Award Number: 2017-IJ-CX-0044

Can Jury Instructions Have an Impact on Trial Outcomes?

Mona Lynch (Principal Investigator)
Department of Criminology, Law, and Society
2340 Social Ecology II
University of California, Irvine
Irvine CA 92697-7080
(949) 824-0047
lynchm@uci.edu

Emily Shaw (Graduate Research Assistant)
Department of Psychological Science
University of California, Irvine
Irvine CA 92697-7080
evshaw@uci.edu

Award Recipient Organizational Address:
The Regents of the University of California
Irvine Office of Research
5171 California Avenue, Suite 150
Irvine, CA 92697-7600

Project period: 01/01/2018-07/31/2020
Award amount: $489,085

This project was supported by Grant No. 2017- IJ-CX-0044 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the US Department of Justice.

Summary of the Project

Major Goals and Objectives

As a long line of legal precedent has established, a defendant's 6th amendment right to a fair trial depends upon the jury being free from racial bias in its decision-making. Most recently, in *Rodriguez-Pena v. Colorado* (2017: 1), Justice Kennedy asserted that the taint of racial bias in the jury's functioning has consequences well beyond a given case or courtroom since "[t]he jury is a central foundation of our justice system and our democracy." The opinion suggested that various jury procedures, including jury instructions that admonish jurors that bias has no place in their decision-making, can function as a check on discriminatory outcomes. To that end, judges around the nation have, in recent years, begun to rely on "implicit bias" jury instructions to help curtail the influence of individual jurors' bias on the group decision-making process.

Despite this development, only scant research has examined the efficacy of implicit bias instructions. The first major study testing their impact failed to find any bias as a function of defendant race, nor did it find any specific effect of the implicit bias instruction (Elek & Hannaford-Agor, 2014). That study, however, used individual mock jurors who read written materials about the case in an online format. Participants rendered individual judgments so the findings were not able to address how such instructions influence group-level deliberation processes. The only other study to address this issue was conducted by Ingriselli (2015), who used an "egalitarian" instruction that explicitly educated participants about implicit bias as one of four instructions conditions in an experiment that again used individual participants who considered a criminal case. She found that those participants who were assessed to be "aversive" racists, who were therefore motivated to appear nonracist, were less likely to find the Black

defendant guilty when they heard the "egalitarian" instruction than when they heard an instruction that focused procedural justice (Ingriselli, 2015).

Because prior research by the Principal Investigator (PI) suggests that racial bias may get activated and/or become more pronounced *through* the group deliberation process (Lynch & Haney, 2009; 2011; 2015), the present study was designed as a more externally valid test of implicit bias instructions that incorporates in-person group deliberation. It also uses audio-visual presentation of the trial that better captures the way evidence, arguments, and jury instructions are presented to juries. Specifically, the project used a 2 x 2 x 2 between-subjects factorial design to examine the influence of defendant's race (Black or White), the government informant-witness's race (Black or White), and the presence or absence of "implicit bias" jury instructions (the control is the standard instructions), as delivered by the judge, on verdict outcomes and assessment of the evidence and testimony.

Jury-eligible participants were assigned to one of 120 small groups that viewed one of the eight trial versions, deliberated to verdict, then individually completed numerous measures to assess their judgments of the case and testimony, their understanding of the jury instructions, their general attitudes about a number of issues, and their basic demographics. This project is an extension of the PI's previous work funded by the National Science Foundation on jury decision-making in drug trafficking cases, which include group deliberations to improve external validity and better measure how decisions are actually made.[1] That study assessed the impact of race on outcomes; the present project revised the trial testimony to clarify the evidence and added the experimental manipulation of jury instructions so that jury groups either heard standard federal

[1] See https://www.nsf.gov/awardsearch/showAward?AWD_ID=1624943

instructions or implicit bias instructions derived from those instructions used by the Western District of Washington.[2]

Research Questions

This study tested two main research questions: (1) Does the race of the defendant and a key witness impact verdict outcomes in a federal drug conspiracy case scenario? and (2) Does the administration of implicit bias instructions impact that judgment process?

The specific hypotheses that are derived from these broader questions are as follows:

1) Race of defendant and race of informant will only influence verdict outcomes in the standard jury instructions conditions.
2) Consistent with prior research demonstrating bias against minority defendants, the Black defendant will be more likely to be found guilty than the White defendant across informant conditions in the standard instruction conditions.
3) Consistent with theory regarding race and credibility (Pager, 2005), the credibility ratings for the White informant will be higher than for the Black informant across defendant conditions in the standard instruction conditions.
4) Consistent with prior research suggesting that same-race pairings of defendant and witnesses increased plausibility of testimony (e.g., Bottoms, et al., 2004 where the victim was witness), same-race defendant-informant pairings will increase credibility ratings, thereby increasing the relative likelihood of guilt findings, compared to different-race pairings in the standard instruction conditions. Within that, the Black pairings will result in the highest guilt findings in the standard instructions conditions.

[2] Available at: https://www.wawd.uscourts.gov/sites/wawd/files/CriminalJuryInstructions-ImplicitBias.pdf

5) Consistent with prior research by the PI that found the deliberation process increased biased outcomes (Lynch & Haney, 2009; 2011; 2015), the group deliberation process will heighten observed race effects, as measured by difference scores between pre-deliberation and post-deliberation verdict and credibility measures in the standard instruction conditions.
6) Consistent with previous work (Lynch & Haney, 2011), the effects of race of defendant will be mediated by participant race at both the individual and group level, such that Whites individually, and White majority groups will demonstrate racial bias against the Black defendant in the standard instruction conditions.

Research Design, Methods, Analytical and Data Analysis Techniques

Research Design and Materials. The study employed a 2 x 2 x 2 factorial design, in which the race of defendant (Black or White), race of informant witness (Black or White), and jury instructions (implicit bias or standard bias instruction) was varied, creating eight experimental conditions (see Figure 1). To better approximate an actual jury trial, the case was presented as a 70-minute voice-recorded and visual trial presentation. The voice recording was completed using actors trained to play each of the speaking roles (prosecutor, defense attorney, FBI agent, informant, judge). The voice recording was then overlaid on a digital capture of a quick-moving slide show of 366 photographs representing the trial. The versions of that slide show varied only on the racial characteristics of the defendant and informant, as well as the instructions given to capture the eight experimental conditions, otherwise they were identical.

Defendant Harold Williams

Testifying informant Sheldon Smith

Figure 1. *Example images across conditions for race manipulations.*

The trial presentation was loosely based on an actual federal narcotics conspiracy trial transcript that involved the testimony of an FBI agent who worked with an informant to make a drug conspiracy case. In the trial presented to our participants, the defendant in the case, Harold Williams, was charged with conspiracy to distribute cocaine in the Central District of California. The government alleged that Mr. Williams entered into an agreement to sell 500 grams of cocaine to an associate, Sheldon Smith, for $10,000. Unbeknownst to the defendant, the associate himself had been arrested and was operating as an undercover informant for the FBI. The trial presentation featured opening statements from the prosecutor and defense attorney, followed by the testimony of two prosecution witnesses.

The first witness was an FBI agent who testified about her experience working in narcotics and managing informants, then about specifics of the case. She testified about text messages the defendant sent to the informant to arrange the drug sale, and how he was ultimately

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

arrested with marked cash from the informant. The second witness was the informant, Sheldon Smith, who testified about his prior history selling drugs for the defendant and about his plan to acquire drugs from the defendant as part of this sting operation. Both witnesses were subjected to cross-examination by the defense after their direct examination testimony.

The defense strategy was to create reasonable doubt through the cross-examination of those two witnesses. The defense attorney questioned the competence and motivation of the FBI agent who had not sufficiently documented the planning between the informant and defendant or the exchange of drugs and cash, and he challenged the informant's credibility on the grounds that the informant had prior criminal convictions, a history of lying to law enforcement, and was expecting a reduced sentence in exchange for his testimony. The defendant did not testify.

After the witness testimony concluded, the judge read the relevant jury instructions. These instructions were drawn from the actual case transcript and reflect the pattern instructions used in federal criminal trials. Specifically, the instructions explain the duties of jurors, the elements of the specific offense charged, the burden of proof and the presumption of innocence, considerations of witness credibility, and the defendant's right not to testify. The instructional manipulation was in the first section of the instructions, explaining the duties of the juror, where they are typically offered. The standard instruction we used was the federal criminal jury pattern instruction 1.1 that includes shorter, more generic language to avoid bias and prejudice (Model Criminal Jury Instructions, Ninth Circuit, 2010), and the implicit bias version was adapted from those developed and used by judges in the Western District of Washington (Doyle, 2017).[3] The trial concluded with closing arguments by the prosecution and defense.

[3] These instructions are sometimes offered at different points in the trial process, however for internal validity reasons we included all instructions in both the standard and implicit conditions at the end of evidence, prior to closing arguments when instructions are normally presented in federal court. The specific implicit bias instructions we included are: "You must decide the case solely on the evidence and the law before you and must not be

Participants and Procedures. To facilitate recruiting participants who were eligible to serve on a federal jury within the district, the study was conducted at a simulated "jury room" suite in a centrally located office building within the Central District of California. Sessions were available at multiple times of the day and days of the week, and the conditions were randomized to ensure there were no time or day-of-the-week confounds with the experimental conditions. Participants were recruited through a multi-pronged outreach strategy. This approach featured the use of business cards advertising the study in the jury assembly room at a local county courthouse, allowing us to directly recruit individuals who showed up to jury service in the area. The advertising business cards were also placed in local businesses and recreational areas in the region. Additional display and classified advertisements were placed in local community newspapers, and on online electronic boards.

Potential participants were screened for jury eligibility by phone; those deemed eligible were then scheduled to participate in an upcoming session. In order to be eligible to participate, interested persons needed to meet the criteria to serve as a federal juror. Thus, eligible participants were U.S. citizens of at least 18 years of age; adequately proficient in English to understand the proceedings and read English; and not currently facing felony charges, nor previously convicted of a felony (without their rights restored). In keeping with federal court procedure, the trial presentation was in English, and all printed materials and measures were in the English language.

---

influenced by any personal likes or dislikes, opinions, prejudices, sympathy, or biases, including unconscious bias. Unconscious biases are stereotypes, attitudes, or preferences that people may consciously reject but may be expressed without conscious awareness, control, or intention. Like conscious bias, unconscious bias, too, can affect how we evaluate information and make decisions. It is important that you discharge your duties without discrimination, meaning that bias regarding the race, color, religious beliefs, national origin, sexual orientation, gender identity, or gender of the defendant, witnesses, lawyers should play no part in the exercise of your judgment."

A total of 639 eligible participants successfully completed the study and were assigned to a total of 123 jury groups that were scheduled over an 8-month period of data collection. An additional 13 participants arrived at the site for scheduled sessions, but because the requisite number of "jurors" were not present to be able to run the session, these 13 participants were paid, dismissed, and invited to reschedule. Participants were each randomly assigned to a small jury group consisting of four to seven individuals. While federal criminal juries are comprised of twelve individuals, the smaller juries allowed for a sufficient number of jury units to meaningfully analyze, while not sacrificing a group decision-making process (Lynch & Haney, 2015; Sommers, 2006).

Data from three groups that completed the procedure, totaling 16 participants, were removed from analyses: two groups were removed due to technical difficulties during data collection rendering the data unreliable, and one group was removed due to the dismissal of a participant for disruptive behavior, resulting in too few participants remaining to comprise a jury unit of at least four people. Finally, data for one participant from a group of 7 mock jurors was removed due to the participant rushing through the individual measures post-deliberation, and providing nonsensical responses. The analyses are thus based on data from 623 participants and 120 jury groups.

A majority of the participants identified as women (62%) compared to men (38%), and they ranged in age from 18 to 88 years old, with a mean age of 42 years. The majority of participants self-identified as White (56%), while 44% identified as non-White. Specifically, 13% identified as Asian, 13% as Latinx, 6% as Black/African-American, and roughly 12% as a different race or ethnicity. About 32% of participants self-identified as Democrat, and 23% as Republican, with the remainder indicating a different political affiliation or none altogether.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Regarding political ideology, the group was relatively balanced, in that 23% of participants identified as conservative, 41% as moderate, and 27% as liberal. Approximately 17% of participants had served on a jury prior to study participation.

When recruited participants arrived at the study site, they were seated in a controlled room with a large video screen and seven chairs around a table. They were supplied with a study information sheet that explained the study and their rights as participants. The researcher then described the study procedure and obtained verbal consent from the participants that they understood what the study entailed and agreed to participate. Next, the 70-minute trial video was presented to the group, which was described as an actual trial that had taken place in the Central District of California. After viewing the trial video, participants each filled out a paper "straw" vote form privately, indicating their personal verdict preference (i.e., guilty or not guilty) and their confidence in that verdict on a five-point scale. This "straw" vote was confidential and nonbinding, and jurors were told they could amend their verdict preference at any time.

After a short break, participants conducted deliberations as a group. Juries were instructed to choose a foreperson who would record the group's verdict, and to deliberate to reach a unanimous verdict—a dichotomous choice of guilty or not guilty—on the single count of conspiracy to sell cocaine. Each jury was provided with copies of the judicial instructions, identical to the ones read by the judge earlier during the trial video.

The groups were left in the closed "jury room" to deliberate and a video camera was used to record these deliberations. Juries were given a maximum of 90 minutes to deliberate; once they reached a unanimous verdict the foreperson was provided a verdict form to record the group members' verdict and jurors' individual confidence in the verdict decision, using the same five-point scale. Failure to reach a unanimous verdict within the 90-minute time limit resulted in a

mistrial, and the nature of the split was recorded on a separate "mistrial voting form," which included individual verdict choices for each juror and the respective five-point scale confidence rating. Some groups declared mistrials prior to the 90-minute time limit, when they were certain they absolutely could not reach a unanimous decision. Approximately 19% of deliberations ended in mistrials. The mean deliberation time was approximately 29 minutes and 14 seconds; those reaching unanimous verdicts had significantly shorter deliberations (24 minutes, 43 seconds) than did those ending in mistrials, which lasted just under 48 minutes on average ($t = -7.75$, Cohen's $d = 1.82$, $p < .0001$).

Although the experiment was a simulated trial, thus reducing its external validity, deliberations were often highly intense, and participants were engaged, expressive, and many defended their positions actively and passionately, highlighting the study's high level of "experimental realism" (Lynch & Haney, 2015). Very few participants commented about being recorded and/or speculated about whether the trial they watched was fictitious; even when such comments were made, other group members spoke up to direct the group back to the task of reaching a verdict.

Following deliberations and documentation of the group's verdict, participants individually completed an electronic survey using laptops provided by the researchers. The survey assessed individual perceptions of the case, evidence, witnesses, attorneys, defendant, and judge; comprehension of jury instructions; influence of jury instructions; memory of case facts; attitudes about a variety of issues; and demographic information. We included a series of scales that measured subtle racial bias among the attitude measures. These scales were drawn from Williams and Eberhart's (2008) Race Conceptions Scale, and from the Subtle Racism Scale (Meertens & Pettigrew, 1997). After participants completed the individual measures, they were

debriefed and paid $100. The study took approximately three-and-a-half to four hours for participants to complete.

Analytic Strategy. The quantitative data was imported from Qualtrics to both Stata and SPSS files, then cleaned and analyzed in both programs. We ran a series of quantitative analyses to see how the instructional condition (implicit bias vs. standard) impacted verdicts and juror comprehension, and how the instructions conditions shaped how jurors thought about their duties as jurors. We also tested our multiple hypotheses on the race impact on outcomes. To do so, we use multiple regression, logistic regression, t-tests, and chi-square analyses, depending upon the analysis and data source.

To prepare for the analysis of the video-recorded deliberations, all recordings were transcribed by a professional transcription service. Two jury groups were not recorded due to experimenter error, so we transcribed 118 deliberations. The transcripts were systematically coded and analyzed by a team of two graduate and seven undergraduate research assistants to evaluate how manipulated aspects of the case (i.e., instructions and race) shaped deliberations. A mixed strategy of concept-driven and data-driven qualitative content analysis (Schreier, 2012), as well as more traditional manifest content analysis (Krippendorf, 2012), allowed for characterizations of how jurors discuss the instructions, informant, and defendant on case-relevant matter, and non-case-relevant matters. The transcriptions were coded for a number of instructions-related and bias-related content, including specific references to the instructions, juror understandings/misunderstandings of instructions, discussions of bias and avoiding bias, and the evaluation of specific evidence and testimony.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Expected Applicability of the Research

Because this research was designed to maximize external validity and obtain robust experimental realism (Weiner, Krauss, & Lieberman, 2011), it has heightened applicability to the real-world conditions of criminal trials. It used highly engaging and legally-accurate stimulus materials derived from a federal trial transcript from an actual federal drug conspiracy case. The use of screened, jury-eligible, non-student participants also increased its external validity, as did the creation of a "jury room" away from the university campus. Most significantly, the research used small groups that deliberated as actual juries do, which is critical for understanding the processes underlying race-based judgment and action by juries (see Devine, 2012 for a discussion of both the rarity and importance of this feature). Given the study's high quality, it has the potential to positively impact our knowledge base and justice policy-making in several key ways.

First and foremost, the research sheds light on whether observed racial disparities in trial outcomes can be attenuated through the use of specific instructions designed to mitigate implicit bias. This is the first study in this area of inquiry to include group-level deliberation (still a rarity across jury studies generally) at a scale that allows for robust, between-groups analyses. Consequently, there are numerous new insights into the impact of implicit bias instructions on jury judgment that we can examine. It also adds to the broader literature as to when jury instructions can be effective in aiding jury judgment (Pfeifer & Ogloff, 1991), and when they are not (Lieberman & Arndt, 2000; Steblay, Hosch, Culhane, & McWethy, 2006).

The utility and applicability of these findings also extends beyond juries. More and more justice agencies are relying upon implicit bias education programs as a remedy to racially disparate outcomes (Smith, 2015). Moreover, a diverse array of organizations and institutions have become attuned to the problem of implicit bias, including schools and universities (Gullo,

Capatosto, & Staats, 2018; Jackson, Hillard, & Schneider, 2014), private industry (Bertrand & Mullainathan, 2004), governmental agencies (Beniwal, 2016; Foley & Williamson, 2019), health systems (FitzGerald & Hurst, 2017; Zestcott, Blair, & Stone, 2016), and legal institutions (Kang, et. al, 2012; Negowetti, 2014), and have instituted various forms of training and remediation to mitigate the potential impact of implicit bias (Burns, Monteith, & Parker, 2017; Casey et al., 2013; Lai, et al., 2014). There is reason to question how effective such programs can be without also addressing contextual and structural conditions that may contribute to bias in organizational settings. Because this work uses small group decision-making as a vehicle for studying bias, it sheds light on how interactive, social processes are influenced, or not, by educationally-based directives on implicit bias.

This research also adds to the small but growing body of research on jury composition and judgment (Lynch & Haney, 2011; 2015; Peter-Hagene, 2019; Shaw, Lynch, Laguna, & Frenda, under review; Sommers, 2006) that address the effects of jury composition for Black defendants, an issue that the U.S. Supreme Court has recently wrestled with in *Foster v. Chatman* (2016). In that case, the prosecutor's exclusion of all Black potential jurors in a capital case was deemed a violation of the Black defendant's 6th and 14th amendment rights, in part because the Court recognized the potential for racial bias that comes with non-diverse juries.

This study's findings also have the potential to inform policy regarding the use of informants, and whether the documented problems with informant evidence are exacerbated by the racial identity of those witnesses. As such, it has the potential to address real-world procedural issues that come from the heavy reliance upon informants in drug case prosecutions. While there is a notable body of work on how defendant characteristics influence case processing and decision-making, there is significantly less empirical work on how witness' race influence

guilt determinations, with much of that work addressing judgments of eyewitness testimony (e.g., Abshire & Bornstein, 2003) and character witness testimony (e.g., Maeder & Hunt, 2011).

The only prior experimental studies examining race and informants are by Maeder and Yamamoto, (2017) and Shaw, Lynch, Laguna, and Frenda (under review). This research also adds to the other important work on informants that indicates laypersons are not able to appropriately gauge how incentives provided to informants may erode credibility (Neuschatz, Lawson, Swanner, Meissner, & Neuschatz, 2008; Neuschatz, Wilkinson, Goodsell, Wetmore, Quinlivan, & Jones, 2012; Wetmore, Neuschatz, & Gronlund, 2014).

Participants and Other Collaborating Organizations

This project did not involve other collaborating organizations or participants beyond the Principal Investigator, a staff research associate during data collection, four graduate student researchers who served at different points in the research process, and a total of 13 undergraduate research assistants at UC Irvine.

Changes in Approach

This project did not require any changes in approach between the study as proposed and the study that was conducted.

Outcomes

Activities / Accomplishments

In this study, 639 jury-eligible adults from the federal jurisdiction of interest (the Central District of California) were recruited, completed the study, and were compensated. Six undergraduate research assistants were trained in the collection of data and proper subject

protocols who then assisted in running the jury groups. Undergraduate research assistants worked in pairs and were supervised by a graduate research assistant and a staff research associate who had previous experience in data collection on a similar mock jury experiment for the PI. Seven additional undergraduate research assistants were trained and supervised in content analysis and coding by two graduate research assistants.

Graduate students funded by the project were able to build professional experience with experimental design, data collection in a complex, off-campus research design, data analysis, and academic writing. They were mentored by the PI in drafting article manuscripts for submission to peer-reviewed journals, presenting research findings at conferences (prior to the COVID-19 disruption), and submitting grant proposals for their own research. Mentees submitted dissertation proposals to the Ford Foundation, Haynes Foundation, National Science Foundation, and the National Institute of Justice.

Results and Findings

Summary of Results. Generally, our hypotheses were not supported in this study. Our quantitative findings indicated no differences in verdict outcomes between those who heard the implicit bias instructions compared to those who heard the standard instructions, either as a main effect or as function of the race conditions. We also found no main effect for race of defendant in our straw vote measures, and no main effect for the informant race. Our hypothesis about the interaction of defendant and informant race was not confirmed, and instead we found the opposite of what we expected: different-race pairings produced more guilty verdicts than did same-race pairings although the results were only marginally significant. On our juror/jury composition-defendant race hypothesis, we found that while White jurors were more likely to

favor a conviction, the proportion of White jurors in a group did not translate to a statistically significant impact at the group level.

In this section, we describe the following specific results:

- Jurors in the implicit bias condition were more likely to consider "being unbiased" the most important duty of a juror, compared to those in the standard instructions condition.
- The instructions did not directly impact verdict at the individual pre-deliberation or group (post-deliberation) level. The instructions also did not significantly interact with either of the race manipulations in predicting verdict outcomes.
- We found that at the individual pre-deliberation level, defendant race and informant race had only a marginally significant interaction ($p = .053$) which contrasted with predictions and indicated that the defendant was most likely to be considered guilty (pre-deliberation) when the informant testifying against him was of a different race.
- White jurors were significantly more likely to prefer a conviction going into deliberations compared to non-White jurors, but we did not find a significant effect of juror race on post-deliberation convictions at the group level.
- We did not find evidence to support the idea that the group deliberation process would amplify the race effects we originally predicted.
- A content analysis of deliberations suggests that jurors used implicit bias instructions in a variety of ways. Some groups used the instructions to foster evidenced-based conversations, while others used the message against bias as a

rhetorical weapon against other jurors or as a tool to impede critical evaluations of witness credibility.

Details of Findings:

Instruction Manipulation and Juror Responses. This study included a manipulation of instructions, comparing the impact of instructions that warned jurors about unconscious or implicit bias to the standard warning about the need to avoid prejudice. An integral part of our hypotheses was the expectation that the implicit bias instructions would have a different impact on juror verdicts and evidence assessments compared to the standard instructions. In hypotheses 1, 2, 3, 4 and 6 (see full hypotheses on pp. 4-5), we predicted a difference in juror verdicts and credibility ratings specifically dependent on the instruction manipulation. However, the data did not support this finding across any of the hypotheses.

It is clear that jurors paid attention to the instructions and noticed bias-relevant portions. We know this because there was a modest detectable effect of these instructions: jurors were significantly more likely to indicate that "avoiding bias or prejudice" was the most important goal of a juror when they were in the implicit bias condition, compared to those in the standard instruction condition (Exp(B) = 1.44, 95% CI [1.01, 2.06], Wald = 4.11, p = .043).

Given the relative attentiveness to the implicit bias instruction, we hypothesized that this group may demonstrate better comprehension of the instructions than those in the standard instructions condition. Instructional comprehension was assessed post-deliberation (i.e., after jurors heard the instructions read out loud by the judge and had finished group deliberations, where they had been provided a written copy of the instructions). The comprehension measures tested whether participants correctly understood the elements of conspiracy to distribute

narcotics, the burden of proof, the juror's duties, what constitutes evidence, and the right of the defendant not to testify. The average comprehension score was 4.81 out of a possible 6 points (with 6 indicating complete instruction comprehension) meaning that participants exhibited a good understanding of these elements of the instructions. Those in the standard instructions condition had a mean score of 4.83, and those in the implicit bias condition obtained a mean score of 4.79. A t-test indicated that this difference in instructional comprehension was not significant ($t(621) = 0.4184$, $p = 0.67$).

The manipulated instructions also had no measurable impact *directly* on verdict preferences (pre- or post-deliberations) in this study. When testing for an impact of this manipulation using a binary logistic regression, there was no significant impact of instructions in predicting pre-deliberation straw poll verdicts, nor any significant interactions between instructions and the other two manipulations (all ps > .40). The same lack of effect was found in final group verdicts post-deliberation; jury groups were no more likely to reach a conviction in the implicit bias condition (56% conviction rate) compared to the standard instruction condition (60% conviction rate; p =.394). There was also no detectable interaction between the instruction manipulation and defendant race in predicting group convictions (p = .454). When the defendant was Black, 52% of groups given standard instructions found him guilty, and 50% of groups given implicit instructions did the same (See Figure 2 for individual condition conviction rates).

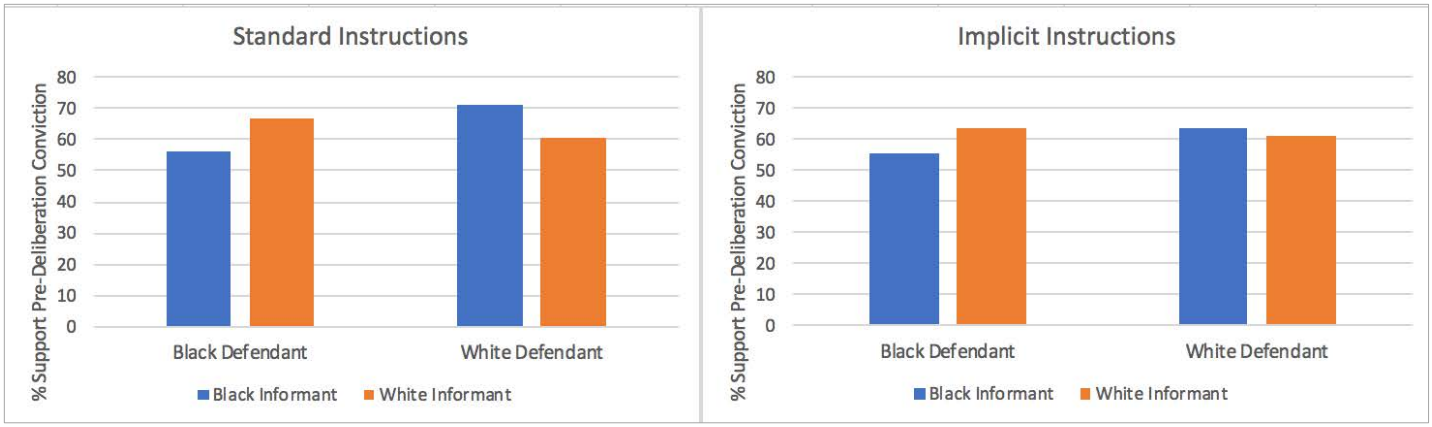

Figure 2. *Pre-deliberation support for convictions across all eight conditions*.

The same pattern of results was found for measures of informant credibility. We performed a three-way analysis of variance test (ANOVA)[4] to assess the impact of all three manipulations on ratings of informant credibility. The three-way interaction term was not significant (p = .995) and the instruction manipulation did not interact with defendant race (p = .866) or informant race (p = .811) to predict credibility ratings. The instruction manipulation also did not have any individual effect on credibility ratings (p = .403).

Together, these findings indicate that our instruction-related hypotheses were not supported by the data. Given the absence of differences in juror responses depending on the instruction manipulations, either directly or in interaction with other manipulations, hypotheses 1, 2, 3, 4 and 6 are unsupported. In light of these findings, for the remainder of the report, we present an adjusted set of tests and findings of interest that collapse the instructions conditions in order to examine the overall effects of race of defendant and informant.

Here, we present results that test the following research questions:

[4] Analysis of variance tests are frequently used for hypothesis testing in experiments. They can measure whether the average (mean) differences in outcomes between the different experimental conditions are statistically significant. The three-way analysis means that three independent variables were included. In this case, that is defendant race, informant race, and type of jury instruction.

A. Did defendant race and informant race interact to predict verdicts (adaptation of hypothesis 1 and 2) or credibility (adaptation of hypothesis 3 and 4)? Juror responses are examined at both the individual level and group level for verdict measures.
B. Did White jurors (individually or as a group) respond differently to the case compared to non-White jurors? Do White jurors show evidence of bias against the Black defendant, as measured by verdicts, or a generalized tendency to support conviction across conditions, relative to non-White jurors? (adaptation and expansion of hypothesis 6)
C. Did the group deliberation process amplify any observed defendant/informant race effects, as measured by difference scores between pre-deliberation and post-deliberation verdicts? (adaptation of hypothesis 5)
D. How did instructions shape juror behavior during deliberations, as measured via a content analysis? (new question)

Adapted Hypothesis A1 – Defendant and Informant Race Manipulations – Individual Pre-Deliberation Straw Poll. We performed a binary logistic regression[5] predicting pre-deliberation straw poll verdicts based on our race manipulations, the interaction between them, and a dichotomous juror whiteness variable. To create the juror whiteness variable, we dichotomized juror race to compare White jurors (n = 346) to members of all other racial / ethnic groups (n = 277) in the sample. The overall model was significant (χ^2 (4, N = 622) = 9.88, p =

[5] Binary logistic regression is a statistical method that can be used to see how one or more independent variables impact an outcome that has only two potential possibilities (a dichotomous variable). In this case, at the individual level, the only two possible responses are "guilty" or "not guilty" so it is an appropriate method to use. We could ask the question of whether the race conditions and/or the instruction condition produced differences in verdict outcomes.

.042), suggesting that at least one of the independent variables or a combination of them impacted verdict outcomes. As we detail in the section, **Adapted Hypothesis B – Juror Race Effects** (p. 24), there was a significant effect of juror race, such that non-White jurors were significantly less likely to prefer a guilty verdict compared to the White jurors.

In regard to the defendant and informant race conditions, there was a marginally significant interaction between defendant race and informant race (Exp(B) = 1.91, 95% CI [0.99, 3.67], Wald = 3.73, p = .053). The defendant was most likely to be considered guilty in the pre-deliberation poll when the informant was of a different race. For example, the conviction rate for the Black defendant was relatively low when the informant was Black (56%), but high when the informant was White (65% conviction rate). The opposite pattern was observed for the White defendant; when the informant was Black, the conviction rate was higher (68% convicting) compared to when the informant was also White (61%). Again, this interaction is marginally significant (p = .053) in this model. See Table 1.

Table 1. Percentage of Jurors Supporting a Pre-Deliberation Verdict of Guilty by Race Condition

Condition	Pre-Deliberation Conviction Support
Black Defendant & Black Informant	56%
Black Defendant & White Informant	65%
White Defendant & Black Informant	68%
White Defendant & White Informant	61%
Overall	62%

We were also interested in the impact of our manipulations and juror race on perceptions of informant credibility, which was measured on a scale from 1 (not at all credible) to 7 (completely credible). We performed a two-way ANOVA using defendant and informant race

(and the interaction between them) to test the hypothesis that the race conditions would impact how credible the informant was perceived to be.[6] The interaction between defendant and informant race was significant, $F(1, 622) = 16.00$, $p < .001$ and paralleled the trends observed in verdicts. Jurors rated the informant's credibility highest when the Black informant testified against a White defendant. In contrast, the White informant was rated as most credible when he testified against a Black defendant. The informant's race was marginally significant ($F(1, 622) = 3.44$, $p = .064$), indicating that our jurors rated the Black informant marginally more credible than the White informant.

Adapted Hypothesis A2: Defendant and Informant Race Manipulations – Group-Level Analyses. Participants were distributed among 120 deliberating jury groups. Of these groups, 56 convicted the defendant (47%), 41 acquitted him (34%), and 23 could not reach a unanimous verdict and were declared mistrials (19%). We performed a binary logistic regression at the group-level to predict final group acquittals, with our race manipulations and the interaction term between them as predictors. The model was significant (χ^2 (3, $N = 120$) = 8.08, $p = .044$). There was a significant interaction between defendant race and informant race in predicting acquittals (Exp(B) = 5.29, 95% CI [1.09, 25.75], Wald = 4.26, $p = .039$), such that the defendant was most likely to be acquitted when the informant testifying against him was the same race (see Table 2 for acquittal percentages in each condition). For example, 53% of groups who evaluated the Black defendant acquitted him when the informant testifying against him was also Black, but just 31% of groups acquitted him when the informant was White. The opposite pattern was observed for the White defendant; 19% of groups acquitted him when the informant

[6] Again, analysis of variance tests measure whether the average (mean) differences in outcomes between the different experimental conditions are statistically significant. The two-way analysis means that two independent variables were included. In this case, that is defendant race and informant race.

was Black, while 33% acquitted him when the informant was White. Again, this overall interaction was significant (p = .039).

Table 2. Final Group Verdicts by Race Condition

Condition	Convict	Acquit	Mistrial
Black Defendant & Black Informant	33%	53%	14%
Black Defendant & White Informant	55%	31%	14%
White Defendant & Black Informant	52%	19%	29%
White Defendant & White Informant	47%	33%	20%

Adapted Hypothesis B – Juror Race Effects. Here, we test two questions: Do White jurors disfavor the Black defendant, and do White jurors convict any defendant (regardless of race) at a different rate compared to non-White jurors?

Per our original hypothesis 6, we tested the possibility that White jurors would be more likely to support a guilty verdict against a Black defendant compared to a White defendant. We tested this question using a binary logistic regression to test for interactions between dichotomized juror race (White / non-White) and defendant race in predicting individual straw poll verdicts. The overall model was not significant (p = .106). Among White jurors, 68% supported a guilty verdict prior to deliberations for the White defendant (123 out of 181) compared to 65% for the Black defendant (106 out of 164).

We also tested the possibility that White jurors would be more likely to support a conviction for any defendant, regardless of defendant race, prior to deliberations compared to non-White jurors. In the binary logistic regression model described in A1 (which tested the impact of both race manipulations, the interaction between them and dichotomized juror race), there was a significant effect of juror whiteness (Exp(B) = 1.43, 95% CI [1.03, 1.99], Wald =

4.57, $p = .033$), such that jurors who were White were significantly more likely to prefer a verdict of guilty (66% conviction rate) compared to non-White jurors (57% conviction rate) in the pre-deliberation straw poll.

We also examined the impact of White jurors at the group level but found no significant effects. To test the possibility that groups with more White jurors were more likely to reach a conviction, we performed a binary logistic regression. Using the proportion of White jurors in the group to predict convictions, the overall model was not significant ($p = .208$). We also tested the possibility that defendant race would interact with the proportion of White jurors in groups to predict group convictions, but the overall model was also not significant ($p = .546$).

Adapted Hypothesis C – Group Deliberations and Race Effects. In our original hypothesis 5, we predicted that the group deliberation process would amplify the race effects we anticipated in hypotheses 1 through 4. For example, we expected that jurors would be more likely to switch their vote preference to favor a conviction for a Black defendant, but not for a White one. We also expected that jurors would be more likely to switch their vote to a conviction when the informant was White compared to when he was Black. (We also expected that the interaction between them would get stronger pre- to post).

However, our findings from earlier hypotheses did not support these predictions. As previously discussed, we did not find evidence of an anti-Black defendant bias at either the group or individual level. For reference, between the pre-deliberation straw poll and the final group verdict, 17% of jurors switched from a verdict preference of guilty to not guilty, 12% switched from a preference of not guilty to guilty, and 71% showed no change. Examining these trends by defendant race, we found that jurors were not more likely to switch votes to favor a conviction when the defendant was Black. Specifically, 11.4% of jurors viewing the Black defendant

changed to support a verdict of guilty, while 12.7% of jurors viewing the White defendant did the same. In contrast, among jurors who viewed the Black defendant, 20.5% switched to favor a *not* guilty verdict, while just 13.4% of jurors viewing the White defendant made that same switch. In other words, we found no evidence that the deliberation process brought out or amplified racial bias against the Black defendant. The same absence of an anti-Black effect was observed for jurors exposed to informants of different races, with no significant differences in juror verdict changes based on informant race condition.

Adapted Hypothesis D - Impact of Instructions on Deliberations: Content Analyses. We previously described how jurors were attuned to the messaging of the implicit bias instructions, evidenced by differences in how jurors identified the most important duty of a juror. However, we did not find a direct impact of instructions on verdicts, overall or as an interaction with the race conditions.

To further examine whether and how the implicit bias instructions influenced the group deliberation process, we first examined deliberations to see if groups were more likely to specifically reference *any* of the instructions in the implicit bias condition compared to the standard condition. The implicit bias groups deliberated for an average of 30 minutes, 22 seconds, whereas the standard instructions groups deliberated for a mean of 28 minutes, 9 seconds ($t = -.765$, *n.s.*). The coding for this measure included all direct references during the deliberations to the instructions generally, or any of their specific elements; reading directly from the instructions; or specific references to the judge's words (he only spoke when reading the instructions in our simulated trial).

The groups in the implicit bias condition were marginally more likely to discuss specific aspects of the jury instructions than those groups in the standard instruction condition.

Specifically, 73% of those in the standard instruction condition discussed the jury instructions at least once during deliberations, and 86% of the groups in the implicit bias condition discussed the instructions at least one time (Exp(B) = 0.44, 95% CI [.17, 1.13], Wald = 2.98, p = .087). Among the groups that did discuss the instructions, the instructions were referenced at roughly the same frequency across the two instruction conditions. Implicit bias juries had a mean of 3.74 distinct mentions of instructions per jury deliberation, while standard instructions juries had a mean of 3.18 mentions (t (116) = 1.02, Cohen's d = 0.21, p = .312).

We then coded explicit references to bias, prejudice, stereotyping, discrimination, and/or racism in the deliberations to test whether the implicit bias groups were more likely to discuss such topics. Those in the implicit bias condition were significantly more attuned to bias in their deliberations, confirming what our previous analysis suggested about how participants perceived the most important duty of jurors. Bias was raised as a discussion point twice as often in the implicit bias condition, coming up at least once during deliberations in 45% (26/58) of the juries in the implicit bias condition, compared to at least once in just 23% (14/60) of the groups in the standard instruction condition (Exp(B) = .375, 95% CI [.17, .826], Wald = 5.92, p = .007). The number of distinct mentions/discussions per group ranged from 1- 13, with a mean of 1.93 mentions within standard instructions groups and 2.22 in the implicit bias groups (*n.s.*).

We also performed a qualitative content analysis of juror deliberations, with detailed results presented in Lynch, Kidd, and Shaw (under review). Our analysis identifies several ways instructions are used. In some groups, we found that the implicit bias instructions appeared to be linked to evidence-based conversations. Some of the deliberations in the implicit bias condition also revealed how groups of strangers tasked with coming to a verdict decision can engage in non-defensive dialogue about the problem of bias, including their own, that can interfere with a

fair and just trial process. In that regard, the instructions sometimes catalyzed an educative process that also could aid in keeping a check on group members' bias.

However, the concept of bias was elastic enough in these deliberations to sometimes be used against criminal defendants, such as when mock jurors would interfere with their peers' lawful and appropriate assessments of witness credibility. Specifically, in some of the deliberations in the implicit-bias conditions, jurors accused others in the group of demonstrating bias against the informant and/or FBI agent when they raised concerns about whether those witnesses had reason to fabricate testimony or shade the truth. This kind of assessment was specifically authorized under the instructions all juries heard.[7] Some jurors also used the anti-bias instructions as rhetorical weapons to counter those they disagreed with, by charging others with being "biased."

Limitations

Although this study was able to provide a number of new insights into how implicit bias instructions influence and shape group-level processes during deliberations, the findings do not offer conclusive evidence as to whether the advantages of using implicit bias instructions outweigh potential costs. Additional research is needed to fully understand if implicit instructions can reduce biased decision-making in specific kinds of cases, among sub-populations of individuals, or using variations in the language of the instructions. More research

[7] The instruction in both conditions on this read, in part: "One of the prime jobs of the jury is to determine the credibility – that is, the believability – of the witnesses. You must decide whether these witnesses were telling the truth or not. A person can tell the truth in whole, in part, or not at all. So I encourage you to run through a mental checklist as you evaluate what you heard today. For example, ask yourself: Did the witness have a good opportunity to see what he or she testified about today? Does the witness have an accurate recollection of the events he or she is recounting? Does the witness have some interest in the outcome, something to gain or lose by what the jury decides? Does the witness have some bias towards one side or the other that might cause him or her to shade the truth? If you believe that the witness has something to gain, ask yourself whether this would make him or her more or less inclined to be truthful."

is also needed to further explore some of the potential complications that come with implicit bias instructions. While we did not specifically uncover any backfire effects, we did find that the elastic interpretation of bias during deliberation could be detrimental to the defendant, whose due process and equal protection rights are at stake in criminal trials. Moreover, given our small cell sizes for our group-level analyses, we recognize this study is under-powered to statistically detect modest effect sizes for our groups. Due to the time-intensive, high-cost nature of this kind of study, we had to sacrifice statistical power at the group level to be able to manipulate all three variables of interest. Future research might consider reducing the number of experimental conditions to increase cell sizes at the group level.

Another consideration is that this study only examined one type of crime–an arranged drug conspiracy that was set up by law enforcement– a case type that may have been especially likely to produce skepticism about both the FBI agent and informant, particularly in the Black defendant conditions. This particular scenario directly called on our participants to consider system actors' fairness and veracity in making a verdict determination, so it likely contributed to our findings. It also likely enhanced the centrality of testimonial credibility in determining verdict. It may be that different kinds of crimes, especially where there is substantial independent evidence, would prompt different responses from jurors and may be more likely to produce outcomes that disadvantage the Black defendant. Prior studies that use a case scenario involving violence have been more likely to produce anti-Black bias (Espinoza & Willis-Esqueda, 2015; Foley & Chamblin, 1982; ForsterLee, et al., 2006; Lynch & Haney, 2000, 2009), so case facts clearly matter in how race may play a role in juror judgment.

A third limitation is external validity with regard to consequentiality. Participants were aware they were not actually determining a true verdict for a real defendant during their

deliberations. However, this did not stop them from engaging actively and passionately in the deliberation process. It was not uncommon for the participants to argue heatedly over matters of fact and law in the case. And given the design we employed, using an audio-visual presentation of the case, recruiting jury-eligible, non-student adults, and including group deliberations, this study was considerably more realistic and higher quality than the majority of mock jury studies (see generally, Devine, 2012).

Finally, this study provides additional evidence of the role of the specific decision-making parameters and contextual factors in catalyzing or impeding biases (Axt, Ebersole, & Nosek, 2016; Devine, 1989; Dovidio, 2001; Kawakami, Dion, & Dovidio, 1998; Payne, Vuletich, & Lundberg, 2017). The particular facts of our case scenario may have elicited a sensitivity to, and recognition of, the problem of racially biased drug law enforcement (including how informants are used) among our participants that make them more skeptical of the evidence when the defendant is Black. It may be that implicit bias instructions would play a very different role in cases where deliberations can exacerbate bias (e.g., Lynch & Haney, 2015). To that end, additional research is warranted on how case features may produce or inhibit different forms of bias, including how implicit bias instructions intersect with those features.

Artifacts

List of Products

- Lynch, M., Kidd, T., & Shaw, E.V. The Subtle Effects of Implicit Bias Instructions. Revise & resubmit for *Law & Policy*.
- Lynch, M. & Shaw, E.V. (in preparation). The Role of Police Legitimacy and Perceived Fairness in Juror Judgments.

- Lynch, M. & Shaw, E.V. (in preparation). Impact of Race and Credibility on Jury Decision-Making.
- Shaw, E. V. & Lynch, M. (March, 2020). *Implicit Bias Instructions, Race and Jury Verdicts*. Paper presentation at the American Psychology-Law Society Conference, New Orleans, Louisiana.

Data Sets Generated

- Individual-Level Data Set. 623 individual juror responses.
- Group-Level Data Set. 120 group jury responses.

Dissemination Activities

To date, dissemination of this research has consisted of one submitted manuscript that is in revision for *Law & Policy,* and one conference presentation listed above. An additional planned conference poster presentation was canceled due to the COVID-19 pandemic.

References

Abshire, J., & Bornstein, B. H. (2003). Juror sensitivity to the cross-race effect. *Law & Human Behavior*, *27*, 471-480.

Axt, J. R., Ebersole, C. R., & Nosek, B. A. (2016). An unintentional, robust, and replicable pro-Black bias in social judgment. *Social Cognition, 34*(1), 1-39.

Beniwal, R. (2016). Implicit bias in child welfare: Overcoming intent. *Connecticut Law Review*, *49*, 1021-1067.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, *94*(4), 991-1013.

Bottoms, B. L., Davis, S. L., & Epstein, M. A. (2004). Effects of victim and defendant race on jurors' decisions in child sexual abuse cases. *Journal of Applied Social Psychology*, *34*(1), 1-33. https://doi.org/10.1111/j.1559-1816.2004.tb02535.x

Burns, M. D., Monteith, M. J., & Parker, L. R. (2017). Training away bias: The differential effects of counterstereotype training and self-regulation on stereotype activation and application. *Journal of Experimental Social Psychology*, *73*, 97-110.

Casey, P., Warren, R., Cheesman, F., Elek, J. (2013). Addressing implicit bias in the courts. *Court Review*, *49*, 64-70.

Devine, D. J. (2012). *Jury Decision-making: The State of the Science*. New York: NYU Press.

Devine, P. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality & Social Psychology, 56,* 5-18.

Dovidio, J. (2001). On the nature of contemporary prejudice: the third wave. *Journal of Social Issues, 57,* 821-850.

Doyle, T. (2017). U.S. District Court produces video, drafts jury instructions on implicit bias. Seattle, WA: King County Bar Association

Elek, J.K. & Hannaford-Agor, P. (2014). Can explicit instructions reduce expressions of implicit bias? New questions following a test of a specialized jury instruction (April 28, 2014). Available at SSRN: https://ssrn.com/abstract=2430438 or http://dx.doi.org/10.2139/ssrn.2430438

Espinoza, R. K., & Willis-Esqueda, C. (2015). The influence of mitigation evidence, ethnicity, and SES on death penalty decisions by European American and Latino venire persons. *Cultural Diversity & Ethnic Minority Psychology*, *21*, 288-299.

FitzGerald, C., & Hurst, S. (2017). Implicit bias in healthcare professionals: a systematic review. *BMC Medical Ethics*, *18*(1), 19. https://doi.org/10.1186/s12910-017-0179-8

Foley, L. A., & Chamblin, M. H. (1982). The effect of race and personality on mock jurors' decisions. *Journal of Psychology, 112,* 47–51.

Foley, M., & Williamson, S. (2019). Managerial perspectives on implicit bias, affirmative action, and merit. *Public Administration Review*, *79*(1), 35-45.

Forsterlee, R., Forsterlee, L., Horowitz, I. A., & King, E. (2006). The effects of defendant race, victim race, and juror gender on evidence processing in a murder trial. *Behavioral Sciences & the Law*, *24*, 179-198.

Foster v. Chatman, 136 S. Ct. 1737 (2016).

Gullo, G. L., Capatosto, K., & Staats, C. (2018). *Implicit Bias in Schools: A Practitioner's Guide*. Routledge.

Ingriselli, E. (2015). Mitigating jurors' racial biases: The effects of content and timing of jury instructions. *Yale Law Journal*, *124*, 1690-1745.

Jackson, S. M., Hillard, A. L., & Schneider, T. R. (2014). Using implicit bias training to improve attitudes toward women in STEM. *Social Psychology of Education*, *17*(3), 419-438.

Kang, J., Bennett, M., Carbado, D., Casey, P., Dasgupta, N., et al. (2012). Implicit bias in the courtroom. *UCLA Law Review*, *59*, 1124-1186.

Kawakami, K., Dion, K., & Dovidio, J. (1998). Racial prejudice and stereotype activation. *Personality and Social Psychology Bulletin, 24,* 407-416.

Krippendorf, K. (2012). *Content Analysis: An Introduction to its Methodology.* Thousand Oaks: Sage Publications.

Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. L., Joy-Baga, J. A., et al. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General, 143,* 1765-1785.

Lieberman, J.D., & Arndt, J. (2000). Understanding the limits of limiting instructions: Social psychological explanations for the failures of instructions to disregard pretrial publicity and other inadmissible evidence. *Psychology, Public Policy, & Law*, *6*, 677-711.

Lynch, M., & Haney, C. (2000). Discrimination and instructional comprehension: Guided discretion, racial bias, and the death penalty. *Law & Human Behavior*, *24*, 337-358.

Lynch, M., & Haney, C. (2009). Capital jury deliberation: Effects on death sentencing, comprehension, and discrimination. *Law & Human Behavior*, *33*, 481-496.

Lynch, M., & Haney, C. (2011). Mapping the Racial Bias of the White Male Capital Juror: Jury Composition and the "Empathic Divide." *Law & Society Review*, *45*, 69-102.

Lynch, M. & Haney, C. (2015). Emotion, authority, and death: (Raced) negotiations in mock capital jury deliberations. *Law & Social Inquiry, 40,* 377-405.

Lynch, M., Kidd, T., & Shaw, E. V. The subtle effects of implicit bias instructions. Under revision for *Law & Policy.*

Maeder, E. M., & Hunt, J. S. (2011). Talking about a Black man: The influence of defendant and character witness race on jurors' use of character evidence. *Behavioral Sciences & the Law*, *29*, 608-620.

Maeder, E. M., & Yamamoto, S. (2017). Attributions in the courtroom: the influence of race, incentive, and witness type on jurors' perceptions of secondary confessions. *Psychology, Crime & Law*, *23*, 361-375.

Meertens, R. W., & Pettigrew, T. F. (1997). Is subtle prejudice really prejudice? *Public Opinion Quarterly*, *61,* 54-71.

Model Criminal Jury Instructions, Ninth Circuit, 2010 Available at: https://www.rid.uscourts.gov/sites/rid/files/documents/juryinstructions/otherPJI/9th%20Circuit%20Model%20Criminal%20Jury%20Instructions.pdf

Negowetti, N. E. (2014). Implicit bias and the legal profession's diversity crisis: A call for self-reflection. *Nevada Law Journal*, *15*, 930-958.

Neuschatz, J. S., Lawson, D. S., Swanner, J. K., Meissner, C. A., & Neuschatz, J. S. (2008). The effects of accomplice witnesses and jailhouse informants on jury decision making. *Law & Human Behavior*, *32*, 137-149.

Neuschatz, J. S., Wilkinson, M. L., Goodsell, C. A., Wetmore, S. A., Quinlivan, D. S., & Jones, N. J. (2012). Secondary confessions, expert testimony, and unreliable testimony. *Journal of Police & Criminal Psychology*, *27*, 179-192.

Payne, B. K., Vuletich, H. A., & Lundberg, K. B. (2017). The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry*, *28*(4), 233-248.

Peter-Hagene, L.C. (2019). Jurors' cognitive depletion and performance during jury deliberation as a function of jury diversity and defendant race. *Law & Human Behavior*, *43*, 232-249.

Pfeifer, J., & Ogloff, J. (1991). Ambiguity and guilt determinations: A modern racism perspective. *Journal of Applied Social Psychology, 21,* 1713-1725.

Rodriguez-Pena v. Colorado 137 S. Ct. 855 (2017).

Schreier, M. (2012). *Qualitative Content Analysis in Practice*. Thousand Oaks, CA: Sage.

Shaw, E. V., Lynch, M., Laguna, S., & Frenda, S. (forthcoming). Race, witness credibility and jury deliberation in a simulated drug trafficking trial. *Law & Human Behavior.*

Smith, R.J. (2015). Reducing racially disparate policing outcomes: Is implicit bias training the answer? *University of Hawaii Law Review*, *37*, 295-312.

Sommers, S. R. (2006). On racial diversity and group decision making: identifying multiple effects of racial composition on jury deliberations. *Journal of Personality & Social Psychology*, *90*, 597-612.

Steblay, N., Hosch, H. M., Culhane, S. E., & McWethy, A. (2006). The impact on juror verdicts of judicial instruction to disregard inadmissible evidence: A meta-analysis. *Law & Human Behavior,* 30(4), 469–492.

Wiener, R. L., Krauss, D. A., & Lieberman, J. D. (2011). Mock jury research: Where do we go from here? *Behavioral Sciences & the Law*, *29*, 467-479.

Wetmore, S. A., Neuschatz, J. S., & Gronlund, S. D. (2014). On the power of secondary confession evidence. *Psychology, Crime & Law*, 20, 339-357.

Williams, M. J., & Eberhardt, J. L. (2008). Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology*, *94*(6), 1033-1047.

Zestcott, C. A., Blair, I. V., & Stone, J. (2016). Examining the presence, consequences, and reduction of implicit bias in health care: a narrative review. *Group Processes & Intergroup Relations*, *19*(4), 528-542.