Implicit Bias and the American Juror

Jennifer K. Elek & Paula Hannaford-Agor

On August 9, 2014, white police officer Darren Wilson shot an unarmed black civilian named Michael Brown. In the wake of the extensive media coverage and public scrutiny that followed, the tragic incident triggered a federal civil-rights investigation and renewed a broader national dialogue about the perception of black males as inherently dangerous, threatening, or criminal—and the role of those perceptions in perpetuating racial inequality in the United States. A grand jury decision not to indict Wilson on murder charges elicited highly polarized reactions from the general public, which ranged from wholehearted support for Officer Wilson to further accusations of racial bias against the prosecutor in the case and against the predominantly white grand jury charged with making the decision. The President of the United States addressed the diverse sentiments of the American people:

There are Americans who agree with it [the Ferguson grand jury decision] and there are Americans who are deeply disappointed—even angry . . . . We have made enormous progress in race relations over the course of the past several decades. I have witnessed that in my own life, and to deny that progress, I think, is to deny America's capacity for change. But what is also true is that there are still problems—and communities of color aren't just making these problems up. Separating that from this particular decision, there are issues in which the law too often feels as if it is being applied in a discriminatory fashion.

On March 4, 2015, the United States Department of Justice (DOJ) released a report of findings in its investigation of the Ferguson, Missouri, criminal-justice system. The report contributed to the still-growing body of literature acknowledging contemporary racial inequality and recognizing that these disparities may not be explained on the basis of people's explicit, intentional biases alone. In the report, the DOJ described evidence of systematic racial discrimination in the community's policing and municipal court practices. Observed disparities in treatment were "unexplainable on grounds other than race and evidence that racial bias, whether implicit or explicit, has shaped law enforcement conduct," resulting in what they concluded to be disproportionate harm to Ferguson's African-American residents. In discussing the DOJ report, Attorney General Eric Holder described Ferguson as "a community where this harm frequently appears to stem, at least in part, from racial bias—both implicit and explicit." Although the possible effects of implicit bias on justice-system outcomes should be considered at each decision point in case processing, we focus in this article on the potential effects of implicit bias in the decision making of everyday American citizens who are randomly selected to serve on grand juries and in jury trials. We begin with a brief explanation of the concept of implicit bias and examine one type of intervention that some believe could address this subtler form of racial bias in jury decision making: a specialized jury instruction.

**IMPLICIT BIAS AND ITS ROLE IN JUROR DECISION MAKING**

In understanding how racial bias can continue to operate in the context of the modern American jury, one must account for both forms of racial bias identified by the Attorney General and the DOJ in the Ferguson report. This includes explicit bias, the form of bias that a person intentionally endorses (and the traditional definition of racial prejudice that most people recognize), but also implicit bias, a form of bias that occurs when a person makes associations between a group of people and particular traits that then operate without self-awareness to affect one's perception of, understanding about, or behavior toward others. People develop these associations (i.e., attitudes and stereotypes) between particular social groups and particular qualities (for example, one study showed that many participants implicitly associate "black" with "guilty," and other

**Footnotes**


6. Anthony G. Greenwald & Mahzarin R. Banaji, Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes, 102 PSYCHOL. REV. 4 (1995). After a Florida jury acquitted Hispanic neighborhood-watch coordinator George Zimmerman of second-degree-murder and manslaughter charges in the death of unarmed black teen Trayvon Martin, one outspoken juror explained her reasoning in the case to the media. She said she had “no doubt” that Zimmerman “feared for his life” in his encounter with Martin and that “anybody would think anybody walking down the road, stopping and turning and looking—if that’s exactly what [Martin did]—is suspicious.” Scientists and legal scholars have concluded that
studies have shown associations between African-Americans and negative characteristics such as aggression and hostility) as they learn from their social environment (e.g., media, parental, or peer role models). The influence of these attitudes or stereotypes in producing a discriminatory response may occur involuntarily and without a person's conscious intent. That is, individuals may explicitly report egalitarian racial attitudes but can nevertheless still make racially biased decisions and behave in racially biased ways.

Findings in the scientific research literature demonstrate how implicit bias can operate to distort a person's interpretations of the evidence in a case. Racial stereotypes have been found to play a role in how people perceive and interpret otherwise ambiguous events. For example, one study found that people interpret ambiguously hostile behavior as more hostile when performed by a black compared with a white actor. Similarly, people who test high on implicit racial bias were found to be more likely to interpret ambiguous expressions in a negative manner (i.e., as angrier) on black faces (but not white faces) compared with those who test low on implicit racial bias. Another recent study found that presenting mock jurors with images of darker-skinned (compared to lighter-skinned) perpetrators biased their interpretations of ambiguous evidence. Biased interpretations of the evidence, in turn, predicted subsequent guilty verdicts.

**ADDRESSING IMPLICIT BIAS WITH JURORS**

In recent years, court leaders across the country have recognized the challenge posed by implicit forms of bias and have focused on addressing this issue in the criminal-justice system through in-depth education and training of judges and court staff. However, unlike with judges and court staff, the courts have limited opportunities to educate jurors about the pernicious effects of complex psychological phenomena like implicit bias. Most jurors in this country serve only for the duration of a trial (typically two to three days) and then are released from service. There is no time available during this short period to provide the type of in-depth education that judges and court staff may receive on strategies to reduce the impact of implicit bias in their decision making. Other solutions that might operate effectively in other settings are not always feasible for use in jury trials, where these time constraints and additional resource limitations (e.g., funding, staffing, technology) often prohibit more elaborate interventions. For primarily practical reasons, courts have historically relied on jury instructions to guide juror decision making.

Presently, pattern jury instructions developed for use in state and federal jury trials typically rely on the simple admonition that jurors should not let “bias, sympathy, prejudice, or public opinion influence [their] decision.” Only a few studies have attempted to explore the utility of standard pattern jury instructions, but they have produced mixed findings. Specifically, one meta-analysis of 16 mock-juror studies (a) found that mock jurors were likely to “render longer sentences for other-race defendants,” (b) found that racial bias in these mock-juror studies was “more pronounced . . . for Black participants; when community members were participants; and in published studies,” but (c) concluded that racial bias in mock juror verdicts “decrease[d] when ecologically relevant procedures [were] used.” Alternatively, one recent study showed that pattern presumption-of-innocence instructions, “a core legal principle specifically designed to eliminate bias,” actually primed greater
attention to race (i.e., to attend more to black individuals) among mock jurors compared to an alternative matched-length instruction. This type of attentional bias, in which decision-makers focus more of their visual attention on black individuals, has been shown in other research to lead to racially biased interpretations of events and racially biased decisions. When considering this mixed evidence in combination with (a) research indicating that people may not be able to consciously correct for the effect(s) of their implicit biases because they are often unaware that these biases exist and (b) evidence of racial discrimination in actual jury trials, standard legal instructions do not appear to offer a complete or reliable solution.

As a next step, to address concerns of both implicit and explicit racial bias among jurors, some judges and lawyers have expressed interest in developing a specialized jury instruction, with at least one midwestern district court judge (now retired) having already used a specialized jury instruction on implicit bias at trial. Crafting clear, effective jury instructions on the topic of implicit bias, however, requires extensive subject-matter expertise for two main reasons. First, subject-matter expertise is necessary to ensure that the language and strategies used in the instruction are accurate reflections of the state of the science. The high level of subject-matter expertise necessary to leverage lessons learned from existing research and provide jurors with appropriate debiasing strategies may not be available among the law-trained professionals who typically comprise committees that draft pattern jury instructions. Second, subject-matter expertise is also needed to ensure that the developed instruction intervention does not incorporate communication strategies known to exacerbate expressions of racial bias in certain subpopulations. For example, strategies that impress an extrinsic authoritarian language typical of jury instructions may provoke hostility and resistance from some individuals, failing to reduce and perhaps even exacerbating expressions of prejudice. Instead, communications designed to foster intrinsic egalitarian motivations may more effectively reduce both explicit and implicit expressions of prejudice without eliciting such backfire or backlash effects. These and other research findings are important to consider for those looking to adopt a jury instruction to minimize expressions of implicit biases in juror judgment.

Any new jury instruction should be carefully evaluated to determine its actual impact on decision making before broadly promoting the instruction as a solution for general use in the courtroom. Jury instructions designed to achieve specific cognitive processing or decision-making objectives are not always a completely effective solution, as has been well documented in prior studies on instructions to disregard inadmissible evidence. Empirical scrutiny is particularly important with any jury instruction on implicit biases given the possibility that a specialized instruction (a) may successfully reduce expressions of racial bias with some jurors yet exacerbate expressions of bias in others and/or (b) may serve to increase juror attention to race in a way that might increase the likelihood of biased outcomes. To date, no known studies have examined the efficacy of a specialized jury instruction informed by the research on reducing implicit forms of bias.

IMPLICIT-BIAS JURY INSTRUCTIONS: A MOCK-JUROR EXPERIMENT

In the present study, we examined for the first time the efficacy of a specialized jury instruction in reducing racial disparities in juror judgments. The National Center for State Courts (NCSC) consulted with nationally recognized social-science and legal experts on implicit bias to develop a specialized jury instruction for use in a web-based experiment with jury-eligible U.S. citizens. In the study, NCSC adapted a trial scenario from a seminal research study that had effectively demonstrated racial bias in juror decision making (an effect which has since been replicated in several experiments using similar if not identical research methodologies). The trial scenario featured a defendant who was charged with assault and battery with intent to cause serious bodily injury. The defendant and victim were described as teammates on a college basketball team, and the alleged assault resulted from a locker-room fight. Prior studies showed that white jurors tended to judge black defendants more harshly.


20. See Kang et al., supra note 12. In addition, the American Bar Association Criminal Justice Section assembled a task force that attempted to develop its own specialized jury instruction on implicit bias. (Personal communication with S. Cox and S. Redfield (June 3, 2013)).


24. Young et al., supra note 17.

25. See pages 113-114 for the specialized jury instruction used in this experiment. We have annotated the experimental instruction we tested, citing the existing scientific research evidence that formed the theoretical basis for each of the components of the instruction.

26. Samuel R. Sommers & Phoebe C. Ellsworth, Race in the Court-
Our Experimental Implicit-Bias Jury Instruction, Annotated

Our system of justice depends on the willingness and ability of judges like me and jurors like you to make careful and fair decisions. What we are asked to do is difficult because of a universal challenge: We all have biases. We each make assumptions and have our own stereotypes, prejudices, and fears. These biases can influence how we categorize the information we take in. They can influence the evidence we see and hear, and how we perceive a person or a situation. They can affect the evidence we remember and how we remember it. And they can influence the “gut feelings” and conclusions we form about people and events. When we are aware of these biases, we can at least try to fight them. But we are often not aware that they exist.

We can only correct for hidden biases when we recognize them and how they affect us. For this reason, you are encouraged to thoroughly and carefully examine your decision-making process to ensure that the conclusions you draw are a fair reflection of the law and the evidence. Please examine your reasoning for possible bias by reconsidering your first impressions of the people and evidence in this case. Is it easier to believe statements or evidence when presented by people who are more like you? If you or the people involved in this case were from different backgrounds—richer or poorer, more or less educated, older or younger, or of a different gender, race, religion, or sexual orientation—would you still view them, and the evidence, the same way?

Please also listen to the other jurors during deliberations, who may be from different backgrounds and who will be viewing this case in light of their own insights, assumptions, and even biases. Listening to different perspectives may help you to better identify the possible effects these hidden biases may have on decision-making.

Our system of justice relies on each of us to contribute toward a fair and informed verdict in this case. Working together, we can reach a fair result.

Footnotes
1. We reprint here the specialized jury instruction used in the present experiment, along with citations to the existing scientific research evidence that formed the theoretical basis for each of the instruction’s components. However, we are not suggesting nor do we recommend that any court proceed to adopt this instruction without further empirical testing to determine the efficacy of this or any other instruction designed to address implicit and explicit forms of bias in juror decision making.
2. When leadership sets an egalitarian example, others may also pursue this goal (see Henk Aarts et al., Goal Contagion: Perceiving Is for Pursuing, 87 J. PERSONALITY & SOC. PSYCHOL. 23 (2004)).
3. To avoid potential backfire effects, instructional language should reduce external pressure to comply (by avoiding authoritarian language) and promote intrinsic motivation to counteract biases (E. Ashby Plant & Patricia G. Devine, Responses to Other-Imposed Pro-Black Pressure: Acceptance or Backlash?, 37 J. EXPERIMENTAL SOC. PSYCHOL. 486 (2001); Jennifer A. Richeson & Richard J. Nussbaum, The Impact of Multiculturalism Versus Color-Blindness on Racial Bias, 40 J. EXPERIMENTAL SOC. PSYCHOL. 417 (2004); Lisa Legault et al., Ironic Effects of Antiprejudice Messages: How Motivational Interventions Can Reduce (But Also Increase) Prejudice, 22 PSYCHOL. SCIN. 1472 (2011)).
6. People often are not aware of their own biases. For people to attempt to correct for bias, they must know that it is a problem and also believe the problem to be self-relevant (see Timothy D. Wilson & Nancy Brekke, Mental Contamination and Mental Correction: Unwanted Influences on Judgments and Evaluations, 116 PSYCHOL. BULL. 117 (1994); see also Tuane T. Wegener et al., Flexible Corrections of Juror Judgments: Implications for Jury Instructions, 6 PSYCHOL. PUB. POLY & L. 629 (2000); Duane T. Wegener & Richard E. Petty, Flexible Correction Processes in Social Judgment: The Role of Naive Theories in Corrections for Perceived Bias, 68 J. PERSONALITY & SOC. PSYCHOL. 36 (1995); Duane T. Wegener & Richard E. Petty, The Flexible Connection Model: The Role of Naive Theories in Bias Correction, in 29 ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY 141 (Mark P. Zanna ed., 1997)).
7. A more deliberative mode of thinking may help to reduce expressions of bias (see Ellen Langer et al., Decreasing Prejudice by Increasing Discrimination, 49 J. PERSONALITY & SOC. PSYCHOL. 113 (1985); Maja Dijkic et al., Reducing Stereotyping Through Mindfulness: Effects on Automatic Stereotype-Activated Behaviors, 15 J. ADULT DEV. 106 (2008); see also Guthrie et al., supra note 4; Jeffrey E. Pfeifer & James R. P. Ogloff, Ambiguity and Guilt Determinations: A Modern Racism Perspective, 21 J. APPLIED SOC. PSYCHOL. 1713 (1991)).
8. For a discussion on processing fluency and perceptions of trust, see Rolf Reber & Norbert Schwarz, Effects of Perceptual Fluency on Judgments of Truth, 8 CONSCIOUSNESS & COGNITION 338 (1999); Adam L. Alter & Daniel M. Oppenheimer, Uniting the Tribes of Fluency to Form a Metacognitive Nation, 13 PERSONALITY & SOC. PSYCHOL. REV. (2009). See also Russell D. Clark & Anne Maass, Social Categorization in Minority Influence: The Case of Homosexuality, 18 EUR. J. SOC. PSYCHOL. 347 (1988), Masaki
than white defendants. Authors of those studies concluded that the race of the victim may not be a critical factor in the expression of this “white juror bias,” but other research suggests otherwise. To account for this possibility, NCSC varied the race and the defendant across experimental conditions of the trial scenario. NCSC also created a video of a trial judge giving instructions on the applicable law, which included either the specialized implicit-bias jury instruction or an alternative instruction of approximately equal length (creating matched control groups for comparison purposes).

NCSC hired a market-research firm to recruit a sample of jury-eligible mock jurors to participate in the study. Recruited participants who met eligibility requirements received a brief description of the study in which they were asked to assume the role of a juror in a trial case. Participants were randomly assigned to one of eight possible conditions in the experiment: They watched one of the two videotaped sets of jury instructions (with either the specialized implicit-bias instruction or the control instruction imbedded) and then read one of four possible versions of the mock-trial scenario describing the evidence in the case against the defendant (varying the race of the defendant and the victim). After the presentation of evidence, mock jurors indicated whether they thought the defendant was guilty and, if so, recommended a sentence. The mock jurors also took the Race Implicit Association Test (IAT), a popular online test developed by researchers to identify, measure, and study implicit bias. A total of 901 jury-eligible adults participated in the study, which was conducted in May and June 2013. On the whole, the composition of the participant group reflected the demographic and attitudinal characteristics of the broader national population. A large majority of participants demonstrated a preference for whites on the Race IAT, which is also consistent with other national studies.

Overall, white participants across all conditions in the present study convicted white defendants at a slightly higher rate (65%) than black defendants (59%), although this difference was not statistically significant. The specialized instruction did not appear to significantly influence mock-juror verdict preference, confidence in verdict, or sentence severity. The authors were unable to replicate with this sample the traditional baseline pattern of “white juror bias” (i.e., the higher rate of guilty verdicts and harsher sentences for black defendants in control conditions) observed in prior similar studies, which precluded a complete test of the value of the specialized instruction. Because prior studies demonstrated or replicated the juror-bias effect successfully in a number of different settings, with a number of different types of trial scenarios, and with designs that varied in ecological validity (i.e., degree of resemblance to natural court settings), it is not likely that the findings of the present study are attributable to the web-based nature of the study design. It is possible that the specific legal instructions provided in the present study differed in a meaningful way from past studies and that those differences were ultimately responsible for eliminating the juror-bias effect. However, it is unlikely that these differences are the primary reason why the juror-bias effect was not observed, as in the months following this study, we learned that other contemporaneous studies in which similar legal instructions were not provided also failed to replicate
the original juror-bias effect. One intriguing potential explanation for the failure to replicate the juror-bias effect in the present study and in contemporaneous studies is that over time and with increased media scrutiny to racial inequality, Americans may have become increasingly aware of implicit forms of bias. They may be particularly sensitive in research and other settings in which they know or suspect that their individual responses are monitored for analysis, and this sensitivity may have been heightened at the time the study was conducted (i.e., during the Florida trial of George Zimmerman). This heightened level of awareness and sensitivity may have prompted many participants to spontaneously self-correct for possible expressions of racial bias, regardless of whether or not they received the specialized jury instruction. Future research should explore this possibility and its implications for contemporary jurors.

Despite the absence of a baseline race-bias effect for jurors, the study provided some preliminary evidence to suggest that a specialized instruction could alter expressions of bias in juror judgments. White jurors who received these specialized instructions produced a different pattern of judgments of the strength of the defendant’s case compared with participants who received a control instruction. Specifically, in control conditions, white jurors perceived the defendant’s case as significantly stronger when the alleged crime occurred between a black defendant and a black victim, compared with the scenario that involved a white defendant and a black victim. This difference was eliminated in the specialized-instruction conditions. Further research is needed to fully examine the impact of such an instruction under a variety of conditions. Additional research could also explore why, absent any bias-reduction intervention, the black-on-black crime in the present study produced the highest strength-of-case ratings for the defense. A complementary effect was not observed in favor of white defendants when the victim was described as white, discounting a same-race explanation for the effect. Finally, we also did not observe any clear evidence of “backlash effects” (in which mock jurors might seem to treat black defendants more harshly) after hearing an implicit-bias instruction, but small sample sizes limited these analyses. Future studies that continue to explore the potential utility of this type of instructional intervention should also be designed to answer this question.

THE VIEW FROM HERE

In this article, we have stressed the importance of empirically testing any new bias-reduction intervention for efficacy before full-scale adoption and implementation. While we have provided the specialized jury instruction used in this experiment, along with citations to the existing scientific research evidence that formed the theoretical basis for each of the instruction’s components, we are not suggesting that any court proceed at this point simply to adopt this instruction and move on: We do not have sufficient data at this time to support a recommendation to use this or any other specialized jury instruction to mitigate juror implicit bias. Based on the results of one study, the specialized jury instruction designed to address implicit and explicit forms of bias in juror decision making does not appear to be the panacea some hoped for. However, the research evidence continues to expand on a variety of strategies for reducing bias in decision making. Social scientists continue in earnest to search for and test innovative bias-reduction interventions. New evidence now exists to demonstrate the utility of some bias-reduction interventions that, at the time this study was conducted, were considered only theoretically promising. Many more strategies continue to show promise but have not yet received empirical scrutiny. As basic research evidence builds on more innovative approaches to addressing bias in decision making, the court community will benefit in time.

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 Master’s in Public Policy from the College of William and Mary.