Version 1.0

Juror Number Six:
Implicit Bias and the Future of Jury Trials

Dave McGowan*
University of San Diego School of Law
dmcgowan@sandiego.edu

* Lyle L. Jones Professor of Competition and Innovation Law, University of San Diego School of Law. Thanks to Kevin Cole, Tony Greenwald, Mona Lynch, Greg Mitchell, and Antony Page for comments. Remaining mistakes are my fault.

Electronic copy available at: https://ssrn.com/abstract=4457050

Implicit bias concepts have been quick to take hold in the law. Their influence may be traced through mandatory continuing education requirements,[1] statutory restrictions on jury selection,[2] jury instructions,[3] and, in two recent rulings from the Washington Supreme Court, judicial limits on juror deliberations and permissible arguments. These limits use implicit bias premises to experiment with altering the balance of interests served by jury trials.

Part I of this Article tells the story of the trial of Tomas Berhe, who was convicted of shooting Everett Williams to death. Mr. Berhe's case led to an extraordinary event: a live hearing in which jurors were shown an implicit bias video, testified about their deliberations, and were examined and cross-examined by the court and counsel, all to try to determine whether implicit bias affected deliberations. Expert testimony from a preeminent implicit bias researcher was heard as well. Prior cases had allowed jurors to submit affidavits accusing other jurors of using racist language, accusations that supported an inference that the accused jurors cast racist votes. In Mr. Berhe's case the hearing was triggered by a juror's affidavit that her own vote had been tainted by bullying attacks and disrespect that did not involve overtly racist language. The affidavit claimed bias against the juror, not the defendant. The tainted vote was the juror's own.

Part II briefly surveys current research on implicit bias and compares what is presently believed probable with the way the Washington courts use the concept. Data from the most well-known implicit measure, the Implicit Associations Test ("IAT"), were recounted in *Berhe*, but they were uninformative relative to the task at hand. If anything, they may have led the court to truncate analysis in unhelpful ways. That is not a general point against the IAT, which is not intended to measure persons' ability to perform analytic tasks as a member of a group analyzing evidence pursuant to instructions, in which each person has the power to block an affirmative decision, which is what a criminal jury is. Its value as a research tool relevant to social policy is not implicated. Most of the current debate over the IAT and implicit bias theory, including the sharper exchanges on the topic, relate to general policies. This Article's concern is with individual cases and what tools are most apt for analyzing them.

Part III assesses Washington's experiment with respect to deliberations and its increasing use of implicit bias rhetoric. The experiment traces its origins to dissatisfaction with the rule of *Batson v. Kentucky*,[4] which required a showing of racist intent to reject peremptory challenges in jury selection. The *Batson* problem was never that lawyers making challenges were innocent victims of unconscious forces, nor was it that inferences of bias could not be drawn from patterns of challenges. The problem identified by scholars and Washington courts was that judges were reluctant to accuse lawyers of making race-based challenges. Both the challenges and the reluctance were conscious processes. Implicit bias concepts sidestepped that problem by

[1] E.g. Cal. Bus. and Prof'n Code § 6070.5 (adopting mandatory implicit bias CLE training requirement)
[2] Wash. Gen. R. 37; Cal. Code Civ. P. § 231.7.
[3] *E.g.* Ninth Circuit Model Criminal Jury Instruction 1.1 ("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 can affect how we evaluate information and make decisions.").
[4] 476 U.S. 79 (1986).

Electronic copy available at: https://ssrn.com/abstract=4457050

allowing judges to reject a peremptory challenge in a non-accusatory way, without fear of reversal, on the ostensible ground that the lawyer might not know their own mind.

Berhe extends the implicit bias aspect of Washington's jury-selection rule to post-trial analysis of deliberations. It opens jury deliberations to a greater extent than prior exceptions to the default rule that deliberations must remain confidential. But implicit bias arguments are too strong to do the work *Berhe* tried to assign them. Taken literally, the implicit bias premise entails that all verdicts are tainted by bias, because all jurors are people, and all people are biased. If a fair trial means a bias-free trial, then implicit bias premises entail that a fair trial is a myth. There may be a sense in which that is right, but it is not a sense useful to the law at present. Jury trials are not going away and, if they did, they would be replaced by some other mode of decision to which implicit bias arguments would apply as well.

Part III argues that courts should not structure inquiries in cases such as *Berhe* as efforts to discern the influence of implicit bias. There is no point to a hearing whose goal is to detect something the Washington Supreme Court has declared to be pervasive, impervious to introspection, and undetectable by direct inquiry, which *Berhe* does. If inquiries into juror conduct are to look for causes, then the inquiries should be framed as attribution problems: the question should be whether the conduct at issue should be attributed (at whatever threshold a court specifies) to dispositional factors--attitudes, implicit or otherwise--or external facts such as the trial record and jury instructions. It is not clear where such an inquiry would have landed in *Berhe*, because it was not done. And it was not done in part because the trial court framed the remand hearing as a search for implicit bias.

Alternatively, such an inquiry could seek not to establish causal relationships (at any threshold that might be posited) but to establish courts as a civil and safe environment for persons for whom courts historically have been hostile places. In neither case, however, will implicit bias premises distinguish between permissible and impermissible conduct. This Article concludes that the utility of implicit bias rhetoric in cases such as *Berhe* is not to provide answers. The utility is to give judges something to point to as a way of sidestepping traditional causal analysis, and thereby free judges to pursue justice as they see it. Whether that turns out to be a good thing depends on what courts do with that freedom.

I

In 2013, Everett Williams was shot dead while sitting in a parked car. In 2016, after five weeks of trial, Tomas Berhr was convicted of first-degree murder for the shooting. Jurors deliberated for three and a half days. They were polled and confirmed their verdict.[5]

[5] 443 P.3d 1172 (2019).

Electronic copy available at: https://ssrn.com/abstract=4457050

A. Juror 6's Allegation and the New Trial Motion.

The next day, Juror 6 contacted defense counsel and the court.[6] She said she had acquiesced in the verdict against her wishes. Defense counsel began contacting other jurors, whose contact information Juror 6 had provided. After these contacts came to light the prosecution began contacting other jurors as well. The trial court ordered counsel to stop contacting jurors and sent letters to each juror providing contact information for counsel on each side.[7] Multiple affidavits were obtained, one by the defense and six by the prosecution.

Mr. Berhe moved for a new trial based in part on juror misconduct. Juror 6 provided an affidavit in support of the motion. She averred that she had been the last juror holding out for a not-guilty verdict and that other jurors had accused her of being "partial," which she understood as referring to her race because she was the only African-American juror and the defendant was African-American.[8] She stated that other jurors were dismissive of her comments and characterized as stupid or illogical some of her suggestions, such as that the Defendant (who was arrested in a car containing the murder weapon) "could have taken the gun from the real shooter." [9] The affidavit did not accuse other jurors of using explicit racial language, but Juror 6 stated she was attacked[10] and "felt emotionally and mentally exhausted from the personal and implicit race-based derision from other jurors."[11] The prosecutors' affidavits asked affiants to answer two questions focused on motivation.[12] These jurors denied that bias played any role in deliberations.[13]

The trial court denied the motion for new trial. It noted that Juror 6 did not contend that any racial language had been used and stated: "I understand about implicit bias. But we can't just assume. I think it's equally likely that she felt pressured because she was the lone holdout as she did because of any other reason."[14] The trial court said "[i]t is not inappropriate for jurors to press other jurors on their respective positions during deliberations" and that "[t]he remaining hold-out juror is frequently subject to pressure by fellow jurors and such pressure is not

---

[6] Juror 6's name is public. She did not choose to be a juror, however, nor to be at the center of the Washington Supreme Court's emerging implicit bias trial jurisprudence. I therefore use her juror number, for two reasons. It seems more respectful of such privacy as she retains and (less importantly) it makes the point that she could be anyone in a category protected by law. Her gender may be material, however, so it is reported here.
[7] 443 P.3d at 1176.
[8] Declaration of Juror # 6 ¶ 16. The affidavit uses the term "African-American" and I follow that usage when referring to the affidavit. In her live testimony Juror 6 used the term "Black," and I follow that usage except when referring to the affidavit.
[9] *Id.* ¶ 17.
[10] *Id*. ¶ 21.
[11] *Id.* ¶ 20.
[12] "Each declaration stated that the prosecutors had asked each juror `to answer in my own words the following two questions`: (1) `Did you personally do anything to Juror #6 which was motivated by racial bias during deliberations?' and (2) `Did you observe any other juror do anything to Juror #6 which appeared to be motivated by racial bias during deliberations?'" 443 P.3d at 1177.
[13] *Id.*
[14] *Id*. at 1178.

Electronic copy available at: https://ssrn.com/abstract=4457050

inappropriate."[15] The trial court held that Juror 6's affidavit failed to make a *prima facie* showing of racist juror misconduct, implicit or explicit.

The court of appeal affirmed in an unpublished opinion that did not treat the bias argument as anything more than one of several challenges.[16] The court noted that Juror 6 did not accuse other jurors of using explicitly racist language or of being biased against Berhe, and it distinguished Berhe's case from a prior decision in which one juror accused another of using overtly racist language.[17]

B. The Washington Supreme Court Decision.

The Washington Supreme Court reversed.[18] The Court held that the trial court abused its discretion by not holding an evidentiary hearing to determine whether a *prima facie* case of racist misconduct occurred. The Court recognized that such a hearing ran counter to the principle that secrecy of deliberations is foundational for a jury system,[19] a principle reflected in the common law rule that jurors may not impeach a verdict and in provisions such as Federal Rule of Evidence 606.[20] It noted, however, that in *Pena-Rodriguez v. Colorado*,[21] the United States Supreme Court has recognized a constitutional exception for cases in which one juror accused another of overtly racist comments that raised a question whether the accused juror's vote was tainted by racist bias. The Court held that

---

[15] State v. Berhe, 2 Wash. App. 1024 at *15 (2018). In a recent Washington case in which a juror was dismissed for cause because he had an outburst in which he began punching himself in the face, the defense position was: "juries often `get heated.'" State v. Norman, 2023 WL 1456755, at *2 (Wash. Feb. 2, 2023). This view would be consistent with defense-side efforts to protect holdout jurors, or those perceived to be holdout candidates, from dismissal. *Berhe* does not address that question, but its reasoning may have implications for it. *See infra* text accompanying note [N].

[16] *Id.*

[17] *Id.*

[18] State v. Berhe, 443 P.3d 1172 (2019).

[19] *Id*. at 1174. The importance of confidentiality drives strict standards for conducting inquiry during deliberations to determine whether a juror should be discharged based on allegations that they are engaged in nullification. *E.g.* United States v. Thomas, 116 F.3d 606, 614 (2d Cir. 1997). One question, discussed below, is whether *Berhe* will be extended to expand inquiry in such cases.

[20] Rule 606(b) provides:

> During an Inquiry into the Validity of a Verdict or Indictment.
> (1) Prohibited Testimony or Other Evidence. During an inquiry into the validity of a verdict or indictment, a juror may not testify about any statement made or incident that occurred during the jury's deliberations; the effect of anything on that juror's or another juror's vote; or any juror's mental processes concerning the verdict or indictment. The court may not receive a juror's affidavit or evidence of a juror's statement on these matters.
>
> (2) Exceptions. A juror may testify about whether:
> (A) extraneous prejudicial information was improperly brought to the jury's attention;
> (B) an outside influence was improperly brought to bear on any juror; or
> (C) a mistake was made in entering the verdict on the verdict form.

[21] 580 U.S. 206 (2017).

Electronic copy available at: https://ssrn.com/abstract=4457050

> [A] showing that one or more jurors made statements exhibiting overt racial bias that cast serious doubt on the fairness and impartiality of the jury's deliberations and resulting verdict. To qualify, the statement must tend to show that racial animus was a significant motivating factor in the juror's vote to convict.[22]

Washington had reached a similar conclusion earlier in *State v. Jackson*,[23] a case also involving a juror affidavit stating that another juror repeatedly expressed overtly racist views directed at a category of persons that included the defendant, certain alibi witnesses, and the juror who submitted the affidavit. The Washington court of appeal held that where a juror affidavit established a "clear inference of racial bias" a trial court should conduct an evidentiary hearing into allegations of bias before ruling on a new trial motion.[24]

Berhe differed materially from these cases. The juror statements in *Pena-Rodriguez* and *Jackson* alleged overtly racist statements by other jurors, and the premise of those opinions was that the juror affidavits provided grounds to believe the jurors who made racist statements cast racist votes because they were biased against the defendant or alibi witnesses. Juror 6 did not accuse other jurors of using racist language, nor did she accuse them of voting to convict Berhe because he was Black. She instead accused them of attacking her, treating her dismissively, and accusing *her* of partiality, because both she and Berhe were Black.

Berhe thus creates a broader exception to the confidentiality of deliberations than was found in either *Pena-Rodriguez* or *Jackson.* The Court used implicit bias theory to justify the extension. The premise of the Court's ruling was that racial bias was "common and pervasive," and "uniquely difficult to identify" in part because "implicit racial bias exists at the unconscious level."[25] The Court stated that "juror's racial bias is uniquely difficult to identify because many people who harbor explicit biases will not admit to doing so, and everyone harbors implicit biases that are difficult to recognize in oneself."[26]

Berhe holds that when an allegation of bias is made the trial court must take control of any investigation, so counsel do not taint jurors. The Court chided the prosecutors in *Berhe* because, the Court held, their questions (which focused on racially motivated conduct) "could well exacerbate whatever prejudice might exist without substantially aiding in exposing it" by making the jurors feel accused and defensive," a claim for which it cited *Pena-Rodriguez.*[27] The cited passage traced back to a comment by Justice Rehnquist, concurring in *Rosales-Lopez v. United States*,[28] which held it was not error for a judge conducting *voir dire* to decline to

[22] *Id*.
[23] 75 Wash. App. 537, 879 P.2d 307 (1994), *review denied*, 126 Wash.2d 1003, 891 P.2d 37 (1995).
[24] 75 Wash App. at 660.
[25] *Id*.
[26] 443 P. 3d at 1180.
[27] *Id*.
[28] 451 U.S. 182 (1981).

Electronic copy available at: https://ssrn.com/abstract=4457050

question the venire about possible racial prejudice against the defendant.[29] Justice Rehnquist's concurrence argued that trial judges should be given discretion to ask such questions or not, in part because such questions might have the potential to exacerbate existing prejudice. It is not clear that the data support a conclusion that direct questions invariably produce backlash, however, nor did Justice Rehnquist argue that they did. (Part II examines the literature to see how far the Court's concern was warranted.) To the extent the Court wanted to refocus inquiry from intention to unconscious factors, however, its criticism was fair: the prosecutors' questions focused on motivation.

The *Berhe* Court's instruction that a trial court take control of inquiry as soon as an allegation of racist conduct is made was clear and should be easy to follow. The Court's comments on unconscious bias were less so. Acknowledging that there was no claim that explicit racial language had been used in deliberations, the Court said:

> determining whether a person has been influenced by implicit racial bias is inherently challenging. Courts cannot simply ask the person about it and assess the credibility of his or her response because "people are rarely aware of the actual reasons for their discrimination and will genuinely believe the race-neutral reason they create to mask it . . . . Therefore, a person may honestly believe and credibly testify that his or her actions were not influenced by racial bias, even where implicit racial bias did in fact play a significant role.
>
> Nevertheless, we should not "throw up our hands in despair at what appears to be an intractable problem. Instead, we should recognize the challenge presented by unconscious stereotyping ... and rise to meet it."[30]

The Court posited that implicit bias is unconscious, so biased jurors may truthfully testify that they acted without bias.[31] That premise implied that testimony alone cannot be decisive

[29] The defendant had asked that the trial court pose the following question: ""Would you consider the race or Mexican descent of Humberto Rosales-Lopez in your evaluation of this case? How would it affect you?" 451 U.S. at 185. Presumably a retrospectively worded version of this question would be impermissible under the Washington Supreme Court ruling in *Berhe.*

[30] 443 P.3d at 1181, quoting *State v. Saintcalle*, 178 Wash.2d 34 (2013).

[31] For this premise the Court cited *State v. Saintcalle*, 178 Wash.2d 34 (2013), which in turn cited literature such as Antony Page, *Batson's Blind–Spot: Unconscious Stereotyping and The Peremptory Challenge,* 85 B.U. L. REV. 155 (2005), which in turn cited primarily work from the early 1970s through the early 1990s. The IAT was introduced in 1998, and the article does capture some of the early reports, including a citation to the foundational paper, Anthony G. Greenwald & Mahzarin R. Banaji, *Implicit Social Cognition: Attitudes, Self-esteem, and Stereotypes,* 102 Psychol. Rev. 4, 4 n.1 (1995). Professor Page, but not the Court in *Saintcalle* or *Berhe*, does note the difficulty in specifying a referent for the term "unconscious." 85 B.U. L. Rev. at 182 n. 118 ("Unconscious" itself is an extremely challenging word to adequately define. Fortunately, there is no need in this article to closely parse the many sometimes subtly different definitions of unconscious and conscious."). 2013, the year *Saintcalle* was decided, was something of a turning point in IAT-related research. Professor Oswald and colleagues published a meta-analysis of IAT data that called into question whether IAT scores could generate valid inferences regarding behavior. Oswald et al *Predicting Ethnic and Racial Discrimination: A Meta-Analysis of IAT Criterion Studies*, 105

Electronic copy available at: https://ssrn.com/abstract=4457050

because truthful testimony is insufficiently probative of unconscious truth. Similarly, the Court held that "courts cannot base their decisions on whether there are equally plausible, race-neutral explanations" for behavior because "[t]here will almost always be equally plausible, race-neutral explanations because that is precisely how implicit racial bias operates."[32] The Supreme Court instead announced a standard derived from Washington's General Rule 37, which is directed to the pre-trial exercise of peremptory challenges:

> The ultimate question for the court is whether an objective observer (one who is aware that implicit, institutional, and unconscious biases, in addition to purposeful discrimination, have influenced jury verdicts in Washington State) could view race as a factor in the verdict. If there is a prima facie showing that the answer is yes, then the court must hold an evidentiary hearing.[33]

Because the Court had altered the standard for establishing a *prima facie* case, it remanded to the trial court to determine whether such a case existed and, if so, to hold a hearing. With respect to how the hearing was to be conducted, the Court said:

> Where the evidence is unclear or equivocal, as it will often be in cases of alleged implicit racial bias, the court must conduct further inquiries before deciding whether a prima facie showing has been made, for example, by asking the juror making the allegations to provide more information or to clarify ambiguous statements.[34]

The opinion in *Berhe* identifies no possible source of unbiased information. At face value, its premises effectively declare that no such sources exist because everybody suffers from unconscious bias and because credible testimony is not enough to dispel an inference of bias. Overt questions are suspect because they might provoke defensiveness, but the opinion does not tell trial judges what to do with the answers to more oblique questions. And the Supreme Court's premises extended to Juror 6, whose presumed unconscious bias logically would be relevant to the probative weight of her testimony. And what of other biases accepted in psychology, such as the tendency to attribute one's own actions to circumstances while attributing the actions of

---

Journal of Personality and Social Psychology 171 (2013). IAT proponents responded with an article acknowledging modest statistical predictability for individual behavior but presenting a simulation in which modest individual probabilities implied significant social effects. Greenwald, A. G., Banaji, M. R., & Nosek, B.A., *Statistically Small Effects of the Implicit Association Test Can Have Societally Large Effects*, Journal of Personality and Social Psychology 2015, Vol. 108, No. 4, 553–561. The terms of debate shifted thereafter.

[32] 443 P.3d at 1182.

[33] *Id*. at 1181. General Rule 37 provides that a party may object to a peremptory challenge exercised by another party. The party exercising the challenges then must state (out of the present of the jurors) the reasons for the challenge. The court then consider the totality of the circumstances and, "if the court determines that an objective observer could view race or ethnicity as a factor in the use of the peremptory challenge, then the peremptory challenge shall be denied. The court need not find purposeful discrimination to deny the peremptory challenge." Wash. Gen. R. 37(e). The rule defines in part the nature of the objective observer, who is deemed to be aware "that implicit, institutional, and unconscious biases, in addition to purposeful discrimination, have resulted in the unfair exclusion of potential jurors in Washington State." *Id*. §(f).

[34] *Id*. at 1182.

Electronic copy available at: https://ssrn.com/abstract=4457050

others to their disposition?[35] Though stating its determination not to throw up its hands, the *Berhe* Court's embrace of implicit bias premises offered the trial court little guidance, and no criteria, for weighing competing inferences or for assessing unconscious motivations.

C. The Remand Hearing.

The remand hearing was extraordinary. Jurors testified live and were examined by the parties and the court. Expert testimony from Professor Anthony Greenwald, a preeminent implicit bias researcher, was heard. The author is not aware of any comparable inquiry focusing on a juror's perception of her own behavior, nor any that were not triggered by allegations of overtly racist conduct. *Berhe* is the closest thing we have to a true test of the utility of implicit bias premises in assessing the lawfulness of real-world jury deliberations.

1. The evidence at trial

Recall that Everett Williams was shot to death in a car, and Berhe was charged with first degree murder for the crime. The prosecution introduced evidence that: (i) Berhe's fingerprints were on the outside of the car; (ii) an eyewitness saw a Black man wearing dark clothes and a white cap standing outside the car pointing at the car an object that sparked and made popping sounds; this witness also saw a black sedan leaving the scene; (iii) Berhe was arrested in a black Chevy Impala, and his clothes matched the description above; (iv) the murder weapon (identified by ballistic firing tests) was found in the Impala; (v) a second eyewitness was taken to the arrest scene and identified Berhe's clothing and the Impala as being present at the murder scene; (vi) the driver of the Impala, Mr. Washington, testified that earlier on the day of the shooting he heard Berhe say that Williams had been involved in shooting Berhe's cousin and that Berhe stated he planned to "take care of" Williams; and (vii) Mr. Washington testified that he picked Berhe up shortly after the shooting, near the scene of the shooting, that Berhe pulled a gun from his waist when he got in the car, and that Berhe then admitted to shooting Williams.

Berhe's defense was misidentification: someone else shot Williams. The defense argued that: (i) eyewitness testimony varied and was unreliable; (ii) no eyewitness actually saw the shooting; (iii) the victim used drugs, had drugs on his person when killed, had been accused by someone other than the defendant of stealing drugs, and had been heard wondering whether this other person might try to kill him; and (iv) Mr. Washington's testimony was unreliable because he had been given immunity and had admitted to lying to police several times. The defense attacked the credibility of Mr. Stuckenberg, the driver of the car Wiliams was shot in, and who was also hit by a bullet in the shooting. The defense suggested Stuckenberg had arranged for someone other than Berhe to kill Williams. The defense also suggested an empty-chair defense,

[35] Sometimes called fundamental attribution error, or correspondence bias. *E.g.* Daniel T. Gilbert & Patrick S. Malone, *The Correspondence Bias*, 117 PSYCH. BULL. 21 (1995)("People may make inferences about the dispositions of others even when situational forces explain the behavior quite nicely. In scores of experiments, subjects have violated attribution theory's logical canon by concluding that an actor was predisposed to certain behaviors when, in fact, those behaviors were demanded by the situations in which they occurred.").

Electronic copy available at: https://ssrn.com/abstract=4457050

attempting to throw suspicion on a person named Dominic Oliveri, who was at the scene of the murder and disappeared quickly thereafter, emerging later represented by counsel.

Mr. Williams, Mr. Washington, and Mr. Berhe were Black. Mr. Stuckenberg and other prosecution witnesses were White.[36] On remand, the trial court characterized the jury as "diverse."[37] The jury deliberated for three and one-half days after the five-week trial.

2. Juror 6's affidavit and testimony.

Juror 6 was female and the only Black juror. As noted above, for the new trial motion, she submitted an affidavit in which she stated she felt "personally attacked and belittled" and that "I felt these attacks carried an implicit racial bias."[38] She noted that three other jurors initially doubted Berhe's guilt,[39] but she was the only juror that other jurors accused of being partial:

> My opinion about the evidence was disregarded and disparaged. Other jurors would repeatedly question my concerns about the evidence, accusing me of being partial. Here are some examples: I was concerned that the State's case rested on the wholly circumstantial evidence. I tried to discuss the lack of concrete evidence that Mr. Berhe was the shooter, but other jurors were outright dismissive. (2) When discussing how Mr. Berhe could have had possession of the gun, I suggested that he could have taken the gun from the real shooter. My idea was mocked as "stupid" and "illogical," despite the fact that the same jurors believed that Mr. Washington took the gun from Mr. Berhe. I felt personally ridiculed in a way the other dissenting jurors were not. (3) During a discussion of Mr. Berhe's clothing on Tuesday morning March 1st following the experiment,[40] one

---

[36] The race of the witnesses is taken from Berhe's appellate brief: "This racial hostility arose in the context of a case that depended on assessing the credibility and behavior of African-American participants (Berhe, Washington, and Williams) and Caucasian observers or accusers. Ex. 20 (pictures of Berhe, Washington, Williams, Stukenberg, Villiott, Cascioppo, and others in social group). Brief at 15-16.

[37] The court stated:

> The jury was diverse in its makeup and Juror 6 was not the only non-white juror. One juror described her father as Hispanic, and testified that she had a "biracial" nephew and granddaughter. Another juror testified that she thought a fellow juror was Filipino. A third juror testified that the jury appeared to include someone who was Native American and "at least two people of Asian heritage."

Final Order, *supra* note **[n]** at 6. The latter reference is to an affidavit from Juror five.

[38] Affidavit, *supra* note [**n**] ¶¶ 13-14.

[39] Two of whom changed their minds "pretty quickly" according to Juror 6's live testimony. Motion for New Trial Attachment C at 10. Juror 6 testified that these other jurors "were allowed to express themselves without feeling I guess badgered or embarrassed about the answers or anything," *id.* at 9, but the transcript does not provide foundation (such as statements by the jurors in question) for testimony about these other jurors' state of mind.

[40] The experiment referred to the jurors examining the defendant's clothes with the lights off in the jury room, to see for themselves how visible the clothes would be in poor light. The point of the experiment was to assess the credibility of the witness who testified to seeing a man in dark clothes and a white cap standing outside Williams's

Electronic copy available at: https://ssrn.com/abstract=4457050

> juror pointed out that if his pants were so big and baggy that he had to hold them up to walk down Eastlake, why didn't he do the same when he exited the car during the stop by police. I rhetorically commented, "What would have happened if he did that?" Several of the jurors as a group interpreted that as a comment on police misconduct towards African-Americans and loudly mocked me for my response, while others sat quietly.[41]

Juror 6's affidavit recounted that two other jurors "surrounded" her, "taunting" her, such that she felt "like they were animals and I was their prey. I can only describe it as feeling emotionally and mentally attacked. I felt emotionally abused; so much so that it became debilitating. I couldn't handle the pressure of being a hold-out anymore."[42] Her affidavit concludes: "I think Mr. Stukenberg should be charged with murder."[43] (Recall that Stukenberg had been sitting in the car when Williams was shot.)[44]

Juror 6's live testimony was consistent with her affidavit but milder. She confirmed that other jurors did not state that she was being partial because she and the Defendant were both Black.[45] Regarding what turned out to be a salient issue when the defense argued for a *prima facie* showing,[46] the jurors debated whether the defendant had baggy pants. The topic was relevant because some witnesses testified that a man wearing clothes consistent with the defendant's clothing when arrested was running and holding his waistband as he got into a car; Mr. Washington (the driver of the Impala) testified that the defendant pulled a gun from his waist when he got into the car.[47] In the police video of the arrest, the defendant's hands were raised but some jurors testified that his pants were not falling down. Those jurors reasoned that if the pants

---

car and holding something that sparked and made popping noises (the prosecution testimony recounted as (ii) above). Juror 6 stated that this testimony carried great weight with the jurors. Affidavit, *supra* note [n] ¶¶ 24-25.

[41] *Id.* at ¶ 17.

[42] *Id.* at ¶ 18. Her live testimony stated:

> It was just during a little break we had among ourselves. We were still in the room, and the room is small, and I was just sitting there, and they was just standing behind and talking to themselves but yet talking loud enough for me to hear.

Motion for New Trial Attachment C at 18.

[43] *Id.* at ¶ 28.

[44] Part of Juror 6's affidavit identified with Williams, the driver of the Impala who testified that Berhe admitted shooting Williams: "At that moment, I could relate to Elijah Washington telling the Detectives what they wanted to hear so he could get out of the interrogation room." *Id.* at ¶ 19.

[45] *Id.* 29:18-25.

> The jurors didn't say that. That was my impression, because they used terms like partial and different, yeah, just partial, using that by saying that I'm black and "You're just partial to the defendant because he's black."
> Q But they never used the term, you're black, the defendant's black?
> A No.

[46] 12/13/19 Transcript at 35 ("I think juror 6 articulated specific topics of conversation, including the baggy pants issue and as it related to violence between police and African American citizens. And the baggy pants issues, I think that clearly gets into the topic of implicit bias, particularly her response of describing the other jurors saying, "Get out of here with that.").

[47] Respondent's brief at 11.

Electronic copy available at: https://ssrn.com/abstract=4457050

did not fall down while the defendant's hands were raised, it was fair to infer that the defendant might have been holding his waistband to conceal something, such as a gun.[48]

To counter this line of argument, Juror 6 commented that wearing baggy pants was common in hip-hop culture, and thus should not be considered evidence of the defendant's guilt.[49] She read the reaction to her comments as evidence of implicit racism:

> Well, I felt a lot of it was more racially motivated cuz when things that I said didn't -- it didn't make sense to them, or it was more me being impartial -- being partial towards what I was saying, and it was like, oh, that's stupid or that's illogical or that don't make sense or something.
>
> You know, it's like but cussing like in paragraph 17, we was talking about him running down the street with his pants, holding up his pants or something, and when -- I guess when he got stopped by the police and they wanted him to get out of the car and they said, "Well, if his pants was so big, why didn't he pull up his pants then?" And so I said, "Well, what do you think would have happened to him if he took his hands down?" And so – and the whole -- sorry, I'm being nervous, the whole jury room, most of the jury room said, whole, the majority of it, was like, "Oh, that, get out of here with that, we don't want to hear that," as though I was saying the police -- well, I was saying the police would have shot him, but that was illogical to them. They wouldn't have done that.
>
> So that really hurt me, like really like what is wrong, these people are not getting it, you know, so it was things like that that touched me and made me upset. So... [50]

Regarding the affidavit's statement that Juror 6 was personally attacked and belittled, her live testimony states:

> THE WITNESS: Well, again, being as far as -- well, at one point when -- as far as the personal attack, I mean they didn't call me names, or anything like that, it was just the feeling and the little things that they were kind of I guess doing in a sense, like once we were just taking a break, and I was just sitting there reading and

[48] New Trial Motion, 7/24/20 Tr. at 11-12 (Juror 7). Juror 6's testimony was roughly consistent on this point, but she interpreted the other jurors as arguing that there would be no reason for a person to run down the street holding their pants unless they were hiding a gun. *Id*. at 30-31 (Juror 6).

[49] Juror 7 testified that it was Juror 6 who raised the issue of hip-hop culture as an alternative explanation for why the defendant might have been holding the waist of his pants (assuming the person identified by witnesses was the defendant). New Trial Motion, 7/24/20 Tr. at 13. Juror 3 testified that she supported Juror 6's comments. New Trial Motion, 8/5/20 Tr. at 60 (Juror 3). Juror 3 was not sure whether she or Juror 6 raised the issue of hip-hop culture.

[50] 12/13/19 Transcript at 35 (Juror 6). Juror 6 later testified that "a lot of people were like, "`Oh, get out of here with that, we don't want to hear all that," and so forth. And one of the jurors that sat next to me was a young man, `Well he would have been shot.' I said, `Yeah, that would have happened,' so that's the end of that." *Id*. at 31-32. Juror 6 testified that three or four jurors had the reaction she described. *Id*. at 32. Other testimony suggests that the young man who agreed with Juror 6 was the foreperson. New Trial Motion, 7/24/20 Tr. at 24 (Juror 7).

Electronic copy available at: https://ssrn.com/abstract=4457050

> then, you know, people just is being around, the room is not very big anyway, but I still could hear them saying, "Well, she's just going to let him walk, and she's not going to do anything"[51] and, you know, or again reminding or just being told questions that I had asked or said, statements that I had made earlier by saying, "That's stupid, that doesn't make sense, or that's illogical" or, you know, things like that.
>
> So that attacked me personally if they're going to say I'm stupid or that's illogical, because a couple times she came, the person who came and said -- they said, "I'm sorry. I'm sorry I said that." She apologized a couple of times, but still she said it. So that was me being personally attacked.[52]

In the course of deliberations, Juror 6 testified that she began to wonder whether she was being partial but decided she was not. One passage reads:

> THE COURT: What did the jurors say to you?
>
> THE WITNESS: Well, I think basically just saying, "You're being partial, you just don't want to convict him," or that kind of stuff, and that really seemed for me seeing evidence in a different way and looking at things in a different way than they was looking at it and like as though they never said as far as being black or something; they just saying, "You being -- I think just she's being partial, you just don't want to see, you just don't want to."
>
> And then I began probably thinking, am I, am I being partial, do I not want to see, but -- but in the end I know I wasn't, because it wasn't beyond a reasonable doubt. There's still some doubt for me in the end anyway.[53]

Defense counsel adduced, and then redirected, similar testimony:

---

[51] The affidavit describes this episode differently:

> Later that same morning on Tuesday March 1st, two of the jurors walked by me as I sat alone looking at my notes, and commented on my hold-out status. They surrounded me, taunting: "She's not going to do anything. She's going to let him walk." I felt like they were animals and I was their prey. I can on[y describe it as feeling emotionally and mentally attacked. I felt emotionally abused; so much so that it became debilitating. I couldn't handle the pressure of being a hold-out anymore.

Affidavit, *supra* note [**n**] ¶ 18. Unlike the prosecution affidavits, Juror 6's affidavit is on pleading paper with a Northwest Defenders Division footer. The appellate court stated that "[d]efense counsel worked with the holdout juror to craft a declaration that set forth the juror's concerns." State v. Berhe, 2 Wash.App.2d 1024 (2018). This fact may explain the difference between the strength of the characterization in the affidavit and the live testimony. If true, that would <u>not</u> be improper. Counsel work with witnesses on affidavits all the time.

[52] 12/13/19 Transcript at 35. Juror 6 clarified that only one juror used the term "stupid" and possibly one (at least the same juror) or two used the term "partial." 12/13/19 Transcript at 14-15, 32-33. The terms were used at least two to three times. *Id*. at 33.

[53] 12/13/19 Transcript at12 (emphasis added).

Electronic copy available at: https://ssrn.com/abstract=4457050

> Q Why did you change your vote to guilty?
> A Because I just felt -- you know, I just start -- I don't know. I just gave in to the bullying or the ridicule or maybe I was being partial. I don't know. It's like well --
> Q Is that still accurate, what you said in your declaration?
> A Uh-huh.[54]

3. Testimony from other jurors.

Evidence from other jurors agreed that frustrations were expressed but generally disagreed with Juror 6 with respect to the tone of discussions.[55] On the new trial motion the foreperson submitted a declaration averring that:

> Frankly, it is hard for me to accept that Juror #6 would make these claims without being influenced by someone that was not present. As presiding juror, I worked hard to make sure everyone had their say and all their questions addressed.
>
> Juror #6 was challenged many times because of her opinion and for that reason alone. The other jurors disagreed with her and often she could not support her position with any of the evidence. Many times she [said] I just don't feel like he's guilty which would cause the other jurors to challenge her.

---

[54] *Id*. at 20. The term "redirected" in the text is intended as a compliment. Mr. Berhe's counsel's job was to get him a new trial. If Juror 6 suggested that she might have been partial, even if she reconsidered the idea later, that testimony would tend to weaken Berhe's case, for reasons discussed more fully in Part III. According to the portion of the transcript copied in the text, counsel intervened to direct Juror 6 back to the stable foundation of the affidavit. A follow-up question adduced testimony that indicates the exhaustion Juror 6 likely felt in scrutinizing the evidence, her reactions to it, and her reactions to other jurors:

> So there's the pressure of being a lone holdout?
> A Yeah.
> Q And then the pressure, I think as you described, as how you were treated maybe differently?
> A Right.
> Q And which was it, or was it a combination of both?
> A It was a combination of both, and maybe as a black person saying, well, am I really seeing this, am I trying to not want to see it, I don't know, but I knew for me I didn't have -- the evidence wasn't saying, yeah, he's guilty.

Id. at 21.

[55] For example:

> Everything was pretty contained and calm. People expressed their opinions, but there wasn't really any animosity or feeling of just -- you know, it was just trying to understand and everybody understanding why we were making the decision, but I wouldn't say that it got heated or, you know, any emotional discussion that way.

New Trial Motion, 7/24/20 Tr. at 5 (Juror 7). Juror 5 testified that Juror 6 was treated similarly to jurors inclined to hold out for a second degree murder conviction rather than first degree. 8/3/20 Tr. at 32.

Electronic copy available at: https://ssrn.com/abstract=4457050

> I too was challenged by 10 jurors at once for my opinion about premeditation so I understand what it feels like but I also believe that is the way that deliberations are supposed to work.
>
> I told those last few undecided jurors o three different occasions that if they decided to vote Not Guilty I would not allow any to [sic] challenge them but if they said they were undecided they were fair game to be part of the debate. Juror #6 sat right next to me so I know she was aware of this.
>
> When Juror #6 was the last undecided juror she changed her vote to guilty without anyone saying a word to her.[56]

The foreperson's live testimony was consistent in saying Juror 6 did not provide logical reasons for her views.[57] Other jurors also testified that Juror 6 would not provide reasons for her views and that this fact was a source of some frustration.[58] The foreperson's live testimony elaborated on his affidavit. He testified that Juror 6:

---

[56] New Trial Motion Exh. C, Decl. of Juror 14. The tone of the first paragraph could be read in different ways—defensive, because the foreperson apparently felt it was his job to facilitate discussion, condescending toward Juror 6 on the ground that she might not know her own mind, or shrewd, if one thinks that Juror 6 might in fact have talked to others before coming forward. That question seems not to have been asked. Another juror testified:

> And he asked her, he says, "Now, I want to make certain. Are you sure that this is how you feel, and that you don't feel like you're being pressured to vote this way, because we will back you, whichever way you choose?" And she said, no, this is how I am going to vote on this. He asked her three times, and each time she had made her decision that this was the way she was going to vote.

7/24/20 Tr. at 10 (Juror 13).

[57] 8/5/20 Tr. at 37 (Juror 14): "Logically. She couldn't explain why she didn't feel he was guilty, and that was what frustrated us."

[58] 7/24/20 Tr. at 11-12 (Juror 13):

> she couldn't give us anything concrete. It was she just knew. That's all she would tell us. There was no -- there wasn't any reasoning behind it other than she just knew he wasn't guilty, and that's all she would say . . . . nobody ever got angry with her or shouted at her, not like that, no. . . . I guess where we were frustrated was that she didn't have any -- her only reasoning was the color of his shoes. I mean, that was the only thing that she could come up with. It didn't matter what all the other evidence showed, it was just the color of the shoes, was the only thing that she could come up with.

Juror 10 testified:

> She said, "I don't -- I don't believe he did it because it's not clear." But then we were saying that because the video show us where the suspect flee the scene, the gun was under the seat, so it was like a perfect, like, evidence right there, like, that person did it. But we don't understand why she didn't believe that he did it, so --

7/24/20 Tr. at 56-57. Juror 7 testified that Juror 6 was not engaged and was on her phone a lot. *Id.* at 32-33.

Electronic copy available at: https://ssrn.com/abstract=4457050

> said things like well, what if the cops planted a gun? Well, there was no evidence of that, you know. Well, what if this? And, yeah. She just didn't want to convict him even though that everybody else saw the facts. Did it. I was at the point where I was about to call a mistrial honestly on that second day, and I told her, look it, this is what you believe, stick to your guns, we will just tell them again we can't come to a conclusion. And then right at the last minute when I said I was going to call a mistrial she said, without anybody forcing her, she said okay. I -- he's got to be guilty.[59]

The foreperson testified that questioning of Juror 6 became repetitive:

> It wasn't badgering or anything, but it was over and over again. And she would be, you know, they would ask well, do you see what we see. And she wouldn't, you know, she didn't -- she didn't want -- it's almost like she just did not want to say he was guilty even though she knew he was, and she actually came up and said, you know, I don't trust the criminal justice system. And I had problems with it in the past. The people that I know.[60]

The foreperson testified that in response to Juror 6's expressions of distrust in the criminal justice system "several people said well, you can't bring that into here. And we went back and looked at the rules, and told her that you have to look at this thing from what we are seeing here. Not from your past."[61] Several jurors confirmed that they told Juror 6 that at least some of her statements were effectively inadmissible in deliberations.[62] Juror 6 apparently referenced persons she knew who she thought had been treated unfairly by the criminal justice system; other jurors argued that those examples were not relevant.[63] And one juror testified that

---

[59] 8/5/20 Tr. at 11 (Juror 14). Juror 14 testified that several jurors supported Juror 6 and took the position that she had to do what she thought was right. *Id*. at 19.

[60] Id. at 12 (Juror 14). The foreperson testified that Juror 6 may have mentioned her distrust of the criminal justice system twice, but not more than that. *Id*. at 15.

[61] *Id*. at 15. Some jurors indicated a normative affinity for the instructions as embodying the law. One juror testified "I am a law-abiding person, and I believe that the law is to be obeyed. And when it's set forth as to what that means, then following those instructions is important to me." 8/3/20 Tr. at 22 (Juror 5). The foreperson testified "[w]e would consistently review the rules or the instructions that we were given to make sure that we followed things, you know, the way that we were supposed to with the instructions." 8/5/20 Tr. at 8 (Juror 14).

[62] Declaration of Juror 9. Juror 9 stated in an affidavit "Juror 6 was challenged by myself and most other jurors to help us understand why she had her opinion, but when she said over and over again `I just feel he is not guilty' we reminded her that feelings could not be brought into these deliberations that we have been told to reach our verdict based on evidence that was seen in trial and given to us to review."

[63] One juror characterized her reaction to Juror 6 raising these examples:

> I just said to everybody as an opening that, you know, "I'm concerned that we are not that -- we all need to make sure that we're really looking at what we're presented with and not bringing other past experiences to this decision," which is -- I mean, hard to do. We all have things that we bring. And then she made the comment and then I said again, you know, "I don't think you can bring that here. You need to really look and make sure that you're separating the two instances."

7/24/20 Tr. at 9-10 (Juror 7).

Electronic copy available at: https://ssrn.com/abstract=4457050

she argued that Juror 6's reference to hip-hop culture was irrelevant because in the arrest video the defendant's pants were not in fact baggy.[64] All the other jurors denied that racism played any part in deliberations.[65]

4. Professor Greenwald's Testimony.

Berhe's counsel introduced expert testimony from Professor Anthony Greenwald, a preeminent scholar whose work has been foundational for implicit bias concepts. He did not offer an opinion regarding whether implicit bias played a role in the *Berhe* trial.[66] His testimony provided background and might be interpreted as opining on a base rate of bias in the general population. Professor Greenwald testified that he defines implicit bias to refer to "operation of attitudes and stereotypes in ways that can cause bias in judgment or action without one being aware of this operation."[67] He testified that implicit bias "is measurable at a nonzero level" in over 75% of the general U.S. population and in over 80% of Whites and Asian Americans.[68] He testified that "about 1/3 of African American subjects show a preference for Whites, 1/3 show a preference for Blacks, and 1/3 show no preference."[69]

---

[64] 7/24/20 Tr. at 12 (Juror 7):

> [S]omebody made -- I think there was a comment about, "Well, that's just -- you know, that's really typical, or that's -- that could be just because, you know, a lot of young black men run and hold their pants up." I remember that there was a brief discussion about that. But it was like, "Well, but I don't know that that had anything to do with it. The fact was, here they're holding and here they're not. So that was what was supposed to be what we were discussing about, not whether it was a -- and I think that she did bring something up about that, I just remember there being a little bit of a discussion there and it's like, "No, no, no, we're not talking race, or culture, or anything here, it is fact. Here no, here yes."

[65] Ironically, given the weight the court placed on the baggy pants issue, Juror 11 stated in an affidavit

> I do not believe that I personally did anything to nor acted in any way toward Juror #6 which was motivated by racial bias during deliberations. In fact, during deliberations I appreciated some of the insights Juror #6 had, specifically around the norm in hip-hop culture of baggy pants without a belt and the need to hold them up with one's hands. This was a norm that some of the other jurors seemed unaware, and I believed we appropriately discussed that as one possible explanation regarding witness testimony that Berhe was walking with his hands at his waistband.

It is hard to believe jurors were unaware of baggy pants as a fashion that correlates with a specific culture, though not an exclusively Black culture. The prosecution's response to this argument on appeal was, in part: "contrary to the defendant's assertion, hip-hop is not synonymous with African-American youth. See Eminem, Beastie Boys, Machine Gun Kelly, Mac Miller, G-Easy, Lil Dicky, Aesop Rock, Macklemore & Ryan Lewis." Opp. at 21. The correlation might not be a synonym, but that does mean it does not exist as a stereotype.

[66] 8/5/20 Tr. at 105.

[67] *Id.* at 84.

[68] *Id.* at 85.

[69] *Id.* at 92. His data source was the IAT. *Id.* at 113.

Electronic copy available at: https://ssrn.com/abstract=4457050

He described the IAT, of which he was an inventor,[70] and explained its measurement hypothesis.[71] Professor Greenwald testified that implicit bias is caused by "exposure to sociocultural environment starting from a very early age" in which persons encounter stereotypes. "These stereotypes pervade our cultural environment. You can no more avoid being exposed to them, and learning them than you can stop breathing."[72]

Professor Greenwald testified that IAT scores correlate with biased behavior. He described the correlations as "modest," "not huge," but opined that they can predict societal effects.[73] With respect to videos or other media informing persons about implicit bias, Dr. Greenwald testified that "It raises awareness of implicit bias. It may increase motivation to try to do something about it, but it doesn't tell you what to do about it."[74] Professor Greenwald had read the Washington Supreme Court opinion and commented on the difficulty facing the court on remand:

> The thing that I attended to most was the State Supreme Court decision that is producing this hearing because I am just intrigued by the difficulty of the question that the Supreme Court is bringing up in the challenge they are propose -- confronting the court with to decide whether implicit bias is operating in a juror's -- in a jury. A very difficult question. . . . as impressed with the Supreme Court's reference to the idea of reasonably objective observer, I'm not sure that's the exact phrase, someone who is aware of the implicit bias concept of this bias. Do they think that it's going on. You know, I am an objective observer in terms of the way the Supreme Court decision described what is wanted. This is not an easy decision for me.[75]

[70] Briefly, "The IAT measures the strength of associations between concepts (e.g., black people, gay people) and evaluations (e.g., good, bad) or stereotypes (e.g., athletic, clumsy). The main idea is that making a response is easier when closely related items share the same response key. When doing an IAT you are asked to quickly sort words into categories that are on the left and right hand side of the computer screen by pressing the "e" key if the word belongs to the category on the left and the "i" key if the word belongs to the category on the right." About the IAT (harvard.edu).

[71] "it's just a measure of how much you -- how much faster you go in one version than the other version of the task. If you are at equal speed, and what we call these two combined tasks you have no automatic racial preference. If you are faster when you have to give the same response to white and pleasant with one hand, and black and unpleasant with the other we say you have an automatic white preference. And if you are faster than the other version we would say you have automatic black preference." 8/5/20 Tr. at 88.

[72] *Id.* at 91.

[73] *Id.* at 94.

[74] *Id.* at 96.

[75] *Id.* at 104-105.

Electronic copy available at: https://ssrn.com/abstract=4457050

5. The implicit bias video on remand

No implicit bias video was shown at trial. Each juror was shown an implicit bias video before testifying in the remand hearing. Jurors varied in their reactions. No juror testified that watching the video before deliberations would have affected their decision.[76]

6. The order granting the motion for new trial.

The trial court held that Juror 6's affidavit and testimony established a *prima facie* case that implicit bias was "a factor in the verdict."[77] With respect to the *prima facie* showing, the court held that "Juror 6 ended up internalizing the assertion that she was "partial"[78] and that "Juror 6 changed her vote and found the defendant Guilty because, having been tarred with the label of "partial," she began questioning her own right to participate in jury deliberations."[79]

Berhe specified the standard for assessing a *prima facie* case but not for a subsequent evidentiary hearing. Drawing on *Jackson*, the trial court held the state bore the burden of rebutting the *prima facie* showing that implicit bias was "a factor" in the verdict.[80] The trial court was right that the Supreme Court thought the State should bear the burden of proof.

The *Jackson* court envisioned a remand hearing where the juror accused of making racist comments would be examined and the trial court would determine "whether or not juror X, in fact, held a racial bias such that he could not have decided the case fairly and impartially."[81] The State thus bore the burden of showing the juror could have acted fairly notwithstanding his alleged statements. That is a different question then the one asked in *Berhe*—whether implicit bias was "a factor." The two might be harmonized with a bit of effort but, to take an obvious example, evidence of guilt might be so overwhelming that even a racist juror "could" act impartially because the evidence was a sufficient condition for the verdict. If the state could prove that, then it might, in theory, prevail in a hearing under the *Jackson* standard. *Berhe*'s "a

---

[76] For example, Juror 13, who testified that her father was Hispanic, thought the video made good points and would be good to show jurors before deliberations. She thought it might possibly have affected how she related to other jurors. She also testified that implicit bias did not play a role in deliberations: "Q. Let me ask you this. After watching the video on implicit bias, do you think it's possible that implicit bias could have played a role in this jury's decision? A. No." 7/24/20 Tr. at 45 (Linda Schmidt).

[77] Order at 1. On this point the court held:

> It is reasonable to presume that the juror(s) who stated this may have improperly believed at some level that Juror 6 could not or should not participate equally in the deliberative process because her race biased her. This is reflected both in the juror(s) statements that she was "partial," and also in the jurors' calling her contributions to jury deliberations "stupid" and "illogical" – derisive remarks that were not attributed to other jurors who questioned the defendant's guilt. More importantly, Juror 6 changed her vote and found the defendant Guilty because, having been tarred with the label of "partial," she began questioning her own right to participate in jury deliberations.

Id. at 4.

[78] Order at 4.

[79] *Id*.

[80] *Id*.

[81] State v. Jackson, 75 Wash.App. 537, 544 (1994).

Electronic copy available at: https://ssrn.com/abstract=4457050

factor" standard is different. Implicit bias might be "a factor" if only because it presumably influences a juror's disposition upon encountering evidence,[82] even if at the end of the day the evidence provided a sufficient explanation for a vote.

The trial court granted the new trial motion. It concluded that Juror 6 did not vote under duress[83] but that her will was nevertheless overcome.[84] It is not clear that this conclusion was necessary to grant the motion—implicit bias could be "a factor" even if Juror 6's will was not overcome—but the trial court may have sensed that some form of causation would be necessary both to explain its decision and to give the inquiry a purpose not already captured by the Supreme Court's presumption that implicit bias is pervasive. The court concluded that the record supported a conclusion that the foreperson went out of his way to ensure that everyone was heard, including Juror 6, the jurors discussed the evidence multiple times at Juror 6's request, and that after Juror 6 changed her vote to guilty the foreperson asked if she was sure, and indicated she would be supported if she was not.[85] The record could also support the conclusion, not reached by the court, that the derogatory comments ("stupid" "illogical" and "partial") and the eye-rolling came predominantly from one or two jurors, one of whom apologized for her statements.[86] That might or might not have mattered to the court.

The court's ruling recited implicit bias premises but there was a *realpolitik* element to its ruling as well:

> The court is mindful of the significant resources that has gone into prosecuting the defendant and getting to this verdict, and of the civic duty performed by each of the jurors. This was a five-week trial. Each juror made a significant sacrifice to serve on this jury. In the words of the presiding juror: "It was a huge weight. It felt terribly heavy. I lost sleep worrying about doing the right thing. It was very difficult." The jurors worked diligently and in good faith to reach a just verdict in this case. However, implicit bias can arise where people are acting in good faith. Implicit racial bias can occur unbeknownst to the perpetrator of the bias, and it can be perpetuated by good people of good intentions because it is the product of our culture. According to the defense expert, it can be very difficult to tease out whether conduct is the result of implicit bias; moreover, even those who are committed to anti-racism can unknowingly perpetuate implicit bias. As the State acknowledges, implicit bias is all around us. Nonetheless, it is the responsibility of the court to ensure that a jury verdict is not a result of implicit bias because otherwise the verdict does not have legitimacy. Legal decisions are not made in a social vacuum. They are a product of our time. And at this time the Washington Supreme Court has directed the entire legal community, including

[82] It might, for example, hasten a juror's arrival at the presumably inevitable conclusion.
[83] The State had argued for a duress standard. Order at 6. The trial court rightly held that a duress standard could not be squared with the Supreme Court's opinion in *Berhe*.
[84] Order at 6.
[85] *Id*. at 5.
[86] New Trial Motion Exh. C at 14-15, 32-33. The terms were used at least two to three times by one or two. *Id*. at 33.

Electronic copy available at: https://ssrn.com/abstract=4457050

> judges, to set a course toward dismantling systemic racial injustice in the legal system. *See, Letter of Washington Supreme Court,* June 4, 2020. . . .
>
> the court is mindful of the anguish that reversing a conviction will cause the family of the victim. However, the court is bound by an obligation to apply the law to the facts of the case and require that the State meet its burden as set forth by Washington appellate courts in *Berhe*, *supra* and Jackson *supra*.
>
> Having considered all of the evidence, the court finds that the State has not met its burden of rebutting the Defendant's showing that implicit racial bias was a factor in the jury verdict.[87]

D. Inferences supported by the record on remand.

The testimony recounted above I supports five conclusions, some of which the trial court acknowledged and some of which it did not. They are:

1. Other jurors discounted or disregarded at least some of Juror 6's contributions, and at least one made dismissive comments to her. It was reasonable for her to feel other jurors shut her down.
2. At least most other jurors consciously thought that in doing so they were following the court's instructions to base their decision on the evidence in the case rather than external considerations. They thought Juror 6 was not doing that.
3. Some of Juror 6's contributions could not be tied off to such evidence; some were based on personal experience alone.
4. It was plausible that at least the juror who accused Juror 6 of being "partial" was influenced by Juror 6's race.
5. It was plausible that Juror 6 was partial, at least in the sense that she believed that Black defendants are treated unjustly in the criminal justice system, and she strongly desired not to perpetuate such treatment.

The trial court accepted and emphasized the first and fourth conclusions. It ignored evidence relating to the second, notwithstanding consistent testimony from other jurors that they were trying to follow the rules, and it did not analyze the substance of Juror 6's contributions, even where testimony agreed on what she said. The trial court therefore assessed the reactions of other jurors without assessing what they at least said they were reacting to. The court sidestepped that issue by resort to the implicit bias concept. The court did not address evidence bearing on the fifth conclusion—one Juror 6 raised (and then discounted) as a possibility.

Part III argues that if *Berhe*-style hearings are meant to be a causal inquiry the trial court should have considered evidence bearing on each point and that courts should do so in future hearings of this kind. Because the implicit bias concept was the ostensible hinge on which the trial court's approach turned, however, I first discuss whether that concept can bear the weight

[87] Order at 7-8.

Electronic copy available at: https://ssrn.com/abstract=4457050

the court gave it. In doing so I do not question the importance of the concept for social policy in general, but only whether it contributes more noise than signal to such hearings.

II

Judicial embrace of a form of implicit bias premises, and judicial willingness to entertain expert testimony in *Berhe* and other cases, makes what psychology has to say relevant to law, subject always to the fact that law is an applied discipline. Its needs may distort contributions from other fields. This section surveys psychological debates relevant to *Berhe* with what is hopefully appropriate modesty for an amateur's survey of the work of experts in a different field.

The specific question is how far implicit bias data tell courts in such situations what needs to be done and how to do it. In simpler terms, how far the data recited in the record justify the result in *Berhe.* This Part largely cites the work of scholars who might be termed IAT optimists—they believe the concept and data such as derive from the IAT are sound and useful. (There are pessimists as well,[88] but I largely omit their arguments in this paper because I think its conclusions hold even if one is an optimist.) This Part reaches a simple and possibly obvious conclusion. For inquires such as the remand hearing in *Berhe,* IAT-style data add little to a folk psychology assumption that culture affects perception and perception influences behavior, subject to mediation by contextual factors. Both the common-sense assumption and IAT-style data are too blunt to draw needed distinctions between verdicts that should be upheld and those that should be overturned.

A. Should the IAT be used to pick jurors or qualify trial counsel or witnesses?

A baseline for assessing the utility of IAT data in assessing deliberations may be derived from the consensus view of IAT optimists that the IAT should not be used for tasks as specific as screening jurors. Implicit attitudes vary in part as a function of environment "such as group membership, political orientation, and geography" and are also "sensitive to relatively transient contextual features, such as shifts in the immediate environment in which the test is taken."[89] Thus, [a]s with any measure of attitudes or other individual differences, the IAT score is not a static measure of disposition: It is an adaptive response produced in a particular situation by an organism with a particular biology, personality, and cultural history"[90] IAT proponents thus "have always advised against using a single intergroup IAT as a device for the selection of people, such as whether to hire someone for a job or admit them to a club. The measure is of value in two contexts: research and education."[91]

---

[88] For a concise summary of this view, *see* Gregory Mitchell & Philip Tetlock, *Popularity as a Poor Proxy for Utility: The Case of Implicit Prejudice*, in PSYCHOLOGICAL SCIENCE UNDER SCRUTINY: RECENT CHALLENGES AND PROPOSED SOLUTIONS, Wiley Blackwell, 164–195 (2017).
[89] Kurdi et al. (2018), *Relationship Between the Implicit Association Test and Intergroup Behavior: A Meta-Analysis,* Am. Psychologist 74 (emphasis added).
[90] *Id.*
[91] *Id.*

Electronic copy available at: https://ssrn.com/abstract=4457050

If the scores should not be used *ex ante* for selection, it is not clear how they can be used *ex post* to characterize deliberations. Professor Greenwald's testimony appeared designed to establish a sort of base rate for bias in the population in general. But the contextual differences between the IAT and juror deliberations are greater than, for example, the differences between the IAT and an employment interview. It is no criticism of the IAT as a research tool to point out that it does not try to measure group behavior based on a common pool of evidence under instructions on how to assess that evidence, under conditions in which each juror has the power to force a specific result—a mistrial. It is unlikely that a scalable test could be designed for such a purpose, but the scalability of the IAT and other implicit measures comes at the price of abstraction and the omission of contextual factors, including those present in jury rooms.

B. Implicit association or implicit racism/sexism?

A recent paper by prominent IAT optimists provides a useful baseline for assessing the law's use of the concepts. *Implicit-Bias Remedies: Treating Discriminatory Bias as a Public-Health Problem*,[92] by Professor Greenwald and colleagues, including law professor Jerry Kang, specifies and corrects certain misunderstandings about implicit bias premises.

Professor Greenwald and his colleagues begin by noting that within a few years of the publication of the IAT, its most active developers "stopped using the words `prejudice' and `racism' to describe the implicit biases that could be measured with the IAT."[93] The point leads to the authors' first correction, which clarifies that the IAT assesses "associative knowledge about groups, not hostility towards them."[94] Referring to early descriptions of the IAT, the authors state "it was a mistake to equate implicit bias with racism or prejudice."[95] This point relates to the modest basis for believing that IAT measurements of implicit bias can reliably predict behavior of concern in any particular context, a point discussed below.[96]

This corrected interpretation of IAT data is not what the *Berhe* Court had in mind when it invoked the concept of implicit bias. When *Berhe* used the term, it referred to unconscious negative attitudes that had some causal influence on a verdict. Its concern was "alleged racial bias that deprives a defendant of his or her constitutional right to a fair trial by an impartial jury."[97] The IAT is not the only measure of implicit bias, of course, but it was the one invoked by Dr. Greenwald in the remand hearing and is the source of the concept's general notoriety. To the extent *Berhe* presumes that "bias" means something other than a willingness to associate persons with certain categories, its holding exceeds the reach of at least IAT data.

---

[92] Greenwald et al, *Implicit-Bias Remedies: Treating Discriminatory Bias as a Public-Health Problem*, 23 Psychological Science in the Public Interest 7 (2022).
[93] *Id*. at 8.
[94] *Id*. at 9.
[95] *Id*.
[96] Dr. Greenwald's expert testimony used the term "preference," which in context is a type of association.
[97] 443 P.3d at 1174.

Electronic copy available at: https://ssrn.com/abstract=4457050

C. Does asking about bias work?

One question relevant to *Berhe* in particular, but more generally as well, is whether undesirable bias can be discerned by asking questions. The trial court in *Berhe* thought not. It thought that bias is unconscious then asking questions will do no good and may do harm by provoking defensive answers that obscure the truth.

IAT optimists point to some data indicating that IAT scores are better predictors of behavior than are more explicit approaches, such as subject self-reporting.[98] Other researchers find the difference immaterial,[99] or argue that implicit behavior evidence could be probative only of "behaviors performed under similar processing conditions (i.e., unintentional behavior resulting from low deliberation) compared with behaviors performed under dissimilar processing conditions (i.e., intentional behavior resulting from high deliberation)," as to which "explicit measures should be more likely to predict behaviors . . . ."[100] More impressionistically, some research suggests people can predict how they will do on IAT tests, calling into question whether the associations it identifies are subconscious.[101]

Professor Greenwald and his colleagues report that measures of implicit bias and explicit bias correlate positively, though the correlation is weaker for intergroup issues than for other issues.[102] Intuitively this means people do not as readily admit biases on issues such as race or gender than on (for example) consumer preferences. But the positive correlation suggests that some portion of a person's bias (to indulge for a moment in the assumption that measurement

[98] Professor Jost argues that "[u]ntil an obviously superior technology for measuring implicit attitudes is devised, researchers are justified in employing the IAT—for it continues to outperform the available alternatives in terms of reliability and validity." John T. Jost, *The IAT Is Dead, Long Live the IAT: Context-Sensitive Measures of Implicit Attitudes Are Indispensable to Social and Political Psychology*, CURRENT DIRECTIONS IN PSYCHOLOGICAL SCIENCE 2019, Vol. 28(1) 10–19. IAT critics have argued that it does not do better than explicit measures in general but they appear to agree there is a difference with respect to intergroup associations, the metric most relevant in *Berhe*.

[99] Franziska Meissner, Laura Anne Grigutsch, Nicolas Koranyi, Florian Müller, and Klaus Rothermund, *Predicting Behavior With Implicit Measures: Disillusioning Findings, Reasonable Explanations, and Sophisticated Solutions*, Front. Psychol. 10:2483 (the "predictive value for behavioral criteria is weak and their incremental validity over and above self-report measures is negligible.").

[100] Bertram Gawronski, Six Lessons for a Cogent Science of Implicit Bias and Its Criticism, Perspectives on Psychological Science 2019, Vol. 14(4) 574–595.

[101] Hahn, A., & Goedderz, A. (2020), *Trait unconsciousness, state unconsciousness, preconsciousness, and social miscalibration in the context of implicit evaluations*. Social Cognition, Supplement, 116–135. Psychologists debate what it means to call a bias "implicit," a term they use in preference to "unconscious." For example, the term could refer to mental processes, in which case it could be close to a synonym for "unconscious," but it also could refer to the test mechanism of, for example the IAT, which is implicit in the sense that it does not pose direct questions about comparative attitudes but draws inferences from a procedure thought to shed light indirectly on such attitudes. One recent paper suggests there is no consensus definition of implicit bias. Kate A. Ratliff and Colin Tucker Smith, *Implicit Bias as Automatic Behavior*, 33 PSYCH. INQUIRY 213 (2022). Arising from the context of peremptory challenges, Washington's use of the term could be defined functionally—an "unconscious" bias is one that is obscure enough for a lawyer to offer a race-neutral explanation for conduct without lying. That concept would include things such as lawyers talking themselves into "believing" a proffered race-neutral justification, an idea that has no analogue in conventional definitions of implicit bias.

[102] *Id*.

Electronic copy available at: https://ssrn.com/abstract=4457050

could be precise) might be detected by asking questions. The legally salient issue would then be, for any given person, how much bias remains undetected by such means.

The point may be significant in deciding whether, for example, jurors could be asked directly whether they did or said anything race-related during deliberations or saw anyone else do so. The court in *Berhe* chided the trial court for not taking control of the inquiry when the issue was first raised, which allowed the prosecution to pose direct questions regarding jurors' motivation. The court worried that asking a direct question might prompt jurors to become defensive and thus obscure the truth. There is some reason to believe that could happen,[103] though typically courts would rely on cross-examination to clarify obscurant answers. (Professor Greenwald testified that leading questions could influence juror responses.)[104] The trial court allowed explicit questions about whether racial language was used.

On balance the data tend to support but not compel the Washington Supreme Court's skepticism of direct inquiry and its rule that trial courts take control of investigating allegations of impermissible bias. They leave open the question whether more explicit questions during live testimony, where answers are subject to cross, should be permitted.

D. Causation

Washington's General Rule 37, and California's comparable rule, declare as fact that implicit bias in jury selection is causal: along with explicit bias it has "resulted in the unfair exclusion of potential jurors in Washington State."[105] The Court in *Berhe* appeared to hold that this principle is divisible—one can infer from a stipulated aggregate effect that implicit bias is causal in particular cases.

Psychologists have argued about whether and to what extent IAT scores predict behavior, and the tenor of the debate has changed over time. In 2005, when Professor Page wrote his influential article, confidence that IAT scores are probative of behavior was somewhat greater than it is now.[106] Writing in 2019, NYU Professor John T. Jost, an IAT optimist, summarized developments in the field to that point: "As a `bona fide' pipeline used to quantify levels of `unconscious racism' as a fixed property of the individual— or as a diagnostic tool to classify people as `having' racism or sexism (like they might `have' clinical depression)—the IAT is

[103] "We expect that two common features of diversity training--mandatory participation and legal curriculum--will make participants feel that an external power is trying to control their behavior." There are some data to indicate that "in corporate settings, diversity at the managerial level either stayed the same or decreased following mandatory anti-bias trainings, possibly because the interventions activate bias and cause backlash." NIH, Is Implicit Bias Training Effective? at 3 (citing Professor Dobbin).
[104] 8/5/20 Tr. at 108.
[105] Wash. Gen. R. 37(f).
[106] Greenwald et al, *supra* note [n] at 14, note three meta-analyses beginning in 2009. A critical assessment of changes in understanding of IAT scores over time may be found in Mitchell & Tetlock, *supra* note [n].

Electronic copy available at: https://ssrn.com/abstract=4457050

dead."[107] There appears to be some consensus that the correlation between IAT scores and tested behavior is "small to moderate," as Professor Greenwald testified.[108] The same is true when IAT scores are used in a jury context. Writing in 2012, Professor Kang and colleagues reported studies finding that jurors harbor small biases against members of other racial groups,[109] and IAT studies suggesting racial differences in how jurors assess evidence (a point for which the remand hearing in *Berhe* is an interesting example) though not in their ultimate votes.[110]

Studies seeking to correlate IAT scores with behavior are only as probative as the behavior they test. What counts as biased behavior in these studies? Here is one example. A 2006 experiment by Professors Amodio and Devine[111] gave psychology students IATs designed to measure race bias.[112] Test subjects were then told that they would be partnered with a student to take a knowledge test. Information provided to the subjects indicated that the other student was Black. The subjects were then taken to a hallway in which identical chairs were spaced an identical distance from each other. A coat and backpack appearing to belong to the (Black) partner were placed on the nearest chair. The measure of racially aversive behavior was how close a subject sat to the coat and backpack, with distance correlating positively with aversion (in lay terms, racism). Students whose IAT scores indicated relatively strong negative associations (greater microsecond differences between Black and White associations) tended to sit farther away than students with lower bias scores. From a lay perspective and taking the early characterization of IAT scores as showing prejudice (rather than mere association), the study implies that students who sat farther away from the objects were relatively aversive. The experimental design avoided the risk that idiosyncratic factors (clothing, hairstyle, etc.) associated with using an actual person instead of a backpack would introduce noise into the

[107] John T. Jost, *The IAT Is Dead, Long Live the IAT: Context-Sensitive Measures of Implicit Attitudes Are Indispensable to Social and Political Psychology*, CURRENT DIRECTIONS IN PSYCHOLOGICAL SCIENCE 2019, Vol. 28(1) 10–19.

[108] Greenwald et al at 14. The authors report that the combined correlation coefficient derived from three meta-studies was 0.165. *Id*. at 11. For a skeptics' summary, *see* Franziska Meissner, Laura Anne Grigutsch, Nicolas Koranyi, Florian Müller, and Klaus Rothermund, *Predicting Behavior With Implicit Measures: Disillusioning Findings, Reasonable Explanations, and Sophisticated Solutions*, Front. Psychol. 10:2483.

> With implicit measures like the IAT, researchers hoped to finally be able to bridge the gap between self-reported attitudes on one hand and behavior on the other. Twenty years of research and several meta-analyses later, however, we have to conclude that neither the IAT nor its derivatives have fulfilled these expectations. Their predictive value for behavioral criteria is weak and their incremental validity over and above self-report measures is negligible.

[109] Kang et al, *Implicit Bias in the Courtroom*, 59 U.C.L.A. L. REV. 1124, 1143 (2012). The effect size for guilt was .092 on a scale in which .2 would be considered small. The cited studies measure individual juror reactions and therefore are not at odds with finding that the presence of even one or two Black jurors tended to equalize conviction rates between Black and White defendants. Shamena Anwar, Patrick Bayer, and Randi Hijalmarsson, *The Impact of Race in Criminal Jury Trials*, The Quarterly Journal of Economics (2012) 127, 1017–1055.

[110] Kang, *supra* note [n] at 1145-46.

[111] Referenced as typical in Kurdi et al, *The IAT and Intergroup Behavior*,

[112] The point of these experiments was in part to try to distinguish stereotyping, which might indicate familiarity with how society views persons of different races, from evaluative race bias, meaning an aversion or attraction to persons of a race.

Electronic copy available at: https://ssrn.com/abstract=4457050

procedure, but one could fairly question how willing courts should be extrapolate from this experiment to any practical trial application.[113]

IAT defenders respond that even small to modest correlations may imply socially significant consequences when extrapolated to a large population.[114] Even accepting that small effects may aggregate, however, it does not follow that effects present in the aggregate are present in each case. To believe in aggregate effects does not tell judges what to do in a specific case. For example, Professor Kang and his colleagues reference studies finding that juries of one race have higher conviction rates for defendants of other races than for same-race defendants. The correlation size is low, a point the authors address by arguing that, over 100 verdicts, the data could imply that all-White juries would convict eight more Black defendants than White defendants.[115] Suppose that is true (it is certainly plausible). The *ex post* problem addressed in *Berhe* is: which eight? IAT data do not help answer that question.

The same point holds for some examples comparing IAT scores to behavior some people find prudent. Professor Kang, for example, argues that some people may take aspirin to reduce the risk of heart attack even though the data on that correlation (r=0.02) is comparably low to the data linking IAT to behavior.[116] For situations such as the remand hearing in *Berhe,* the analogy is imperfect because the aspirin/heart attack correlation points to a discrete act—taking aspirin—which is invariant to contextual factors such as whether aspirin is taken at home, in the office, or in court, and because heart attack is a more precise correlate than teasing out the import of juror comments. To believe in a weak correlation between implicit bias and behavior in general does not yield a discrete behavior to follow in court, a problem compounded by the ineffectiveness of de-biasing techniques, discussed below.[117]

In contrast, the aspirin example illustrates that in some cases one might rationally be indifferent to causal mechanisms. On this point Professor Kang's analogy is apt for some aspects of jury selection. Studies show that having even one Black juror may largely eliminate racial differences in conviction rates.[118] One could posit several theories about why that might be

---

[113] This experiment included data from 21 students. @ 657.

[114] Greenwald et al, *supra* note 1, at 17. The original work on this point may be found in simulations reported in Greenwald, A. G., & Banaji, M. R., *Statistically Small Effects of the Implicit Association Test Can Have Societally Large Effects*, Journal of Personality and Social Psychology 2015, Vol. 108, No. 4, 553–561:

> [P]roblems of limited test–retest reliability and small effect sizes are maximal when the sample consists of a single person (i.e., for individual diagnostic use), but they diminish substantially as sample size increases. Therefore, limited reliability and small to moderate effect sizes are not problematic in diagnosing system-level discrimination, for which analyses often involve large samples.

[115] Kang et al, *Implicit Bias in the Courtroom*, 59 U.C.L.A. L. REV. 1124, 1143 (2012).

[116] Jerry Kang, *Judicial Behavioral Realism about Implicit Bias*,

[117] The analogy might justify broader measures with lower stakes for individuals, such as training requirements, provided that one was confident training at least did no harm.

[118] Shamena Anwar, Patrick Bayer, and Randi Hijalmarsson, The Impact of Race in Criminal Jury Trials, The Quarterly Journal of Economics (2012) 127, 1017–1055:

Electronic copy available at: https://ssrn.com/abstract=4457050

true, but for practical purposes the mechanism does not matter. The important point is that it does appear to be true, and the observation yields a concrete step that courts may implement.[119] In this example, it does not matter whether implicit bias reasoning can assist in *ex post* analysis, and it does not matter whether the balancing effects of at least a minimally diverse jury have anything to do with implicit bias. What matters is that the data provide a sound basis for doing something that is feasible to do.

E. The ineffectiveness of de-biasing techniques

Professor Greenwald and his colleagues provide a skeptical summary regarding attempts to eradicate implicit bias: "[w]ith only occasional exceptions, experimental attempts to reduce long-established biases have not found that they are durably modifiable."[120] They add that "[s]cholarly reviews of the effectiveness of group-administered antibias or diversity-training methods have not found convincing evidence for their mental or behavioral debiasing effectiveness."[121] A skeptical meta-analysis of studies of implicit bias remedies noted that "[t]he modal intervention uses mentalizing as a salve for prejudice,"[122] which is one way to describe jury instructions that caution jurors to be wary of unconscious bias.[123]

In view of such general comments, it is not surprising to learn that mock jury studies have found that implicit bias instructions do not affect jurors' results.[124] UC Irvine Professors

---

> the presence of even one or two blacks in the jury pool results in significantly higher conviction rates for white defendants and lower conviction rates for black defendants. Specifically, in cases with no blacks in the Jury pool, black defendants are convicted at an 81% rate and white defendants at a 66% rate. When the jury pool includes at least one black potential juror, conviction rates are almost identical: 71% for black defendants and 73% for white defendants.

See also Kang et al, *supra* note [n] at 1180-1181.

[119] Setting aside questions of positive law, one could, for example, require that all parties have at least one juror of their race or gender on each jury unless to do so would be demonstrably infeasible. Concerns of intersectionality complicate this recommendation—should one assume that a Black woman juror such as Juror 6 is satisfactory in a case involving a Black male defendant, such as Mr. Berhe—but dealing with implementation at that level would be preferable to ignoring the aspirin analogy altogether. Interestingly, a version of this lesson was implemented in *Berhe*, where the parties agreed with the Court's suggestion that Juror 6 be excluded from the random draw used to choose alternates, on the explicit ground that it would be unfortunate if the only Black juror were selected as alternate. State v. Berhe, 193 Wash.2d 647, 651 (2019). For an influential recommendation of such a right, *see* Sheri Lynn Johnson, *Black Innocence and the White Jury*, 83 Mich L Rev 1611, 1695-1700 (1985)(proposing right to jury including racially similar jurors).

[120] Supra note [**n**] at 10.

[121] *Id*. at 11.

[122] Paluck at 533.

[123] *Supra* note [n]. "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 can affect how we evaluate information and make decisions."

[124] Elek, Jennifer and Agor, Paula Hannaford, 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 ("The present study found no significant effects of the instruction on judgments of guilt, confidence, strength of the prosecution's evidence, or sentence length."); Mona Lynch & Emily Shaw, Can Jury Instructions Have an Impact on Trial Outcomes? https://www.ojp.gov/pdffiles1/nij/grants/300717.pdf ("Our quantitative findings indicated no differences in verdict

Electronic copy available at: https://ssrn.com/abstract=4457050

Mona Lynch, Taylor Kidd, and Emily Shaw state "even though changes to jury instructions that promote bias awareness sound commonsensical and harmless, the social- psychological phenomena underpinning this intervention are more complex and have the potential to produce unanticipated effects in the group decision-making process." They report a recent study in which they "did not find that the implicit bias jury instruction exerted influence on verdict outcomes when compared with the standard instruction condition."[125] The researchers' qualitative assessment was mixed. They found that groups who heard implicit bias instructions were more likely to discuss bias than groups who did not, an observation that suggested the instructions might affect deliberations even if not results. On the other hand, "the concept of bias was elastic enough in these deliberations to sometimes impede other aspects of the jurors' duties in a manner that could potentially harm defendants, especially in the implicit bias instruction conditions."[126] In a number of instances jurors invoked implicit bias instructions to truncate discussion of whether a prosecution witness (snitch) might be suspected of presenting biased or unreliable testimony.[127]

Similarly, another recent study found that the implicit bias video and instruction from the Western District of Washington had no effect on mock jurors' assessment of guilt or culpability. The study found that White and Black jurors reached similar results for White defendants, but White jurors were more lenient than Black jurors toward Black defendants. The study found that the video affected Black jurors' assessment of witness credibility.[128]

III

Berhe raise the question of what is permissible in jury deliberations, relative either to the goal of producing a verdict based on the record or to the goal of making jurors feel safe and respected as jurors. Those two goals may align, but they may not, a potential tension ignored in *Berhe*. Related to each question is how to assess juror allegations of implicit bias, and whether IAT data or other implicit measures provide the right analytic framework for assessing those questions.

This part argues that implicit bias premises did no analytic work in *Berhe* and are not capable of assessing whether bias has exerted a causal force—at whatever level ("a factor" or something else)—on a verdict. To the extent such an inquiry focuses on causation, it is better conceived of as an attribution problem. To the extent such an inquiry aims to ensure that courts are not hostile to persons such as Juror 6, implicit bias concepts will not prove useful, either. Both in the peremptory challenge context and in *Berhe* the Washington courts' use of implicit

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.").

[125] The subtle effects of implicit bias instructions, LAW & POLICY Jan. 2022, https://onlinelibrary.wiley.com/doi/ftr/10.1111/lapo.12181.

[126] *Id*.

[127] *Id*. The defense strategy in the case presented to test subjects entailed an attack on the credibility of an informant testifying for the prosecution.

[128] Christine L. Ruva, Elizabeth C. Sykes, Kendall D. Smith, Lillian R. Deaton, Sumeyye Erdem & Angela M. Jones (2022): *Battling bias: can two implicit bias remedies reduce juror racial bias?*, Psychology, Crime & Law.

Electronic copy available at: https://ssrn.com/abstract=4457050

bias rhetoric serves to give judges leeway to pursue results they deem just. That may turn out to be a good thing; it depends on the pursuit. But implicit bias plays no analytic role in it.

A. What is permissible in the jury room?

In *Pena-Rodriguez v. Colorado*,[129] one of the precedents cited by the Washington Supreme Court in *Berhe*, Justice Kennedy posited that "fair and impartial verdicts can be reached if the jury follows the court's instructions and undertakes deliberations that are honest, candid, robust, and based on common sense."[130] This happy portrait does not tell courts what to do when following instructions puts a juror at odds with what they see as common sense. Some parameters have been established, however. Differences among them give lawyers an incentive to frame juror disputes as implicit bias issues.

1. The general duress standard.

Juror threats amounting to duress may justify an order granting a new trial but the standard for duress is strict. Two years after *Berhe*, for example, the Washington Court of Appeal denied a new trial motion supported by a juror's claim that they felt threatened when another juror said: "karma should come back at me, and someone should come to my house and do that to me, and [juror X] hopes that I am the next person that that happens to if I don't agree with [them]."[131] Surveying precedent from other jurisdictions, the court held that the party alleging juror misconduct has the burden of showing misconduct and must overcome "the policy favoring stable and certain verdicts and the secret, frank and free discussion of the evidence by the jury" and must show "juror misconduct has prejudiced the defendant."[132] The court held:

> a juror commits misconduct only if the alleged coercive acts rise to the level of actual or threatened physical violence or abuse. But mere expressions of frustration, temper, empty threats, and strong conviction against the contrary views of another panelist are insufficient to establish a claim of juror misconduct.[133]

Standards from other states include Colorado's requirement that juror misconduct occurs "only if the alleged coercive acts [first] rise to the level of continuous violent, abusive, and profane language and conduct threatening or amounting to physical violence against a juror."[134] The record must show "more than expressions of frustration, impatience, annoyance, or empty threats."[135] Minnesota will find misconduct based on threats of physical violence but not

[129] 580 U.S. 206, 211 (2017).
[130] *Id*.
[131] *State v. Hill*, 19 Wash. App. 2d 333, 338 (2021), review denied, 199 Wash. 2d 1011, 508 P.3d 675 (2022).
[132] *Id*. at 342.
[133] *Id*. at 344. *Hill* issued September 28, 2021. The order granting a new trial in *Berhe* issued December 4, 2020.
[134] People v. Mollaun, 194 P.3d 411, 418 (Colo. App. 2008).
[135] People v. Rudnick, 878 P.2d 16, 22 (Colo. App. 1993).

Electronic copy available at: https://ssrn.com/abstract=4457050

"[e]vidence of *psychological* intimidation, coercion, and persuasion,"[136] which is inadmissible in Minnesota. Oregon will find misconduct based on any conduct that would subject a juror to prosecution for obstruction of justice, such as coercion.[137] California declined to order a new trial in a case where there was hearsay evidence from a defense investigator that one juror shouted to another: "`If you make this all for nothing, if you say we sat here for nothing, I'll kill you and there'll be another defendant out there—it'll be me.'"[138] A juror affidavit stated that the elderly woman at whom the statement was shouted began crying and shaking and went to the bathroom "where I believe she vomited."[139] The California Supreme Court viewed the statement as one no reasonable person could take literally and "an expression of frustration, temper, and strong conviction against the contrary views of another panelist."[140]

2. The standard for overtly racist conduct by a juror.

Berhe, and both *Pena-Rodriguez* and *Jackson* before it, drew on history to hold that race is different. As the Court in *Pena-Rodriguez* said, relative to the institution of the jury "racial bias implicates unique historical, constitutional, and institutional concerns."[141] This distinction predates *Berhe*, as noted above, but historically strong evidence of overtly racist behavior was required to inquire into deliberations. The standard set in *Pena-Rodriguez* required "a showing that one or more jurors made statements exhibiting overt racial bias that cast serious doubt on the fairness and impartiality of the jury's deliberations and resulting verdict. To qualify, the statement must tend to show that racial animus was a significant motivating factor in the juror's vote to convict."[142]

As noted in Part I, *Pena-Rodriguez* dealt with a claim that a juror's racist conduct showed bias against the defendant and thus tainted the juror's vote. It did not deal with a case such as *Berhe*, in which the claim was of racial bias against a juror rather than a defendant. More importantly, because the *Pena-Rodriguez* standard requires evidence of overtly racist conduct, a court inclined to apply that standard to a claim of inter-juror bias could distinguish a motion based on that ground from a motion based on the more forgiving standard for duress. An allegation of racist conduct would have to be proved factually at some level but, if proved, the allegation would be sufficient to invoke the stricter standard.

3. The standard for implicit bias claims.

Berhe's standard for granting a new trial motion based on implicit racial bias—whether a complaining party can make a *prima facie* case that a person believing in implicit bias *could* view bias as being a factor in a verdict and, if so, whether the prevailing party could then

[136] State v. Jackson, 615 N.W.2d 391, 396 (Minn. App. 2000)(emphasis added).
[137] Hill v. Lagrand Indus. Supply Co., 193 Or. App. 730, 735 (2004).
[138] People v. Keenan, 46 Cal. 3d 478 (1988).
[139] *Id*. at 540.
[140] *Id*. at 541.
[141] Pena-Rodriguez v. Colorado, 580 U.S. 206, 224 (2017).
[142] *Id*. at 225.

Electronic copy available at: https://ssrn.com/abstract=4457050

demonstrate that such bias was not "a factor" in a verdict--is more lenient than the standard for overt bias and significantly more lenient that the general duress standard. Shifting the burden from a party asserting bias to the party who prevailed may be decisive in many cases.

Unlike the explicit bias cases, the implicit bias standard can be triggered without evidence of racial remarks. The line between a case to which the duress standard should apply and a case in which the *Berhe* standard should apply will be harder to discern than the line between *Berhe* and *Pena-Rodriguez. Berhe* makes the *prima facie* showing the test of which test applies. If such a showing is not made, the duress standard applies. If it is made, the burden-shifting inquiry in *Berhe* applies. It is significant, therefore, that the showing in *Berhe* consisted largely of Juror 6's attribution of jury-room-conduct to what she saw as the biased dispositions of other jurors, including use of the key term "partial." The large difference between *Berhe* and the duress standard will create an incentive for defendants to frame juror disputes in implicit bias terms. I discuss the dynamic potential of *Berhe* in Part III(f) below.

B. Fair trial concerns may trade off with juror conduct rules or interests.

In *Berhe* and in a subsequent case extending its reasoning, *Henderson v. Thompson*,[143] the Washington Supreme Court effectively acknowledged that courts historically have been isolating and unwelcoming places for members of certain groups, to the extent courts were not overtly hostile places. The Washington Supreme Court combined this premise with implicit bias arguments to signal that it intends to create courtroom environments safer and more welcome to persons in those groups. This signal was followed in the *Berhe* remand opinion finding that Juror 6 was not subjected to duress but that her own account of her perception of other jurors' motivations was sufficient to place upon the state the burden of disproving race as a factor in the other jurors' conduct. The remand record was not strong enough to show juror misconduct in a conventional sense; legally the burden of proof was decisive.

Two issues arise from shifting the focus of fair trial analysis from parties to jurors. The first is purely conceptual because *Batson* resolved it. If jurors are biased, as IAT data would suggest relative to at least many cases, then juror interests in sitting on a case may trade off with a defendant's interest in a fair trial. This idea is the premise of Justice Thomas's argument that concerns for juror dignity may threaten defendants' right to a fair trial.[144] If peremptory challenges could be calibrated well, the psychological data would at least not exclude a regime in which defendants could make explicitly race-based challenges because doing so might reduce the net bias of a panel. To borrow from the example of juror outgroup bias discussed above, the data would suggest that a Black defendant should be permitted to make explicitly race-based challenges against White jurors so long as doing so increased the probability of seating a Black

[143]518 P.3d 1011 (2022).

[144] Flowers v. Mississippi, 139 S. Ct. 2228, 2243 (2019)(Thomas, J. dissenting)("Instead of focusing on the possibility that a juror will misperceive a peremptory strike as threatening his dignity, I would return the Court's focus to the fairness of trials for the defendant whose liberty is at stake and to the People who seek justice under the law.").

Electronic copy available at: https://ssrn.com/abstract=4457050

juror. That is not the law now,[145] and I am not aware of anyone arguing that implicit bias premises should make it so, but the idea is a fair implication of the premises of *Berhe*.[146]

The second issue is whether there is a relationship between conflict among jurors and fairness and, if so, what it looks like. Conflict is the norm in courts. Conflict is generally presumed to result in—sometimes is almost equated with—fairness.[147] In *Pena-Rodriguez* Justice Kennedy praised "robust" deliberations,[148] and the original trial judge in *Berhe*, who denied the new trial motion, said jurors should be allowed to "press" each other on their positions.[149] A recent Washington Supreme Court opinion holds that "[i]f jurors did not feel free to dissect the credibility and motivations of the witnesses *and each other*, their ability to reach a unanimous verdict would be compromised."[150] As noted above, the general rule for assessing claims of juror duress is permissive. Conduct amounting to criminal assault crosses the line, but anything short of that is not enough to shift a verdict.

Berhe raises the question whether implicit bias findings justify modifying this norm, or whether, if implicit bias theory falls short, modification is desirable based on moral considerations, such as a concern for justice. On the former point, IAT scores and the debates swirling around them are no help. The IAT does not test group interaction, nor does it seek to measure whether arguments based on a common pool of evidence are effective at ameliorating bias or tend to reinforce it. That is not a point against the IAT or the data it generates; it just measures different things. To the extent one was inclined to interpret the IAT as showing that everyone is biased, that conclusion would not explain where the burden of proof should lie or what it should be, since Juror 6 would be presumed biased under that interpretation as well. And

[145] *E.g.* Flowers v. Mississippi, 139 S. Ct. 2228, 2243 (2019)(“A defendant of any race may raise a *Batson* claim, and a defendant may raise a *Batson* claim even if the defendant and the excluded juror are of different races.”); United States v. Walker, 490 F.3d 1282, 1292 (11th Cir. 2007)(pattern of striking White jurors violated *Batson*). For an argument that *Batson* may make Black defendants worse off in the long run by eliminating their chance to exercise peremptory challenges based on race, *see* Georgia v McCollum, 505 U.S. 42, 60 (1992) (Thomas, J, concurring)(“I am certain that black criminal defendants will rue the day that this Court ventured down this road that inexorably will lead to the elimination of peremptory strikes.”).

[146] Professor Lawrence’s retrospective on his foundational article is apt on this point. Describing conversations with liberal White friends who, in essence, sought his exculpation from the charge of racism, Lawrence concluded that exculpation was simply wrong. “My answer to this dilemma was to refuse to issue “good white folks” passes, to say, “Sorry, my friend. You are a racist and so am I.” Charles R. Lawrence III, *Unconscious Racism Revisited: Reflections on the Impact and Origins of "The Id, the Ego, and Equal Protection,"* 40 CONN. L. REV. 931-978 (2008).

[147] A view expressed among other places in Justice Marshall’s dissent in Strickland v. Washington, 466 U.S. 668, 711 (1984)(“Every defendant is entitled to a trial in which his interests are vigorously and conscientiously advocated by an able lawyer. A proceeding in which the defendant does not receive meaningful assistance in meeting the forces of the State does not, in my opinion, constitute due process.”).

[148] *Supra* note [n].

[149] *Supra* note [n].

[150] State v. Norman, No. 100777-9, 2023 WL 1456755, at *6 (Wash. Feb. 2, 2023)(emphasis added).

Electronic copy available at: https://ssrn.com/abstract=4457050

if one were inclined to use IAT data to distinguish between Black jurors and others, such as Whites or Asian-Americans,[151] the positive law problems would be insurmountable.

The moral problem in drawing a line between creating a welcoming environment and preserving debate as a norm of deliberations is that there are fair concerns on each side. As the Washington Supreme Court held in *State v. Norman*,[152] where a juror's self-harm during deliberations posed a risk that other jurors would self-censor, debate can sharpen analysis, including by challenging juror positions. If, as the Washington courts have said, everyone has unconscious bias, and if to be fair trials should be free from such bias, then jurors should be allowed to push each other, to some extent, to uncover bias. If, through the process of deliberation, a juror comes to realize that their reactions are not cashing out in the record, the juror should think about what else might be driving their reactions. From that perspective, Juror 6's testimony that she began to question whether she was being partial is a sign that deliberations were serving their purpose, not a sign that they had gone off the rails.

Other jurors' perspectives are relevant, too. Everett Williams was dead, after, all, shot four times. The case was not a minor drug bust. Other jurors testified, seemingly in good faith, that what they saw was one of their number make conjectures not grounded, or only weakly grounded, in the evidence (the cops planted the gun; Berhe may have got the gun from the real killer), refer to personal matters not in the record (acquaintances unfairly charged) as a reason for reluctance to convict, and substitute what they may have seen as a stereotype—members of hip-hop culture often wear baggy pants—for what other jurors saw on tape as evidence—Berhe's pants weren't baggy. Having been instructed to base their votes on the evidence, what should they have done? The traditional answer is that the other jurors were supposed to push back and give reasons grounded in the evidence in doing so. On the traditional view of courts as domains of conflict, that is how the verdict is pushed closer to the record and thus, by assumption, closer to a fair result. The basic hope is that the pressure of logic applied to evidence will strip away

[151] Though Washington has declared that "everyone" has unconscious biases, Professor Greenwald's testimony was more nuanced. He testified that over 80% of Whites and Asian-Americans display a pro-white preference, while Black persons are divided—1/3 display a pro-White preference, 1/3 display no preference, and 1/3 show a preference for Black persons. Using IAT nomenclature, the aggregate score for Black persons on a Black-White IAT is close to 0, which the IAT interprets as unbiased. To the extent the IAT captures only associations, these scores may suggest only that Black respondents are aware of social stereotypes that favor Whites, not that Black respondents disfavor Black persons. More substantive interpretations—that IAT scores reflect "preference," as Professor Greenwald testified, have suggested that Black persons have internalized racial stereotypes, leading them to disfavor their own group at a rate higher than Whites. Such thinking has a long history in law, tracing to the citation in Brown v. Board of Education, 347 U.S. 483, 494 n.11 (1954), to studies including the Clarks' childhood development studies. A current example is system justification theory, which holds that persons find ways to rationalize that status quo and in particular that "members of disadvantaged groups often (but not always) exhibit outgroup favouritism by expressing more positive attitudes about other groups that are higher in status or power than their own group." John T. Jost, *A quarter century of system justification theory: Questions, answers, criticisms, and societal applications*, British J. Soc. Psych 2018 at 15. An interesting unpublished paper by Ulrich Schimmack and Alicia Howard questions the validity of IAT findings on this point. *The Race Implicit Association Test is Biased: Most African Americans Have Positive Attitudes Towards Their In-Group*, The Race Implicit Association Test Is Biased | Replicability-Index (replicationindex.com).

[152] 2023 WL 1456755, at *6 (Wash. Feb. 2, 2023)(emphasis added).

Electronic copy available at: https://ssrn.com/abstract=4457050

bias or sentiment. The Washington Supreme Court's opinion in *Norman* represents that side of the coin.

Berhe represents the other side. To say that jurors should push each other to some extent does not answer the question of what extent is permissible. The premise that conflict is a measure of fairness is hardly self-evident. Lawyers might like jousting and might think argument is good just because lawyers are good at it. Ordinary people might place less weight on verbal dexterity, and lots of people find conflict alienating. A juror who checks out and ignores what sounds like noise might be rationally expressing a limited tolerance for a closed room full of strangers attempting to be amateur forensic logicians. Even on the traditional view, a fair trial should be measured by juror assent, not acquiescence. Fairness is undermined if conflict engenders unwilled acquiescence in peer pressure. There are no purely logical grounds for choosing between rules that favor strong exchanges and rules that try to temper them.

Though the remand hearing in *Berhe* is unprecedented, as far as the author is aware, judges have interviewed jurors during deliberations to respond to juror complaints and decide whether a juror should be discharged. A juror may be discharged for refusing to follow instructions, but the standard is strict and as a practical matter would be extremely strict in *Berhe*, where Juror 6 was the only Black juror.[153] The Second Circuit has adopted a standard a request to dismiss a juror for cause on nullification grounds must be denied if the record discloses any possibility that a complaint about a juror's conduct was based on the juror's view of the sufficiency of the government's evidence.[154]

The strict standard for discharge is based largely on the need to keep deliberations confidential. Washington explained its rule by stating "investigation into a claim that a juror is engaging in nullification risks violation of the cardinal principle that juror deliberations must remain secret."[155] Added to this concern is the risk that judges might make things worse if they asked jurors to explain the thinking behind actions that might be hard to interpret from purely factual descriptions. Inquiry could cause jurors to dig in further to their positions, or to silence themselves, for example.[156] *Berhe's* holding that racial bias by a juror is not protected by the rule

[153] In United States v. Thomas,116 F.3d 606, 614 (2d Cir. 1997), the Second Circuit held that a juror intending nullification could be discharged for cause under Federal Rule of Criminal Procedure 23, and that this rule "does not include an exception for jurors who violate their sworn duty on the basis of racial or ethnic interests or affinities."
[154] Id. at 622, quoting United States v. Brown, 823 F.2d 591, 596 (D.C. Cir. 1987). California treats the issue as one kind of refusal to deliberate and allows discharge of a juror if "it appears as a `demonstrable reality' that the juror is unable or unwilling to deliberate." People v. Cleveland, 25 Cal. 4th 466, 484 (2001). Washington's standard is that "a deliberating juror must not be dismissed where there is any reasonable possibility that the impetus for dismissal is the juror's views of the sufficiency of the evidence." State v. Elmore, 155 Wash. 2d 758, 761 (2005).
[155] Id. at 770.
[156] The *Thomas* court explained the point:

> The mental processes of a deliberating juror with respect to the merits of the case at hand must remain largely beyond examination and second-guessing, shielded from scrutiny by the court as much as from the eyes and ears of the parties and the public. Were a district judge permitted to conduct intrusive inquiries into-and make extensive findings of fact concerning-the reasoning

Electronic copy available at: https://ssrn.com/abstract=4457050

against impeaching verdicts, and the hearings it authorizes upon a *prima facie* showing of implicit bias, raises the question whether the standard for discharging jurors should be altered to facilitate challenges in cases such as *Berhe*—not against Juror 6, who was deliberating, but against the juror who accused her of being partial. I address such a possible extension below.

One might think that in criminal trials a greater concern for juror peace of mind skews in favor of the defendant. The government needs unanimity to convict, and a form of civility rule would tend to make it easier to hold out and hang a jury. Probably that is right, though the opposite result is possible. Suppose a juror was too inclined to believe testimony from a cooperating government witness of the juror's race, such as Mr. Washington,[157] and thus was too inclined to convict, as Professor Lynch found was possible. In that case a failure to press hard on a juror's reasoning would help the government, not the defendant, and would tend to move a verdict away from the record.

C. The ratio of personal experience to evidence in deliberations.

One purpose behind jury selection rules such as Washington's General Rule 37, and the Supreme Court's opinion in *Batson v. Kentucky,*[158] which held unconstitutional purposeful discrimination in the exercise of preemptory challenges, is to ensure juries reflect the diversity of perspectives in a community. *Batson* gets a bad press in the era of implicit bias because its emphasis on deliberate discrimination is considered naïve. But one rationale in *Batson* matches the logic of *Berhe. Batson's* holding rested in part on juror interests, not just the interest of the defendant. "[B]y denying a person participation in jury service on account of his race, the State unconstitutionally discriminate[s] against the excluded juror."[159] The logic of the rationale is as simple as it is compelling—if a verdict is supposed to reflect the judgment of a community, the chance to sit as a juror is a prerequisite to full citizenship in the community.[160]

The remand order in *Berhe* stressed that a juror's personal experiences and perspective are fair game for deliberations. The context was testimony from other jurors that Juror 6 had made comments about persons she knew whom she believed had been treated unjustly by the criminal justice system.[161] The prosecution argued that use of what appeared to be the crucial

---

> behind a juror's view of the case, or the particulars of a juror's (likely imperfect) understanding or interpretation of the law as stated by the judge, this would not only seriously breach the principle of the secrecy of jury deliberations, but it would invite trial judges to second-guess and influence the work of the jury.

116 F.3d at 620.

[157] Juror 6 disbelieved Mr. Washington out of sympathy—she identified with him because she thought he was just saying what police wanted to hear so he could get out of the situation he was in.

[158] 476 U.S. 79 (1986).

[159] *Id*. at 87.

[160] *E.g.* Flowers v. Mississippi, 204 L. Ed. 2d 638, 139 S. Ct. 2228, 2243 (2019)(" Other than voting, serving on a jury is the most substantial opportunity that most citizens have to participate in the democratic process."). Justice Thomas's dissent in *Flowers* criticizes the shift in emphasis from defendants' interests to juror interests.

[161] Juror 7's testimony is representative:

Electronic copy available at: https://ssrn.com/abstract=4457050

term "partial" referred to Juror 6's references to her personal experience. The court rejected the argument, holding that all jurors bring their experiences to deliberations and that, just as it would be inappropriate to tell a juror who tended to trust the police to disregard that trust, "it is similarly inappropriate to tell a juror who has skepticism about the police that she must set aside this `partiality' in order to deliberate."[162]

The court's statement was fair, and a good distillation of the reasons why diverse juror perspectives are valuable. The statement can be read as a defense of bias, though the court was careful to use the phrase "skepticism." And indeed, if the idea of a bias-free trial is a fiction not worth taking seriously then having an array of biases that might check each other is arguably the next best thing. The difference between skepticism and bias in this context is not particularly clear.[163]

Regardless of what one thinks on that point, the court's framing of the question elides a harder issue, at least if the court meant to endorse verdicts based on the record rather than on offsetting biases. To say that perspectives need not be checked at the jury room door is not to say that perspectives themselves count as evidence. It may be hard to discern, but there is a conceptual difference between examining the evidence from different points of view, which the court rightly defended, and ignoring evidence because one's prior experience is strong enough to displace it.[164] For example, Juror 6 was reported to have said that the police may have planted the gun in the Impala.[165] The defense did not make that argument; other jurors said no evidence supported it. Is that comment an example of a juror making something up to justify their position, in which case the comment is lived experience *instead of* evidentiary analysis, or does the comment count *as* evidentiary analysis? Conventional trial procedure would say the former, though nothing would follow from that conclusion. Juror 6 did not refuse to deliberate, which is

---

> A. Again, I mean, it came up that she was not comfortable with making that kind of decision because she had known somebody that had been wrongly convicted in kind of similar scenario that she was – she couldn't -- she could not make that decision. . . .
>
> A. And then I think I made the general -- I made the general comment to the group because I was stewing on it at home that we needed to make sure that we're -- you know, we are looking at the facts that we're given here and that we're not bringing something else that, you know, is influencing us to this, that we just have to look at what we're doing here. And ***she said again that she could not make this -- that decision because of she was afraid of making the wrong decision because of this past experience.*** And I said, "I understand, but that's their leave to look at what we've got here and not bring that to this proceeding."

7/24/20 Tr. at 30-31 (emphasis added).

[162] Order at 7.

[163] In the psychology literature the term bias would be associated with a deviation from a statistically correct probability assessment, or from the most logical inference.

[164] Ideally *voire dire* is the time to identify (or self-identify) a juror whose distrust in the criminal justice system was so great that deliberation would not be possible, but one should not expect too much from *voire dire*. It is hard to answer questions about whether one could perform a task that may be unfamiliar in a case that is unfamiliar. And who wants to say publicly that they could not be fair or could not follow the law?

[165] 8/5/20 Tr. at 11 (Juror 14).

Electronic copy available at: https://ssrn.com/abstract=4457050

possible cause for dismissal as a juror,[166] and nothing in the transcript from her or other jurors suggests she was consciously engaged in nullification.

In every case every juror will have life experiences that inform their assessments of evidence. In the conventional conception of a trial, in which jurors are instructed to focus on evidence, a juror's decision ideally will be a product of that experience applied to the evidence in a case. Setting the IAT and implicit bias theory aside, one need only believe that culture and experience affect perception to accept that no vote will ever be based on an un-intermediated view of the evidence. There is no such thing. Life experience may play greater or lesser roles for different jurors, but each will contribute. It is thus fair to conceive of juror's votes, and contributions to deliberations, in terms of a ratio—some portion derives from experience, and some portion derives from evidence. The ratio of experience to evidence will likely appear different to jurors making comments and jurors listening to them, which makes discussion hard.

D. The attribution problem.

One way to read *Berhe* is that it allows courts to police conduct among jurors regardless whether that conduct alters a verdict. On that reading, *Berhe* aims to make courts and jury rooms less hostile places for persons in protected classes, and it does not matter whether implicit bias caused the sort of interactions recounted on remand in *Berhe*.

The opinion also may be read, and more naturally, as looking to determine whether racial bias had a causal effect. To the extent the case is about causation, the remand hearing in *Berhe* effectively presented two hypotheses: (i) other jurors treated Juror 6's comments dismissively because they were implicitly biased to do so based on her race; and (ii) other jurors treated Juror

---

[166] At some hard-to-specify point a juror's reliance on their own views to the exclusion of evidence may be deemed a refusal to deliberate, and the juror may be discharged and replaced by another juror. The standard is relatively strict, however, to avoid lessening the burden of proof by removing unpersuaded jurors. People v. Cleveland, 25 Cal. 4th 466, 485 (2001), offers a nonexclusive list of examples:

> A refusal to deliberate consists of a juror's unwillingness to engage in the deliberative process; that is, he or she will not participate in discussions with fellow jurors by listening to their views and by expressing his or her own views. Examples of refusal to deliberate include, but are not limited to, expressing a fixed conclusion at the beginning of deliberations and refusing to consider other points of view, refusing to speak to other jurors, and attempting to separate oneself physically from the remainder of the jury. The circumstance that a juror does not deliberate well or relies upon faulty logic or analysis does not constitute a refusal to deliberate and is not a ground for discharge. Similarly, the circumstance that a juror disagrees with the majority of the jury as to what the evidence shows, or how the law should be applied to the facts, or the manner in which deliberations should be conducted does not constitute a refusal to deliberate and is not a ground for discharge. A juror who has participated in deliberations for a reasonable period of time may not be discharged for refusing to deliberate, simply because the juror expresses the belief that further discussion will not alter his or her views.

It is not clear whether the original trial judge in Berhe would have granted such a motion if other jurors had testified as they did during the remand hearing. Juror 6's change of vote mooted the question. It is tempting to dismiss the result in Berhe because Juror 6 could have forced a new trial anyway by adhering to her not-guilty vote, but her change in vote did foreclose this inquiry. As the jury appeared ready to declare themselves deadlocked, however, the prosecution did not lose much of a chance.

Electronic copy available at: https://ssrn.com/abstract=4457050

6's comments dismissively because those jurors attempted to follow instructions to assess the evidence in the case and rationally rejected Juror 6's comments because her comments did not tie off to the evidence. Each explanation could be true in part, and which explanation seems more likely might vary based on the specific interaction in question.[167]

IAT data on how quickly persons associate terms with categories of persons do not provide much help on how to assess these competing explanations. The problem is not one of association, which the IAT is designed to test, but of attribution: should the conduct Juror 6 complained of be attributed to other jurors' disposition, in the form of implicit bias, which was Juror 6's conclusion, or to the evidence and the jury instructions, which is what other jurors testified to? In general, attribution analysis implies that the more external factors constrain behavior the less plausible a dispositional explanation for that behavior becomes, and vice versa.

A significant finding within social psychology suggests people are systematically too inclined to presume that the actions of other people are best explained by the dispositions of those people when circumstances might be sufficient to explain observed behavior.[168] In other words, we explain our own behavior by pointing to the contexts that constrain us, while attributing others' behavior to their character. Focusing on what appears to have been the key issue in *Berhe*—the use of the word "partial" by one juror—the problem is one of competing attributions. Juror 6 attributed other jurors' antagonism to their dispositions in the form of implicit bias. At least one other juror attributed Juror 6's comments to her disposition—

[167] The hypotheses could be flipped: (i) other jurors thought Juror 6 was partial because they over-attributed her conduct to disposition; (ii) other jurors thought Juror 6 was partial because her comments did not tie off to the record.

[168] Sometimes referred to as correspondence bias or fundamental attribution error. *E.g.* Lee Ross, *The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process*, in L. Berkowitz (Ed.), Advances in experimental social psychology 173; Daniel T. Gilbert & Patrick S. Malone, The Correspondence Bias, 117 PSYCH. BULL. 21 (1995). The error or bias may be characterized as the divergence between the (notional) fraction of conduct explained by circumstances and the fraction an observer would attribute to circumstance rather than disposition. If an action was 90% constrained by external factors and an observer thought it 30% likely that disposition explained the action, the difference in the probability estimates would be counted as bias. An interesting paper by Professor Walker and colleagues posits that correspondence bias may not be a bias and instead reflect a rational assessment of conditional probabilities. Walker et al, *Reconsidering the "Bias" in "The Correspondence Bias,"* 9 DECISION 263 (2022). Because jury instructions tell jurors to base decisions on evidence introduced at trial, jury deliberations are a plausible example of a partly deterministic environment (partly because jurors are told to follow instructions but may decline to do so without material risk) in which there are competing sufficient explanations for observed behavior—i.e., jurors are implicitly biased or jurors are sincerely trying to apply rules provided in instructions to the evidence admitted at trial. This approach has the promise of assessing the relative weight of competing factors; it encounters the (likely insurmountable) difficulty of needing some foundation to specify the probabilities of competing explanations in particular cases. In *Berhe*, for example, variables would included at least: (i) juror baseline levels of implicit bias; (ii) juror attentiveness to instructions; (iii) strength of the evidence; (iv) juror attentiveness to the evidence; (v) strength (relative to instructions or some other metric) of juror arguments; and (vi) the degree to which one juror's responses to another juror's argument tracked (v) in view of (i)-(iv). Notwithstanding such debates, which are to be expected for any prominent thesis in psychology, correspondence bias is widely enough documented and accepted to make it a fair question to pose in cases such as *Berhe*.

Electronic copy available at: https://ssrn.com/abstract=4457050

partiality--because she and Berhe were both Black.[169] Each attribution discounts other explanations. Juror 6 did not accept that other jurors might rationally reject her contributions as unfounded in the record; other jurors did not accept that Juror 6 believed she was commenting on evidence rather than, as they perceived, ignoring it.

The risk of attribution error suggests that both Juror 6 and other jurors would be prone to overestimate the degree to which behavior in the jury room was caused by internal, dispositional forces. That means that Juror 6's attribution of bias should have been scrutinized relative to the external record, as should the interpretation of other jurors, who thought that Juror 6 was ignoring the evidence and repeating her conviction without giving reasons grounded in the record. Though only one juror appears to have used the phrase "partial," the other jurors testified that Juror 6 did not have reasons for her positions and simply did not want to vote to convict.

To assess fairly the claim that other jurors' behavior was driven by dispositional factors in the form of implicit bias, the trial court should have considered whether the reactions of other jurors could be explained by contextual factors constraining their behavior including: (i) the evidence; (ii) the jury instructions; and (iii) how Juror 6's comments related to each of (i) and (ii). In simpler terms, Juror 6 testified that other jurors called her statements illogical—were they? Juror 6 testified that at least one other juror accused her of being partial: Did the record support that attribution?

The trial court on remand asked the latter question but not the former. It scrutinized the actions and reactions of other jurors, but it did not assess substantively the comments to which those jurors reacted. That was a mistake. Probably it was a mistake primed by the Washington Supreme Court's opinion—there was not much doubt that a ruling granting the motion for new trial was less likely to be reversed than a ruling denying it—but the trial court's opinion on remand exemplifies the risk that IAT-style implicit bias concepts will truncate rather than enhance analysis. That is not because attribution analysis is antagonistic to research on associations. Attribution bias is, after all, a form of bias. But the principal contribution of implicit bias rhetoric could be to induce (or allow, if motivated) a judge to overlook the attribution question and conduct a hearing as a search for bias, the presumption of which was the basis for the hearing. Such an inquiry is effectively circular and unlikely to produce a sound basis for attributing conduct to disposition or circumstance.

E. The rule for peremptory challenges is not right for all purposes.

Berhe extended Washington's *ex ante* peremptory challenge rule to an *ex post* inquiry into deliberations. Though in each case the problem is one of attributing behavior to a cause, that extension is problematic because institutional considerations are different. That is particularly true because the implicit bias rhetoric animating Washington's jury selection rule was in effect a cover story—it was based on a need to counteract what was presumed to be conscious

[169] Juror 6 was the only juror who testified affirmatively to the use of the word "partial," but one would expect the term to be at least potentially more salient, and thus more memorable, to her.

Electronic copy available at: https://ssrn.com/abstract=4457050

institutional bias among judges, not on the demonstrated utility of implicit bias concepts in making sound attributions of cause regarding peremptory challenges.

In addition, peremptory challenges and verdicts should not be assumed to pose equivalent risks of racist taint. Peremptory challenges have faced strong criticism as a practice. To take one example, Justice Marshall's concurring opinion in *Batson* argued that "[t]he inherent potential of peremptory challenges to distort the jury process by permitting the exclusion of jurors on racial grounds should ideally lead the Court to ban them entirely from the criminal justice system."[170] Strict standards for peremptory challenges make sense if one believes they are barely tolerable in the first place. It is at least not obvious that verdicts should be subject to the same suspicion. Concern that racism will not taint verdicts is present in each case, but there are important institutional differences as well.

1. Implicit bias and the peremptory challenge problem.

Washington's General Rule 37 was adopted to "eliminate the unfair exclusion of potential jurors based on race or ethnicity."[171] Rule 37 modified the rule of *Batson v. Kentucky*,[172] under which peremptory challenges would be disallowed only if the party opposing the challenge could show the challenge was racially motivated. Procedurally, under *Batson*, a party challenging the exercise of a peremptory (mostly defense attorneys) had to show a *prima facie* case of discrimination. The prosecutor then would have to offer a race-neutral explanation for the challenge. If they did, the defendant would have to show the explanation was pretextual, or the challenge stood.[173] That is the origin of the structure in *Berhe*—a requirement for a *prima facie* showing followed by a hearing.

The nominal problem addressed by Rule 37 was that intent is hard to prove, particularly if the subject of inquiry is a skilled advocate aware of what needs to be said to preserve a peremptory challenge. For that reason, many well-informed people thought *Batson* did little if anything to eliminate race-based peremptory challenges. At the same time, there was consensus that many if not most peremptory challenges were motivated—consciously—by race. By 2005 Justice Breyer observed that "the use of race– and gender-based stereotypes in the jury-selection process seems better organized and more systematized than ever before."[174] He catalogued an industry of service providers offering to help lawyers use stereotypes in a sophisticated manner to pick juries. He rightly pointed out that demand for high-quality stereotyping "reflect[s] a professional effort to fulfill the lawyer's obligation to help his or her client."[175] Evidence was

[170] Batson v. Kentucky, 476 U.S. 79, 107 (1986)(Marshall, J. concurring).
[171] WASH. GEN. R. 37(a).
[172] 476 U.S. 79 (1986).
[173] Miller-El v. Dretke, 545 U.S. 231, 267 (2005)(Breyer, J. concurring).
[174] *Miller*, 545 U.S. at 270.
[175] *Id.*

Electronic copy available at: https://ssrn.com/abstract=4457050

uncovered suggesting that some prosecutors were trained on how to provide pretextual justifications for race-based challenges.[176]

In other words, even after *Batson* everyone knew the real problem was that race matters, that lawyers know that race matters, and that they act accordingly—tactically, consciously--to advance their clients' interests.[177] It is true that in *Batson* itself—in 1986, long before the foundational implicit bias paper or the IAT—Justice Marshall suggested in his concurring opinion that unconscious racism might prompt a peremptory challenge.[178] That was a fallback position, however. The lead risk Justice Marshall identified was "outright prevarication by prosecutors."[179] Over 20 years later Professor Paul Butler summed the point up nicely: "the prosecutor who says he doesn't consider race when choosing jurors is either stupid or a liar."[180] They're not stupid, and neither are defense counsel.

Batson sensibly acknowledged that race matters while holding that lawyers are supposed to act as if it didn't. It should have been—and wasn't—a surprise that lawyers, on both sides of the caption, continued to base their tactical decisions on what everyone involved seemed to agree (and IAT optimists would posit) was reality. Washington expressed dissatisfaction with the *Batson* standard in *State v. Saintcalle*,[181] an opinion that called for a new *Batson* standard.[182] *Saintcalle* relied extensively on Justice Breyer's 2005 concurrence in *Miller-El v. Dretke*,[183] which in turn had cited a 2005 law review article by Professor Antony Page, called Batson's *Blind–Spot: Unconscious Stereotyping and The Peremptory Challenge.*[184] Professor Page argued that lawyers exercising peremptory challenges for racial reasons might not know they were doing do because "attorneys, like everyone else, may lack self-awareness."[185] The citations for this point focus on psychological observations tending to show that people often do not know their own minds.[186] Professor Page argued "[a]n attorney may well in good faith think she has identified her reasons, without knowing that her reasons were distorted by her unconscious expectations. The decision-maker is not dishonest; she just lacks adequate self-awareness."[187] Those poor people.

---

[176] Alafair S. Burke, *Prosecutors and Peremptories*, 97 IOWA L. REV. 1467, 1477 (2012); Paul Butler, *Mississippi Goddamn*: Flowers v. Mississippi's *Cheap Racial Justice,* 2019 S. CT. REV. 73, 98 (discussing training to be a prosecutor and the use of group affiliation in jury selection).

[177] *E.g.* Butler, *Mississippi Goddamn, supra* note [n] at 84 ("it's a commonplace among lawyers, judges, and scholars that Batson is underenforced and easy to evade.")

[178] Batson v. Kentucky, 476 U.S. 79, 106 (1986)("Marshall, J. concurring).

[179] *Id*.

[180] Paul Butler, LET'S GET FREE: A HIP-HOP THEORY OF JUSTICE 12 (2009).

[181] 178 Wash. 2d 34, 43 (2013), abrogated by City of Seattle v. Erickson, 188 Wash. 2d 721, 398 P.3d 1124 (2017).

[182] *Id.* at 55. The Court did not write a new standard because the parties had not proposed one or briefed the issue.

[183] 545 U.S. 231, 267 (2005)(Breyer, J. concurring).

[184] 85 B.U. L. REV. 155 (2005). A classic earlier argument for the relevance of unconscious influences on behavior was Charles R. Lawrence III, *The Id, the Ego, and Equal Protection: Reckoning With Unconscious Racism*, 39 STAN. L. REV. 317, 327–28 (1987).

[185] *Id.* at

[186] The relevant sources are notes 368-404 of the article, citing sources including Spinoza, Freud, Kahneman & Tversky (the biases and heuristics literature), Ross & Nisbett (important theorists in attribution analysis), and many others.

[187] *Id*. at 235.

Electronic copy available at: https://ssrn.com/abstract=4457050

Apropos of Justice Marshall's first-order concern with *Batson*, "outright prevarication by prosecutors,"[188] and Professor Butler's equally trenchant observation, the problem Page identified was unlikely to reflect the sort of implicit bias studied by the IAT. A peremptory challenge is issued by: (i) a person acting in a representative capacity, and thus under obligations to a client;[189] with (ii) personal reputational capital at stake and (iii) expertise (through experience or advice) in how juror biases affect verdicts; with the challenge reflecting (iv) an estimate of the probability that a juror would favor the client's case; (v) relative to the next-most-likely alternative juror; provided that (vi) sufficient non-racial grounds for the challenge could be invoked. If we assume that acting in a role affects behavior,[190] the IAT captures none of this context. That is not its fault. It is not meant to do so, and probably could not scale to large populations if it tried to do so.[191]

One could not exclude the possibility that unconscious stereotypes—in the common-sense understanding of the terms—had marginal effects, but they were not the main problem. The problem was more that lawyers (again mostly, but not only, prosecutors) knew exactly what they were doing but that *Batson* made it too hard to prove they were doing it. The majority opinion in *Flowers v. Mississippi*[192] posited that "*Batson* immediately revolutionized the jury selection process that takes place every day in federal and state criminal courtrooms throughout the United States." [193] The smarter money would bet on Professor Butler's assessment: "This assertion is ludicrous."[194]

If everyone knew that racial challenges were common, however—if Justice Breyer could point to a service industry overtly offering to help lawyers make them—why didn't judges just disallow race-based challenges? The real problem Professor Page identified was institutional bias within the judiciary. He noted that judges might be reluctant to "stigmatize a lawyer" who offered race-neutral reasons for a peremptory as "a liar, and maybe racist or sexist as well,"[195] and his insight on this point was sound. Judges are lawyers, judges and lawyers may know each

---

[188] *Id*.

[189] As Professor Butler rightly says, "There are cases in which it would come close to legal malpractice for either the prosecutor or the defense to ignore race." Butler, *Mississippi Goddamn, supra* note [n] at 76. That fact should not be assumed to be lost on lawyers. He considered *Flowers* such a case:

> It would amount to something close to malpractice for a defense attorney in Mr. Flowers's case to ignore race. Her case would depend on getting as many African Americans as possible seated on her jury, and therefore to strike nonblack jurors. And, likewise, it would be virtually impossible for a prosecutor to be indifferent between a white and black potential juror.

Id. at 100.

[190] A reference to Zimbardo seems obligatory.

[191] *Cf* Kang et al, *supra* note [n] at 1146 ("Unfortunately, because of the incredible difficulties in research design, we do not have studies that evaluate implicit bias in real criminal trials. Accordingly, the existing body of research, while strongly suggestive, provides inferential rather than direct support that implicit bias accounts for some of the race effects on conviction and sentencing").

[192] 588 U.S. – (2019).

[193] *Id.* at [n].

[194] Butler, *Mississippi Goddamn, supra* note [n] at 84.

[195] *Id*. at 177.

Electronic copy available at: https://ssrn.com/abstract=4457050

other and socialize together or at least have common acquaintances, lawyers are repeat players in the judicial system, and judges understand that a finding that a lawyer lied to a judge would materially harm the lawyer's standing and career. A judge might want to be very sure they had the goods on a lawyer before causing such harm. Page understood that the unconscious bias idea could lower the threshold for rejecting a peremptory challenge because courts would not have to respond to a prosecutor's non-racial justification for a challenge with "the difficult finding that the lawyers before them are dishonest."[196]

Somewhat ironically, in *Saintcalle* the Washington Supreme Court agreed that the problem was judges: "A requirement of conscious discrimination is especially disconcerting because it seemingly requires judges to accuse attorneys of deceit and racism in order to sustain a Batson challenge. . . . Imagine how difficult it must be for a judge to look a member of the bar in the eye and level an accusation of deceit or racism."[197] (Apparently so difficult that the judge would prefer to accept an argument they felt was likely to be untrue.) Those poor people?

In *Saintcalle* as elsewhere, the move to unconscious bias was a solution to the problem of judicial reluctance to call out lawyers for doing what everyone knew they were doing. The move to unconscious bias made life easier for judges as a way of making life harder for lawyers, thereby protecting prospective jurors. Because juror interests in participating in the judicial process trump party interests that might be served by a race-based peremptory challenge, that was a good result. But the resort to the rhetoric of implicit bias—which on its face *exculpated* lawyers for what was likely intentional race-based conduct--was an unfortunate concession to the reality of judicial reluctance to reject challenges.

Professor Sam Bagenstos refers to moves such as this as "depersonalization," which is part of what he referred to the "political project" of implicit bias research.[198] "Depersonalization says that people should not be upset when others call out implicit bias. To have one's acts attributed to implicit bias is not to be accused of having done something wrong. It is to be accused of being human."[199] In the context of peremptory challenges, depersonalization was understandable. It was a lot easier to change the *Batson* rules by arguing that lawyers suffer unconscious bias than by accusing judges of being clubby or timid, which is where the institutional bias arguments pointed. Instead of condemning lawyers who exercise peremptory challenges on racial grounds, depersonalization almost commiserated with them for their unfortunate but all-too-human foible.

To be clear, however, the move was more than a bit disingenuous. Justice Marshall's first-order criticism, and Professor Butler's trenchant comment, were right. Peremptory challenges are conscious acts made by experts as a tactic in a complex game. Washington's General Rule 37 was a sound tactical response to the institutional judicial bias identified by

[196] *Id*. at 260.
[197] 178 Wash. at 53.
[198] Samuel Bagenstos*, Implicit Bias's Failure*, 39 Berkeley J. Emp. & Lab. L. 37 (2018).
[199] *Id*. at 41.

Electronic copy available at: https://ssrn.com/abstract=4457050

Professor Page and the Court in *Saintcalle*, but that is the bias it really responded to. The rule was not based on a judgment, much less evidence, that implicit measures could explain even peremptory challenges, much less other aspects of trial, better than explicit measures or standard forensic techniques such as cross-examination.

2. Implicit bias and scrutiny of deliberations.

This exculpatory depersonalization strategy does not scale well from judicial evaluation of lawyers making peremptory challenges to judicial evaluation of witness testimony. It is true that both assessing a peremptory challenge before trial and the *ex post* analysis in *Berhe* pose attribution problems. But some obvious structural differences are easy to identify. *Voire dire* may give jurors a small sense of the case but cannot realistically test their reactions to the evidence they will see if seated. Jurors answering questions in *voire dire* face the tough task of predicting how they may perform an unfamiliar job. Post-verdict, there will be a much richer record to assess how jurors did. Most significantly, peremptory challenges occur before the parties, the court, and jurors invest time in a case, and before parties get whatever semblance of closure a verdict may provide. It is to the trial court's credit that it paused in its order granting a new trial to acknowledge "the anguish that reversing a conviction will cause the family of the victim."[200]

Against this background, the Washington Supreme Court's comment in *Berhe* that "a person may honestly believe and credibly testify that his or her actions were not influenced by racial bias, even where implicit racial bias did in fact play a significant role,"[201] is too strong. The authority for that assertion was *Saintcalle* and, as we have seen, *Saintcalle* referred to reasons prosecutors gave to justify peremptory challenges, to the reluctance of judges to call out pretext, and to the strong logical and empirical reasons to believe that lying about peremptory challenges was common. In that context the premise that speakers might not know their own mind, while likely window-dressing in most cases, was useful.

The judicial biases identified in *Saintcalle* are unlikely to apply to juror witnesses, or even to witnesses in general. On average, witnesses are not likely to be repeat players, are not part of the judicial peer group, and while they may take a career or social hit if they are found to have engaged in a racist act an adverse credibility finding is less likely to hit a witness in the same way a finding that a lawyer lied to a judge would hit a lawyer. Nothing in the *Batson* background suggests that judges sitting as factfinders have trouble with witness credibility determinations, and there is no real reason to think they would. Implicit bias premises might give judges an easy way to discount witness testimony in the same way they give judges an easy way to discount lawyer justifications, but the two situations are materially different.

[200] Order at 8. This is probably a safe assumption, subject to the qualification that Mr. Williams was Black and one would not want to ignore the possibility that his family might understand the court's concerns.
[201] 193 Wash. 2d at 664.

Electronic copy available at: https://ssrn.com/abstract=4457050

As with peremptory challenges, however, the most important feature of witness testimony on race-related issues is the risk that witnesses will lie. They have an incentive to lie because it is no longer socially acceptable to admit to racially motivated decisions (other than in contexts in which doing so would be considered compensatory, as with faculty or judicial appointments). A witness who admitted to racial motivation would face a risk of social sanction. But lying and dissembling are conscious processes, not implicit. As noted in Part II, data at present show a positive correlation between implicit and explicit measures of bias, meaning that courts should not simply assume that candid answers will never reveal bias.[202] The data do not support a categorical assumption that cross-examination cannot detect implicit bias.

It is not clear that the Supreme Court's strong statement influenced the trial court's treatment of juror testimony in the remand hearing. It is notable, however, that the trial court spent no time examining the reasons jurors other than Juror 6 gave for their treatment of her views. Under Rule 37, a trial court would examine whether a lawyer asked of each prospective juror the questions asked of a challenged juror, which is one way of framing what the trial court did on remand in *Berhe*—it asked whether other jurors' comments were dismissed as illogical, stupid, or irrelevant. But in a Rule 37 challenge the trial court would also consider the answers jurors gave on *voir dire*,[203] which the trial court on remand in *Berhe* did not do.

To the extent the Supreme Court's opinion was read as an instruction to discount testimony positing non-racial causes, that instruction is not justified by implicit bias data. It may be justified by standard reasons to scrutinize exculpatory testimony, but that is standard judicial prudence, to which citations to implicit bias literature add little of use.

F. Dynamic effects

Part of Justice Breyer's critique of *Batson* was that lawyers will take race into account when doing so is good for their clients, and they will act for race-based reasons to the extent permitted to do so, again because lawyers are competitive and want to get good results for their clients, i.e., to win. In the *Batson* context that meant offering race-neutral reasons for actions, but the general point was that lawyers will make the kinds of arguments judges accept. If judges signal receptiveness to arguments, lawyers will make them

Which brings us to the brilliant defense lawyering in *Berhe*. His counsel told the opposite tale of a *Batson* story. Instead of nothing being about race, anything that plausibly could be portrayed as evincing bias was portrayed that way. (To be clear; this is professional praise, not a criticism.) Juror 6's affidavit was written using implicit bias terminology familiar in Washington's case law (and eventually in Rule 37, which was adopted in 2018, after Berhe's

[202] And, as noted above, the experimental conditions meant to detect bias may not scale well to the jury room, weakening inferences that might be drawn from differences between explicit and implicit measures on intergroup interactions.

[203] *E.g.* State v. Omar, 12 Wash. App. 2d 747, 754 (2020)(analyzing juror responses in a Rule 37 challenge). Review under Rule 37 is *de novo.*

Electronic copy available at: https://ssrn.com/abstract=4457050

2016 trial).[204] It worked brilliantly. The trial court order on remand struggled with the thinness of the record and omitted inquiries into topics that should have been considered in a causal inquiry, such as whether other jurors' characterizations were apt, but that the trial court would order a new trial without such inquiry is a testament to the power of the defense position. Once the Supreme Court was persuaded to shift the burden of proof, and once implicit bias premises were detached from the peremptory challenge context, the decision was effectively made. The Washington Supreme Court's extension of *Berhe* in *Henderson*, in some ways a harder case, confirms the significance of defense counsel's achievement in *Berhe.*

Washington's General Rule 37 applies to issues of race and ethnicity, but the implicit bias premise reaches farther, as do provisions such as California Code of Civil Procedure section 231.7.[205] Because implicit bias claims do not require overt conduct, cases alleging inter-juror implicit bias will be harder to distinguish from cases to which the more lenient duress standard applies. For that reason, lawyers advancing new trial arguments based on conduct among jurors, rather than expressions of bias directed to a defendant or witnesses, have an incentive to frame their motions using implicit bias concepts. The more implicit bias claims succeed, the more they will be made.

Drawing on its discussion of *Jackson*, the rule of *Berhe* applies to conduct among jurors and to juror conduct directed at parties or witnesses. One question it raises is what would happen under the current standard if jurors had complained to the trial court during deliberations that Juror 6 was engaged in nullification. The remand record would not support such a charge under Washington's standard,[206] but a note to the court would alert the court that the juror in question was the only Black juror. *Berhe* might be read to imply that concerns for the secrecy of deliberations should give way at that point to the need to investigate what was going on in the jury room. Suppose the court ascertains the facts recounted above. Should the court then discharge the juror who accused Juror 6 of being partial? *Must* a trial court do that unless the state could disprove an inference of implicit bias?[207] That would be a fair extension of the logic of *Berhe*, though the Washington Supreme Court had no reason to think about that problem and I doubt a court would have discharged any juror based on the remand record in *Berhe*. But one could conceive of defense counsel shifting the usual defense position, of protecting holdouts by limiting inquiry, to go after the jurors giving the holdout juror a hard time. Nothing in *Berhe* would limit that strategy.

[204] Juror 6 worked for a lawyer in some capacity, 8/5/20 Tr. at 69 (Juror 3), so the phrase "I feel these attacks carried an implicit racial bias" also may have been hers, or may have reflected how she thought about things. Either way, it fit very well into the *Saintcalle* rhetoric.
[205] Which forbids peremptory challenges based on "race, ethnicity, gender, gender identity, sexual orientation, national origin, or religious affiliation, or the perceived membership of the prospective juror in any of those groups."
[206] *See supra* note [n].
[207] If implicit bias is encompassed within Wash. Rev. Code Ann. § 2.36.110 (West), which states it is the "duty of a judge to excuse from further jury service any juror, who in the opinion of the judge, has manifested unfitness as a juror by reason of bias, prejudice" and other factors, the answer would seem to be "yes." Again, the burden of proof would likely be decisive. One likely could not prove other jurors acted from bias, but the State could not disprove it either.

Electronic copy available at: https://ssrn.com/abstract=4457050

The logic of *Berhe* extends beyond conduct among jurors. Suppose a Black woman witness gives testimony that a juror, during deliberations, characterizes as "combative." Suppose the juror argues that the witness should be disbelieved because their testimony was suspiciously argumentative and hostile. Extending the principle of *Berhe,* the Washington Supreme Court in *Henderson* held that such comments, made in closing argument by a White female attorney about a Black female plaintiff, established a *prima facie* case of racial taint. Would the comments of the juror in the hypothetical above establish such a case as well? One important difference is that the juror's comments would not be transcribed; an investigator would have to uncover them after a verdict. *Berhe* creates an obvious incentive for such investigation, which was encouraged by Justice McCloud's concurrence.[208] In criminal cases the limiting factor on investigations will probably be resource constraints; in civil cases it may be only juror cooperativeness.

If and to the extent Washington or other courts continue to extend *Berhe*, lawyers will have an incentive to inject race, gender, or other bias-related categories into trials, if only to be able to accuse their opponents of misconduct.[209] Lest one think this the dystopic paranoia of an out-of-touch academic, it is happening already. One prominent attorney was sanctioned in a civil case for accusing opposing counsel of antisemitic prejudice, [210] which one presumes would fall within the rule of *Berhe*, and antisemitism has been raised as a charge in other civil cases as well. A recent trademark case in California concerning dolls has yielded multiple sanctions motions, a mistrial, and accusations of racism based on trial counsel's quotations of a lyrics written by a witness.[211]

Implicit bias research is a largely scholarly endeavor, with some arguable implications for broader policy. Litigation is tactical. Lawyers will present arguments to which judges are receptive, and they will lay the foundation for such arguments where they can. In this context that means a rational lawyer will maximize the chance of making an implicit bias argument to reverse an unfavorable verdict. In practical terms that means selecting witnesses for whom such arguments are plausible, to the extent counsel has a choice, and injecting such issues strategically where possible, either through argument or through examination. The *Berhe* standard presents the prospect of a sort of reverse-*Batson* problem, with a favorable standard for a new trial motion serving as an incentive to frame as racial bias anything an objective observer believing in the causal power of implicit bias *could* interpret as bias. If a *prima facie* case can be shown, and the

[208] 443 P.3d at 1184. It is notable, of course, that Juror 6 came forward and was not first contacted by defense counsel. That voluntary action enhanced her credibility with the courts.

[209] A strand of research suggests that doing so might backfire for counsel to the extent they depended on a biased reaction—jurors otherwise inclined to be biased might self-monitor more closely, potentially offsetting what otherwise might be the effect of casual if not implicit bias. Samuel R. Sommers & Phoebe C. Ellsworth, *Race salience" in juror decision-making: misconceptions, clarifications, and unanswered questions*, 27 BEHAV. SCI. & L. 599 (2009). For counsel worried about bias, however, this finding implies the approach stated in the text, which has the additional benefit of creating a record that might support a new trial motion.

[210]Freshub, Inc. v. Amazon.com, Inc, Memorandum Opinion and Order, December 17, 2021 Case 6:21-cv-00511-ADA (W.D. Tex.). The sanction order might suggest that raising such issues is risky, but that will depend on the jurisdiction. The argument was raised in Waco.

[211] Craig Clough, *MGA Atty's N-Word Use Looms Over Doll Retrial Featuring T.I.*, Law 360 Feb. 2, 2023.

Electronic copy available at: https://ssrn.com/abstract=4457050

standard in *Berhe* is modest, the burden will shift to the prevailing party to make a showing that, if implicit bias premises are taken literally, will not be made. At this point, implicit bias is a meme with a life of its own, and it seems almost certain that lawyers will wield it to the extent possible.

CONCLUSION

Somewhat surprisingly, close focus on implicit bias research and on *Berhe* shows that, for all the ink spilled on the implicit bias concept, it does not explain the result. The shift in the burden of proof provided the formal explanation, and the trial judge's opinion was admirably direct in acknowledging that the most parsimonious explanation comes from the legal scholar Bob Dylan: You don't need a weatherman to know which way the wind blows.

In the courtroom implicit bias premises have the vices of their virtues. The premise that everyone's behavior is influenced by biases inaccessible to introspection, undetectable by inquiry, and not reliably remediable, is so strong it cannot generate useful means of attributing behavior to implicit rather than explicit biases, nor useful means of attributing behavior to disposition (implicit or explicit) rather than a trial record or other external forces.[212] It cannot draw the distinctions the system needs to operate in specific cases. The most tangible use of implicit bias theory in court has been Washington and California's use of implicit bias rhetoric in revising the rules for jury selection, but at least in Washington those revisions were not based on tangible insights from implicit bias theory that allowed judges to make sound attributions about peremptory challenges. Implicit bias rhetoric was used to provide an explanation for reform that did not accuse judges of accepting insubstantial justifications for such challenges.

There is nothing inherently wrong with using implicit bias rhetoric to give judges cover to pursue an agenda, just as there was nothing inherently wrong with Earl Warren's citation of the Clarks' doll studies in *Brown.*[213] Whether that move is desirable depends on the agenda.[214] It would be useful, however, not to confuse the cover with the inquiry. Professor Butler makes the point well:

---

[212] A recent paper by IAT optimists provides little basis for optimism on this score:

> [T]heorizing about implicit cognition is relatively unsophisticated at this time. Why might this be? First and foremost, implicit cognition by its nature refers to aspects of human thought that are relatively less accessible to conscious awareness. As such, scientists themselves have a harder time generating good intuitions about the mechanisms of implicit cognition.

Kurdi et al, *supra* note [N] at 32. The point seems obvious, but the problem is even more acute in court, where attribution decisions, not theories, are required.

[213] No one stays up nights worrying that the studies were underpowered, or whether the *r* scores were too weak to justify desegregation, either. No one cares, and they shouldn't.

[214] It would be nicer to the jurors in *Berhe* not to disregard their testimony of the reasons for their actions, and the statements to which they responded, by effectively accusing them of false consciousness, as the trial court did. But the reality of litigation, at least from the defense side, is that if those jurors have to take a hit for Berhe to get a new trial, so be it.

Electronic copy available at: https://ssrn.com/abstract=4457050

> It cannot actually surprise anyone—especially any lawyer—that race matters with regard to jury deliberation. But when people pretend that race is irrelevant, it makes honest conversations more difficult and encourages bad law and policy.[215]

Berhe can be read as a marker cautioning judges to ensure that courts are not hostile places for persons in protected classes. That is a laudable goal. Implicit bias theory does not tell courts how to pursue it and suggests instead that the task will be hard and perhaps Sisyphean. Jurors are not repeat players, they are unlikely to be de-biased with an instruction, and most of them are not trained to argue well in groups. Those are not reasons to throw up one's hands, to borrow from the Court in *Berhe*, but they counsel caution. Implicit bias premises do not point to a non-arbitrary way to do such work.

If, on the other hand, the problem in *Berhe* is to decide whether race (or presumably gender or other protected classification) exerted some degree of causal influence on a verdict, then it is a problem of attribution not of a notional base rate of bias. Should a challenge or a juror comment be attributed to disposition or facts in the external world? Professors Gilbert and Malone offered a lovely simile to describe this problem: A balloon rises in the air; is it full of helium or have winds swept it upwards?[216] Implicit bias theory suggests all balloons have some helium. It does not suggest all balloons have enough to gain lift in all circumstances. Either way, the theory is not a substitute for a wind gauge. To the extent it may have thought otherwise, the trial court on remand in *Berhe* erred.

Which takes us back to the question of how courts are using implicit bias ideas. What is the agenda and how should we think about it? At the end of the day, Juror 6 always had the ability to hang the jury and to force a new trial. Washington let her change her mind back to where it had been for most of the deliberations. Implicit bias concepts can't get you there, but it's not a crazy result. It is possible that even Everett Williams's family would understand. If courts keep an eye out for practical, sensible results then the references to implicit bias won't do any harm, even if the concepts entail no conclusions in particular cases.

[215] Butler, *Mississippi Goddamn*, *supra* note [n] at 107.
[216] Daniel T. Gilbert & Patrick S. Malone, The Correspondence Bias, 117 PSYCH. BULL. 21 (1995).

Electronic copy available at: https://ssrn.com/abstract=4457050