 (2011).

45. Samuel R. Sommers, *Determinants and Consequences of Jury Racial Diversity: Empirical Findings, Implications, and Directions for Future Research*, 2 SOC. ISSUES & POL'Y REV. 65 (2008).

46. Shari Seidman Diamond, Mary R. Rose, Beth Murphy & John B. Meixner, *Damage Anchors on Real Juries* (Sept 20, 2011) available

at <http://ssrn.com/abstract=1883861> or <http://dx.doi.org/10.2139/ssrn.188386>.

47. See, e.g., EQUAL JUSTICE INITIATIVE, *ILLEGAL RACIAL DISCRIMINATION IN JURY SELECTION: A CONTINUING LEGACY* (2010); COUNCIL FOR COURT EXCELLENCE, *DISTRICT OF COLUMBIA JURY PROJECT, JURIES FOR THE YEAR 2000 AND BEYOND: PROPOSALS TO IMPROVE THE JURY SYSTEMS IN WASHINGTON, D.C.* 19-37 (1998); Kenneth J. Melilli, *Batson in Practice: What We Have Learned About Batson and Peremptory Challenges*, 71 NOTRE DAME L. REV. 447 (1996).

48. See, e.g., *JURY TRIAL INNOVATIONS* (G. Thomas Munsterman, Paula L. Hannaford-Agor & G. Marc Whitehead eds., 2d ed. 2006).

each new jury. However, research studies have provided mixed evidence of its utility in practice. On one hand, most studies confirm that jurors take their responsibility to apply the law provided by the trial judge seriously, spending up to one-quarter of their deliberation time focused on jury instructions.<sup>49</sup> On the other hand, although most jurors in actual trials report that they understand the law relatively well,<sup>50</sup> research shows that jurors have fairly low levels of comprehension regarding the basic legal principles articulated in jury instructions.<sup>51</sup> But when Diamond and her colleagues observed actual jury deliberations in 50 civil trials, they found that nearly 80% of the jurors' comments about the instructions were accurate and nearly half of the incorrect comments were ultimately corrected during deliberations. This led Diamond and her colleagues to surmise that jurors in actual trials might be "able to assist one another in ways not captured on post-deliberation questions or in studies of individual respondents."<sup>52</sup> The implication from this research is that to understand the full impact that any jury instruction (including a specialized implicit-bias jury instruction) may have on juror decision making, one should examine it in a context in which group deliberations take place.

It is not yet known whether a well-crafted jury instruction could help to mitigate the effect of implicit racial bias in juror decision making. Studies show that individuals can control the behavioral expression of implicit biases in specific laboratory contexts if provided with a concrete strategy for bias reduction.<sup>53</sup> In addition, whether or not jurors are motivated to produce a fair and just outcome can determine whether debiasing instructions are followed. However, pattern jury instructions developed for use in state and federal jury trials rarely incorporate these characteristics, relying instead on the simple admonition that jurors should not let "bias, sympathy, prejudice, or public opinion influence your decision."<sup>54</sup> Moreover, jury instructions tend to be written in an authoritarian legal style

that, in the context of implicit bias, may ultimately prove counterproductive by triggering a backlash effect.<sup>55</sup>

## CONCLUSIONS

It is clear that members of the court community are coming to understand the general problem posed by implicit bias and are clamoring for readily available solutions on which they can act. As the court community has become more knowledgeable about implicit bias and more aware of the potential for harm in judicial decision making, judges and lawyers have expressed a great deal of interest in extending intervention efforts to jurors through the development of a specialized jury instruction on implicit bias. Judge Mark Bennett of the U.S. District Court, Northern District of Iowa, was the first trial judge known to have incorporated this approach in jury trials.<sup>56</sup> More recently, the Criminal Justice Section of the American Bar Association has convened a committee to develop a toolbox of options intended to reduce the impact of implicit bias in court proceedings, including a jury instruction on implicit bias.<sup>57</sup> The topic of implicit-bias instructions has also been a recurring theme on listserv discussions among members of pattern jury instruction committees.<sup>58</sup> Through these efforts and others, several versions of implicit-bias jury instructions are now or will soon be available for use.

Unfortunately, existing research suggests the possibility that an implicit-bias jury instruction may produce a backlash effect that actually exacerbates expressions of both implicit and explicit bias. This effect may not be universal: Specialized

**It is not yet known whether a well-crafted jury instruction could help to mitigate the effect of implicit racial bias in juror decision making.**

49. Shari Seidman Diamond, Beth Murphy & Mary R. Rose, *The "Kettleful of Law" in Real Jury Deliberations: Successes, Failures, and Next Steps*, 106 NW. U. L. REV. 1537 (2012); see REID HASTIE, STEVEN PENROD & NANCY PENNINGTON, *INSIDE THE JURY* (1983).

50. See B. MICHAEL DANN, VALERIE P. HANS & DAVID H. KAYE, *TESTING THE EFFECTS OF SELECTED JURY TRIAL INNOVATIONS ON JUROR COMPREHENSION OF CONTESTED mtDNA EVIDENCE: FINAL TECHNICAL REPORT FOR U.S. DEPARTMENT OF JUSTICE* (2004), available at <https://www.ncjrs.gov/pdffiles1/nij/grants/211000.pdf>; Paula Hanaford-Agor, Valerie P. Hans, Nicole L. Mott & G. Thomas Munsterman, *The Timing of Opinion Formation by Jurors in Civil Cases: An Empirical Examination*, 67 TENN. L. REV. 627 (2000); Hanaford-Agor, Hans & Munsterman, *supra* note 14.

51. Phoebe C. Ellsworth, *Are Twelve Heads Better Than One?* 52 LAW & CONTEMPORARY PROBLEMS 205 (1989); Alan Reifman, Spencer M. Gusick & Phoebe C. Ellsworth, *Real Jurors' Understanding of the Law in Real Cases*, 16 LAW & HUM. BEHAV. 539 (1992); Geoffrey P. Kramer & Doreen M. Koenig, *Do Jurors Understand Criminal Jury Instructions? Analyzing the Results of the Michigan Juror Comprehension Project*, 23 U. MICH. J.L. REFORM 401 (1990); Bradley Saxton, *How Well Do Jurors Understand Jury Instructions? A Field Test Using Real Juries and Real Trials in Wyoming*, 33 LAND & WATER L. REV. 59 (1998).

52. See Diamond, Rose & Murphy, *supra* note 49.

53. See Kim, *supra* note 16; Saaid Mendoza, Peter Gollwitzer & David Amodio, *Reducing the Expression of Implicit Stereotypes: Reflexive Control through Implementation Intentions*, 36 PERSONALITY & SOC. PSYCHOL. BULL. 512 (2010); Brandon D. Stewart & B. Keith Payne, *Bringing Automatic Stereotyping Under Control: Implementation Intentions as Efficient Means of Thought Control*, 34 PERSONALITY & SOC. PSYCHOL. BULL. 1332 (2008).

54. Judicial Conference Criminal Jury Instructions, CALCRIM No. 101 (2012).

55. See *supra* notes 18-20 and accompanying text.

56. Judge Bennett, a former civil-rights lawyer, shares his unnerving discovery of his own disappointing IAT results in Mark W. Bennett, *Unraveling the Gordian Knot of Implicit Bias in Jury Selection: The Problems of Judge-Dominated Voir Dire, the Failed Promise of Batson, and Proposed Solutions*, 4 HARV. L. & POL'Y REV. 149 (2010).

57. For a description of the Committee and its work, see [http://www.americanbar.org/groups/criminal\\_justice/voir\\_dire.html](http://www.americanbar.org/groups/criminal_justice/voir_dire.html).

58. The PJI Listserv is an online discussion group hosted by the NCSC and composed of chairs and reporters and state and federal pattern-jury-instruction committees.

59. SJI-13-N=082 (Pilot Test of an Implicit Bias Jury Instruction).



implicit-bias jury instructions may successfully reduce expressions of bias with some types of jurors but elicit backlash from others. Consequently, to prevent the dissemination of harmful jury instructions that produce backlash effects, we strongly recommend that new jury instructions be carefully evaluated using rigorous empirical methods to determine their overall and differential effectiveness before they are broadly promoted for use in the courtroom.

To begin this process, the National Center for State Courts (NCSC) is currently engaged in an effort to test the efficacy of an implicit-bias jury instruction. With funding from the State Justice Institute,<sup>59</sup> the NCSC has undertaken a project to draft a model jury instruction on implicit bias and, using mock-jury methods with a vignette of a fictitious trial, to pilot test the instruction to determine its effectiveness in minimizing the impact of implicit bias in juror decision making. The results of the pilot study, which should be available in late 2013, will help inform the direction of future efforts to address implicit bias in jury trials.



*Jennifer K. Elek, Ph.D., is a court research associate with the National Center for State Courts. In Dr. Elek's recent work at NCSC, she has focused on promoting gender and racial fairness in the courts; on improving survey-based judicial performance evaluation programs; on educating the court community about offender risk and needs assessment and its role in evidence-based sentencing; and on identifying and evaluating the efficacy of problem-solving court programs. She holds a Ph.D. in social psychology from Ohio University, an M.A. from the College of William and Mary, and a B.A. from Vassar College.*



*Paula Hannaford-Agor is the director of the NCSC Center for Jury Studies. She joined the NCSC Research Division in May 1993 and routinely conducts research and provides courts and court personnel with technical assistance and education on the topics of jury system management, jury trial procedure, and juror decision making. She has authored or contributed to numerous books and articles on the American jury, including *Jury Trial Innovations* (2d ed. 2006), *The Promise and Challenges of Jury System Technology* (NCSC 2003), and *Managing Notorious Trials* (1998). She is faculty for the ICM courses *Jury System Management* and *Promise and Challenges of Jury System Technology*. As adjunct faculty at William & Mary Law School, she teaches a seminar on the American jury. She holds a J.D. and a Masters in Public Policy from the College of William and Mary.*