

U.C.L.A. Law Review

Race, Intellectual Disability, and Death: An Empirical Inquiry Into Invidious Influences on *Atkins* Determinations

Sheri Lynn Johnson, John H. Blume, Amelia Courtney Hritz,
& Caisa Elizabeth Royer

ABSTRACT

In *Atkins v. Virginia*, the U.S. Supreme Court held that the execution of a person with intellectual disability violates the Eighth Amendment's Cruel and Unusual Punishment Clause. After more than a decade of *Atkins* litigation, we perceived there to be a substantial risk that race influences intellectual disability—and consequently, life and death—determinations. Due to the difficulty of demonstrating the influence of race in a particular case, we decided to investigate its potential effects in a controlled experiment. We did so by manipulating race in three different ways and by presenting cases with both strong and ambiguous evidence of intellectual disability. We found statistically significant race effects when we showed the face of the defendant and when the evidence of intellectual disability we provided was ambiguous. The influence of race was more pronounced when we limited our sample to white mock jurors. Even with a relatively weak manipulation, the size of the race effect is substantial. We also discovered that many participants weighed the facts of the criminal case and the consequences of their decision (death penalty eligibility), even though it was not relevant to the determination of whether the claimant was (or was not) a person with intellectual disability. These findings shed light on why claims of intellectual disability almost never succeed before juries: death-qualified jurors may not make the diagnostic determination based on the evidence, but instead likely upon their own assessment of death-worthiness.

AUTHORS

Sheri Lynn Johnson is the James and Mark Flanagan Professor of Law, and Assistant Director of the Death Penalty Project, at Cornell Law School. John H. Blume is the Samuel F. Leibowitz Professor of Trial Techniques and Director of the Death Penalty Project at Cornell Law School. Amelia Courtney Hritz is the Robert B. Kent Public Interest Fellow at Cornell Law School. Caisa Elizabeth Royer is a Post-Doctoral Fellow at the University of Utah. The authors wish to thank Emily Eagleton, Maddie Feldman, Meghan Flyke, Ali Franz, Gina Garrett, Anna Kallmeyer, Jackie Katzman, Cathryn Masloff, and Michelle Morris for their excellent research assistance. We are also grateful to Lynn Marie Johnson at the Cornell Statistical Consulting Unit for her advice.



TABLE OF CONTENTS

INTRODUCTION.....	1508
I. THE EMPIRICAL STUDIES.....	1511
A. Experiment 1.....	1512
1. Methodology.....	1512
2. Results.....	1513
a. Context.....	1513
b. Race.....	1515
c. Intellectual Disability Definition.....	1516
B. Experiment 2.....	1518
1. Methodology.....	1518
2. Results.....	1519
a. Race and Strength of Evidence.....	1519
b. Explanation for Decision.....	1519
C. Experiment 3.....	1521
1. Methodology.....	1521
2. Results.....	1522
a. White Participants.....	1523
b. Explanation for the Participants' Decisions.....	1524
CONCLUSION.....	1526
APPENDIX.....	1529

INTRODUCTION

In *Atkins v. Virginia*,¹ the U.S. Supreme Court held that the execution of persons with intellectual disabilities violates the Eighth Amendment's Cruel and Unusual Punishment Clause. Although the Court has determined that other categories of offenders are ineligible for execution, most notably juveniles, the intellectual disability ban is the only one that is based on a medical or psychological diagnosis. The *Atkins* Court cited the American Association of Mental Retardation (AAMR) (now the American Association of Intellectual and Developmental Disabilities (AAIDD)) and the American Psychiatric Association (APA) definitions as the touchstones for determining intellectual disability. The two nearly identical definitions refer to substantial limitations in present functioning characterized by three prongs: (1) deficits in intellectual functioning; (2) deficits in adaptive functioning; and (3) onset of these deficits before the age of eighteen.² *Atkins* does not, however, mandate any particular procedures for determining whether a person has an intellectual disability, nor does it address the possibility that bias may affect the determination.

After almost two decades of *Atkins* litigation, the Supreme Court has intervened only to insist that states employ clinically approved formulations of intellectual disability.³ It has refused to consider whether jurors or judges should decide the intellectual disability question or whether a defendant may be required to prove his intellectual disability beyond a reasonable doubt.⁴ It has also declined to review cases in which courts have explicitly relied upon the

-
1. *Atkins v. Virginia*, 536 U.S. 304 (2002). For a more detailed discussion of the Court's decision in *Atkins*, see MARC J. TASSÉ & JOHN H. BLUME, INTELLECTUAL DISABILITY AND THE DEATH PENALTY: CURRENT ISSUES AND CONTROVERSIES 143 (2018).
 2. AM. ASS'N INTELLECTUAL & DEVELOPMENTAL DISABILITIES, INTELLECTUAL DISABILITY: DEFINITION, CLASSIFICATION, AND SYSTEMS OF SUPPORTS (11th ed. 2010) [hereinafter AAIDD]; see also AM. PSYCHIATRIC ASS'N, DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS 33 (5th ed. 2013) [hereinafter DSM-5]. The third prong, onset during the developmental period, distinguishes intellectual disability from traumatic brain injury.
 3. See *Hall v. Florida*, 572 U.S. 701, 724 (2014) (holding that Florida's brightline IQ cutoff approach to prong one, subaverage intellectual functioning, violates the Eighth Amendment because it is inconsistent with clinical definitions of intellectual disability); *Moore v. Texas*, 137 S. Ct. 1039, 1044 (2017) (holding that Texas's use of the Briseno factors to determine prong two, adaptive functioning deficits, violates the Eighth Amendment because it is inconsistent with clinical definitions of intellectual disability).
 4. *Hill v. Humphrey*, 662 F.3d 1335 (11th Cir. 2011), cert. denied, 134 S. Ct. 115 (2013).

defendant's race, nationality, or language in deciding that he does not have an intellectual disability.⁵

Lower court cases explicitly relying on African American and Latino⁶ defendants' group membership first prompted us to assess the scope of the problem. Robert Sanger documents cases where prosecutors have argued for adding points to the IQ scores of African American and Latino (or as the courts more often say, Hispanic) individuals, purportedly to account for perceived biases in IQ testing.⁷ But such ethnic adjustments have no scientific basis, and they permit the execution of African American and Latino defendants when identically scoring white defendants would be exempt from execution.

We suspected that race may silently influence other facially-neutral *Atkins* determinations. There are at least two reasons to think it might do so. First, racial stereotypes of intellectual ability are pervasive. Interestingly, these stereotypes do not lead in an obvious way to predict the direction that race might bias intellectual disability determinations. Might a decisionmaker believe that African Americans or Latinos are less intelligent than members of other racial groups, and therefore, at least in close cases, be more inclined to find a defendant from one of those groups to be a person with intellectual disability? Alternatively, might a decisionmaker who accepts that stereotype dismiss evidence of cognitive impairments as normal *for that group* rather than finding it to be probative of intellectual disability? Or perhaps some decisionmakers will be affected by stereotypes in the first way and others in the second, which might produce no observable difference based on race, but nonetheless embody race-based arbitrariness.

Second, we know from the broader literature on motivated cognition that a decisionmaker may first evaluate an issue based on preferences regarding the outcome and then look to the evidence to confirm his or her judgment.⁸ More particularly, our own research has revealed that mock jurors are more likely to find intellectual disability on a particular set of facts when they are told their

-
5. See, e.g., Brief for United Mexican States as Amicus Curiae Supporting Petitioner at 14, *Hernandez v. Stephens*, 134 S. Ct. 1760 (2014); *Commonwealth v. DeJesus*, 58 A.3d 62, 72–73 (Pa. 2012).
 6. We use *Latino* rather than *Latino/a* or *Latinx* both because courts use either *Latino* or *Hispanic*, and because all of the claimants of which we are aware are male. Similarly, we present participants with hypothetical defendants with male names and faces because *Atkins* claimants are virtually all male.
 7. Robert M. Sanger, *IQ, Intelligence Tests, "Ethnic Adjustments," and Atkins*, 65 AM. U. L. REV. 87 (2015).
 8. Ziva Kunda, *The Case for Motivated Reasoning*, 108 PSYCHOL. BULL. 480, 490 (1990); see also, e.g., Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 812, 814–6 (2001) (demonstrating this phenomenon with judges).

finding will result in the awarding of benefits than when told it will make a defendant ineligible for the death penalty.⁹ And we also know that white jurors—mock and real—are more likely to favor the death penalty for African American defendants who have killed white victims than for other racial combinations.¹⁰ Although much less data regarding such predispositions for Latino defendants is available, the frequency and nature of bias against Latinos would predict a similar pattern, though possibly with smaller race effects. Thus, it might be that a greater preference for death sentences for some defendant-victim racial combinations affects the likelihood that a defendant will be found to have an intellectual disability because a juror knows that the intellectual disability finding will determine whether death is imposed.

Although we have collected and reported data on all of the litigated *Atkins* cases,¹¹ that data cannot answer whether or how race (or nationality or language) is affecting *Atkins* ineligibility. The first obstacle to drawing reliable conclusions from win-loss rates in *Atkins* claims is the high likelihood of selection effects. How can we know whether attorneys are more or less likely to raise *Atkins* claims based on race? Put differently, how can we know that the intellectual disability claims raised on behalf of Latino clients on average are equally as strong as the claims raised on behalf of white clients? Stereotypes may influence an attorney's decision to raise an intellectual disability claim (in either direction) just as they may influence a judge or juror. Moreover, controlling for the strength of the case is very difficult, given the nature of both IQ scores and adaptive functioning deficits. Almost all defendants will have more than one IQ score, usually on more than one kind of IQ test. An average IQ of 69 may be produced by scores of 68 and 70 or by scores of 79 and 59; the consistency of the first pair suggests reliability whereas the second pair poses a disparity so large it suggests some form of invalidity. Moreover, the weight that should be accorded a particular score

9. Data on file with the authors (as of yet unpublished).

10. See, e.g., David C. Baldus, George Woodworth, David Zuckerman, Neil Alan Weiner & Barbara Broffitt, *Racial Discrimination and the Death Penalty in the Post-Furman Era: An Empirical and Legal Overview with Recent Findings from Philadelphia*, 83 CORNELL L. REV. 1638, 1659–60, 1676 (1998); John H. Blume & Sheri Lynn Johnson, *Unholy Parallels between McCleskey v. Kemp and Plessy v. Ferguson: Why McCleskey (Still) Matters*, 10 OHIO ST. J. CRIM. L. 37, 39 (2012); Tara L. Mitchell, Ryann M. Haw, Jeffrey E. Pfeifer & Christian A. Meissner, *Racial Bias in Mock Juror Decision-Making: A Meta-Analytic Review of Defendant Treatment*, 29 L. & HUM. BEHAV. 621, 622, 627–29, 631 (2005).

11. See John H. Blume, Sheri Lynn Johnson, Paul Marcus & Emily Paavola, *A Tale of Two (and Possibly Three) Atkins: Intellectual Disability and Capital Punishment Twelve Years After the Supreme Court's Creation of a Categorical Bar*, 23 WM. & MARY BILL RTS. J. 393 (2014); John H. Blume, Sheri Lynn Johnson, & Christopher Seeds, *An Empirical Look at Atkins v. Virginia and its Application in Capital Cases*, 76 TENN. L. REV. 625 (2008).

depends both upon the reliability of the instrument and the testing conditions. Evidence of adaptive functioning deficits is even more variable; this is in part because persons with intellectual disability have both strengths and weaknesses which are difficult to compare,¹² but also because the weaknesses may be reported by more or less trustworthy informants. Thus, comparing win-loss rates in decided cases is not a valid way of assessing whether *Atkins* decisions are racially biased.

Given both the substantial risk that race influences intellectual disability—and consequently, life and death—determinations, and the difficulty of demonstrating the influence of race in cases without explicit references to race, we decided to investigate its influence in controlled settings. We did so by manipulating race in three different ways and by presenting cases with both strong and ambiguous evidence of intellectual disability. We found statistically significant race effects in the ambiguous evidence condition when we manipulated race by showing faces of the defendants. Participants were less likely to find an African American or Latino defendant to be a person with intellectual disability as compared to a white or Asian defendant. That we found those effects only in cases with ambiguous evidence is consistent with previous research on motivated cognition;¹³ it is also likely to replicate the kind of evidence presented to juries in real cases because cases with very strong evidence of intellectual disability are likely to be settled by pleas.¹⁴ That we found significant race effects only when we made race more salient—by using faces rather than names—was not surprising, but given that race would be much more salient in a real trial than in our strongest experimental manipulation, we conclude that the effects we observe would almost certainly be stronger in real *Atkins* determinations.

I. THE EMPIRICAL STUDIES

In order to explore the influence of race on intellectual disability determinations, we conducted a series of empirical studies. In each study, we presented participants with a vignette and asked them to determine if the subject

12. Although standardized instruments for assessing adaptive functioning are available, they are designed for administration assessing current functioning, and not the retrospective use that occurs in *Atkins* determinations, so courts rarely rely heavily on scores from those instruments, and thus, such scores cannot be fruitfully compared.

13. See Avani Mehta Sood, *Motivated Cognition in Legal Judgments—An Analytic Review*, 9 ANN. REV. L. SOC. SCI. 307, 309–10 (2013) (discussing that motivated cognition can shape the interpretation of ambiguous stimuli).

14. Blume et al., *supra* note 11, at 399.

was a person with intellectual disability. The vignette included a recitation of evidence by two sides: one in favor of the diagnosis and one disputing it. In each experiment we manipulated the race of the claimant. We also asked participants to explain their reasoning, answer questions about their views on the death penalty and the *Atkins* decision, and provide demographic information.¹⁵

A. Experiment 1

1. Methodology

In Experiment 1, we varied the race and ethnicity of the claimant by changing the name in the vignette. Participants were randomly assigned to read about an African American, Latino, or white claimant. We used one of three first names for each condition: for the African American condition we used DaShawn Washington, Jamal Washington, or Tyrone Washington; for the Latino condition Diego Hernandez, Mateo Hernandez, and Tomas Hernandez; and for the white condition Billy Nielson, Greg Nielson, or Seth Nielson.¹⁶ We selected these names based on U.S. Census data and previous research.¹⁷

In addition to varying the name of the claimant, we varied the context of the claim. Participants were either given a death penalty case or a disability benefits case. In both situations, participants were instructed that the criteria for intellectual disability are significant limitations in intellectual functioning, significant limitations in adaptive behavior, and onset prior to the age of eighteen. They were also given the same evidence relating to intellectual disability: the defendant or claimant failed grades in school, had IQ scores within the range of intellectual disabilities (between 62 and 72), was unable to hold a job for more than a few weeks, never had a driver's license or bank account, never lived alone, and was unable to cook or manage money. The scenario also provided other facts about the defendant, including that he played football and had been previously married, that are within the range of behaviors that a person

15. Participants stated whether they generally favor or oppose the death penalty for convicted murderers (strongly favor, somewhat favor, unsure, somewhat oppose, strongly oppose), and whether they agreed that “intellectual disability isn’t an excuse for a crime, at least not if the defendant is capable of telling right from wrong,” “anyone with intellectual disability shouldn’t be sentenced to death,” “persons with intellectual disability should not be excluded from the death penalty,” and “persons with intellectual disability should not be given benefits by the government.” Participants responded: agree strongly, agree moderately, agree slightly, not sure, disagree slightly, disagree moderately, and disagree strongly.

16. The vignette referred to the claimant by his full name initially and then only by his first name.

17. See, e.g., S. Michael Gaddis, *How Black Are Lakisha and Jamal? Racial Perceptions From Names Used in Correspondence Audit Studies*, 4 SOC. SCI. 469, 475–76 (2017).

with intellectual disability may demonstrate, but suggest some strengths. Participants were provided with expert testimony on both sides of the question.

In the death penalty context, the defendant was charged with armed robbery and double homicide. Participants read a description of the crime, which included clear evidence of guilt (video footage) and upsetting details like the murder of a child and a motive of pecuniary gain. In the disability benefits context, the subject was fired from his job on his first day. Participants read the benefits available to the subject if diagnosed with an intellectual disability, which included supervision, a job, and a supported living placement.

We recruited 407 participants from Amazon's MechanicalTurk. Fifteen participants were dropped from analysis for providing answers to the open-ended questions that indicated they were not paying attention to the vignette. The remaining 392 participants were 57 percent male, with an average age of 35 years ($SD = 10.36$), and 36 percent Republican. The racial composition of this sample was 81 percent white, 6 percent black, 6 percent Latino, 6 percent Asian, and 1 percent other.

2. Results

a. Context

Our most striking finding was the difference in the rates of finding a person to have an intellectual disability in the murder case compared to the benefits case, even though the evidence of intellectual disability was the same.¹⁸ In the murder case, only 50 percent of participants found that the person had an intellectual disability, compared to 81 percent in the benefits case.¹⁹ This suggests that either

18. It would be inappropriate to consider the criminal behavior depicted in the hypothetical as evidence that the defendant did not have an intellectual disability. See John M. Fabian, *State Supreme Court Responses to Atkins v. Virginia: Adaptive Functioning Assessment in Light of Purposeful Planning, Premeditation, and the Behavioral Context of the Homicide*, 6 J. FORENSIC PSYCHOL. PRAC. 1, 15 (2006) ("In essence, there is no objectivity or 'rule of thumb' indicating that an individual who functions within the second percentile of cognitive and adaptive abilities pursuant to mild mental retardation, also lacked a substantial capacity to purposely plan and premeditate the homicide."); see also Bethany Young, Marcus T. Boccaccini, Mary Alice Conroy & Kristy Lawson, *Four Practical and Conceptual Assessment Issues that Evaluators Should Address in Capital Case Mental Retardation Evaluations*, 38 PROF. PSYCHOL.: RES. & PRAC. 169, 174 (2007) ("[N]obody has attempted to examine the extent to which mental health professionals agree about the level of functioning needed to engage in certain types of criminal behavior.").

19. See *infra* Appendix. We estimated a logistic regression model to predict which participants would find that the claimant is a perintellectual disability based on the type of case and the race of the claimant. We used dummy codes for the categorical predictors and included the interaction term. The estimates of the raw scores of the predictor variables, standard errors, 95

the presence of criminal behavior or the consequences of an intellectual disability determination (or both) influence some jurors' determination of intellectual disability, despite the fact that neither is a clinically relevant factor.

Attitudes toward the death penalty were an important factor in intellectual disability determinations. Notably, regardless of the context, people who favored the death penalty were significantly less likely to find that an individual had intellectual disability.²⁰ Jury selection in capital cases includes "death qualification," a process that excludes prospective jurors whose opposition to the death penalty would impair their ability to impose a sentence of death. Previous research has found that death-qualified juries are more likely to convict and less likely to accept an insanity defense than would be juries that included jurors who are opposed to the death penalty.²¹ Our data suggests that a death-qualified jury would also be less likely to find that individuals have intellectual disability.

The influence of context is also apparent in participants' open-ended descriptions of their reasoning. Participants in the death penalty condition were far more likely to mention the consequences of their decision in their responses than participants in the disability benefits condition (24.50 percent versus 2.51 percent). One participant wrote, "Thomas should be eligible because murder is murder. He thought about his actions and proceeded with them." Such rationales may explain why participants' death penalty beliefs strongly predicted their intellectual disability determinations. This bottom-up approach, in which the participant let their choice of punishment color their decision about intellectual disability, was also evident in participants who found the defendant did have an intellectual disability. For example, another participant wrote, "I am highly against the death penalty. Leave the boy to rot in jail."

Several participants in the death penalty condition also discussed the quality of knowing the difference between right and wrong as a determinant of intellectual disability. One participant wrote:

Having a low IQ doesn't mean one doesn't understand right from wrong. If he does have the abilities of an eleven-year-old then he must clearly know right from wrong. He must know that murder is wrong. If he hadn't shot the nine-year-old[] potential witness, I

percent confidence intervals, and odds ratios are displayed in the Appendix. Positive estimates indicate that finding intellectual disability is more likely. The odds of participants finding the claimant to be a person with an intellectual disability in the benefits context are 2.78 times higher than the odds in the death penalty context.

20. $X^2(4, n = 392) = 23.62, p < 0.001$.

21. E.g., Phoebe C. Ellsworth, Raymond M. Bukaty, Claudia L. Cowan, & William C. Thompson, *The Death-Qualified Jury and the Defense of Insanity* 8 L. & HUM. BEHAV. 81, 84–85 (1984).

might have thought a bit differently but that murder was clearly in cold blood.

Another participant used similar reasoning: “He was intelligent enough to plan a robbery and obtain a gun. He should face a normal sentence.” This comment exemplifies another pattern present in many participants’ responses: belief that the commission of a murder requires a level of intellectual functioning that precludes a diagnosis of intellectual disability. One participant wrote, “If he’s coordinated enough to rob a store, he’s not retarded. He’s just stupid and violent.” Such reasoning demonstrates how difficult it is for participants to separate the facts of the crime and their normative views about punishment from their decision about intellectual disability. Of course, at *some* extreme, the commission of a very sophisticated crime could—if done without help—rule out intellectual disability. But the crime in the vignette, which is intended to be typical of crimes committed by persons on death row, is not a sophisticated one. Therefore, the participants who rely on the crime to find the defendant did not have an intellectual disability are either relying on gross stereotypes,²² or being influenced by their personal views of what the defendant deserves.

b. Race

Overall, the intellectual disability determination in this experiment (strong evidence, weak manipulation of race) did not *significantly* differ by first or last name,²³ though it is interesting to note that some first names produced much lower rates of intellectual disability findings than did others.²⁴ Several participants, moreover, demonstrated that the defendant’s race influenced *their*

22. Both the AAIDD and the APA require focus on adaptive *deficits* because “[t]he skills possessed by individuals with mental retardation vary considerably, and the fact that an individual possesses one or more that might be thought by some laypersons as inconsistent with the diagnosis (such as holding a menial job, or using public transportation) cannot be taken as disqualifying.” James W. Ellis, *Mental Retardation and the Death Penalty: A Guide to State Legislative Issues*, 27 MENTAL & PHYSICAL DISABILITY L. REP. 11, 21 n.29 (2003). The AAIDD further instructs that significant deficits in adaptive skills are “*not* outweighed by the potential strengths in some adaptive skills.” AAIDD, *supra* note 2, at 47 (emphasis added).

23. See *infra* Appendix.

24. Rates of finding intellectual disability by first name: Billy (70 percent), Greg (70 percent), Seth (50 percent), Diego (57 percent), Mateo (65 percent), Tomas (62 percent), DaShawn (68 percent), Jamal (62 percent), and Tyrone (73 percent). By last name: Hernandez (61 percent), Nielson (64 percent), Washington (68 percent). All p 's > 0.10, indicating a lack of statistical significance. The variation among first names was unexpected, and we did not design the study to have high enough power to test for differences based on first name. Power indicates the probability that a test of significance will identify an effect present in the data. For this study, there is 49 percent power at .05 significance level. The general rule is that power should be above 80 percent in order to test an effect.

decision. One participant in the disability benefit condition defended the participant's decision that the defendant did not have intellectual disability by writing, "No rewards for dumb blacks." Another participant provided a more detailed rationale:

Low intelligence should not be subsidized by the government unless medically diagnosed as a disorder/disease that is genetically/biologically based—possibly and conditionally in such cases. Mateo obviously is a less than capable person, probably Mexican, and society is not obligated to support such people, as unfortunate as they may be.

Even when participants did not mention race explicitly, of the participants who concluded that the defendant did not have an intellectual disability, those given the stereotypical African American or Latino names were more likely to say that the defendant was stupid and yet not find that he had an intellectual disability. For example, one participant in the Latino name condition wrote, "Unable to manage money or do math is not a disability. He's able to get married, find a job, and take basic care of himself. He's just low intelligence, not disabled." Another participant wrote, "I think that Tomas is just slower than the normal person but not disabled in any way. There are some people that are smarter and some that are not as smart. It's more so that Tomas is unmotivated and lazy."

Altogether, the participants' responses demonstrate the complexity of decisions about intellectual disability and the variability of responses to evidence of such disability. Some participants supported their decisions with a variety of legally relevant facts—including IQ and adaptive functioning—while other participants were influenced by the consequences of the intellectual disability decision and their own biases.

c. Intellectual Disability Definition

Some—though not all—participants in both the strong and weak evidence conditions mentioned factors relevant to the three-prong test for intellectual disability in their explanations of their decisions. One-third of participants (33.92 percent) mentioned the defendant's IQ scores in their rationale, while one-half of participants (51.65 percent) mentioned the defendant's adaptive functioning. Participants used the defendant's multiple IQ scores to support both "yes" and "no" votes of intellectual disability. For example, one participant who believed the defendant was a person with intellectual disability wrote this, "In all of the IQ tests he always tested around the borderline of 70 or below, and this has been the case since before he was 18." This response was typical of most

participants who believed the defendant to be a person with intellectual disability, with many participants using the IQ scores to demonstrate that the defendant had a consistent history of below average functioning prior to the age of eighteen. In comparison, participants who did not believe the defendant was a person with intellectual disability tended to use the IQ scores as evidence of the defendant's malingering. For example, one participant said:

Based on his third IQ test, where he received a score of 72 which was significantly higher than the final IQ test where he received a 64 score, it's reasonable that Tyrone purposefully failed the IQ test and gotten [sic] a lower score in order to qualify for government benefits. In addition, Tyrone had shown to be able to function in his daily life in such a way that indicated that he had no visible sign of intellectual disables [sic] before the prosper of receiving the benefits came into play.

Another participant wrote, "His average IQ comes to 70.25. It would have been higher if he wasn't trying to fail." It is important to note that the participants were not given much context or explanation about intelligence testing, which may have increased the likelihood that the participants used the numbers as post-hoc support for whatever choice they preferred to make about intellectual disability. In litigated cases, although more explanation of testing would often be provided, expert testimony interpreting the scores generally is conflicting and therefore would permit the same kind of reliance on preferences.

Participants also demonstrated a variety of approaches for the evaluation of adaptive functioning, which is defined as the skills necessary for effectively navigating through everyday environments. Participants in both conditions used the fact that the defendant was able to get married, hold a job, and play football as evidence that he was not a person with intellectual disability, even though none of these behaviors are beyond the ability of a person with intellectual disability. One participant wrote, "I believe since he was able to care for himself and get married, shows he is able to make important decisions by his own free will." Another wrote, "Greg was able to play football, get employed, and got married. That's pretty functionable[sic]." On the other side, participants tended to cite the defendant's school and employment history as evidence that the defendant was a person with intellectual disability:

Because of his past history in school, he failed many grades throughout school, his previous IQ tests show that he is intellectually below a functioning normal person. Dashawn [sic] is also unable to properly hold down a job, the strongest evidence is the evaluation by the psychiatrist who said DaShawn can't mentally

retain and calculate even very simple word problems when they are told to him verbally.

Every detail from the evidence presented in the scenario was used by at least one participant to support an argument for a finding of intellectual disability and by another participant to reject that finding. This shows the careful consideration that jurors give these kinds of details, but also how facts (like playing football) can influence different people in different ways. For example, some jurors thought the information about the defendant playing football was extraneous, while others saw it as a conclusive proof that the defendant can follow three-stage commands, which they then viewed as evidence that the defendant had average mental functioning. Such differences in interpretation can contribute to arbitrariness of determinations of intellectual disability.

B. Experiment 2

1. Methodology

After finding the strong effect of context (disability versus death eligibility determinations) in Experiment 1, we explored whether bias would play a role in death penalty cases when the facts of intellectual disability were more ambiguous. In Experiment 2, we presented participants with a vignette with either ambiguous or clear evidence related to the defendant's intellectual disability claim. In the clear evidence condition, the defendant failed several grades in school; was in special education classes in school; had IQ scores of 65, 72, 70, and 64; was unable to hold a job for more than a few months; had no driver's license or bank account; and had never lived alone. In comparison, the evidence in the ambiguous evidence condition described poor school performance; IQ test scores of 65, 83, 72, and 69; and that the defendant had briefly lived alone. As in Experiment 1, both vignettes included brief statements from state and defense experts who disagreed about whether the claimant was a person with intellectual disability. All participants were given a death penalty case with the same crime-related facts Experiment 1.

After finding variation among first names in Experiment 1 (possibly due to confounds from other variables associated with names, such as religion, that may introduce other stereotypes of intelligence and criminality), we used a description of the defendant's race in Experiment 2. We added a brief physical description of the defendant to the beginning of the vignette, which included the defendant's race. We removed the name manipulation from Experiment 1 and instead referred to the defendant as "J.M." Drawing from the cognitive motivation literature, we expected to see a greater influence of bias—including racial

stereotypes—on judgments when the facts regarding intellectual disability were ambiguous, compared to when the facts were clear.²⁵

For Experiment 2, we recruited 251 participants from Amazon's MechanicalTurk and had to drop three for nonsensical responses to open-ended questions. The remaining 248 participants were 58 percent male, on average 37.94 years old ($SD = 12.10$), and 32 percent Republican. The racial composition of this sample was 78 percent white, 7 percent black, 5 percent Latino, 6 percent Asian, and 4 percent other.

2. Results

a. Race and Strength of Evidence

As expected, participants were more likely to find that the claimant was a person with intellectual disability when the evidence was clear, compared to when it was ambiguous,²⁶ although somewhat surprisingly, this difference was not significant.²⁷ The one-word description of the defendant's race did not alter rates of findings of intellectual disability in either the strong or ambiguous conditions.²⁸

b. Explanation for Decision

Participants' open-ended explanations for their intellectual disability diagnosis were very similar to the descriptions provided in the first study, with many participants on both sides. Here is one typical example provided by a participant in the weak evidence condition who decided that the defendant did not have an intellectual disability:

JM was able to pass a driver test, get married, held a job. He knows right from wrong even if he can only follow one or two steps after another. He is borderline on the IQ and probably did worse on purpose. He committed a serious crime of murder and knows well

25. See Sood, *supra* note 13, at 309–10 (describing previous studies of motivated reasoning).

26. A total of 50 percent of participants in the clear condition found that the defendant had intellectual disability, compared to 34 percent in the ambiguous condition.

27. See *infra* Appendix. As in Experiment 1, we estimated a logistic regression to predict which participants would find that the claimant is a person with an intellectual disability based on the type of case and the race of the claimant. Again, we used dummy codes for the categorical predictors and included the interaction term. The estimates of the raw scores of the predictor variables, standard errors, 95 percent confidence intervals, and odds ratios are displayed in the Appendix.

28. See *infra* Appendix. The main effect of race and the interaction of race and evidence type were nonsignificant (all p 's > 0.10).

what he did. He can ask for assistance and make himself understood.
He is not that intellectually [disabled] as he received grades of C and
D and showing he knows how to complete work.

Another participant was more succinct: “If he can’t make decisions, why did he shoot the nine-year-old boy that was a witness? If he doesn’t know the difference between right and wrong, why did he flee the scene?” A third participant stated, “[W]here did he get a gun and learn how to use it? If he can figure that out, he’s likely not intellectually disabled.”

Participants who decided that the defendant was a person with intellectual disability expressed concerns about the consequences of their decision. One participant said, “It seems too close to call, so it is better to err on the side of caution and not use the death penalty.” Another participant in the strong-evidence condition said this:

Based on his experiences in school, it is clear that the defendant has a learning disability at the very least. Also, the IQ cutoff is ‘approximately’ 70, and I consider 72 to be within that range. People with intellectual disabilities often get married and hold unskilled jobs, so that didn’t sway my opinion at all. I also don’t think J.M. intended to kill anyone and that it just happened on the spur of the moment. I feel like he lacked the ability to consider the consequences when he first decided to rob the store.

Finally, one participant said, “I made my decision like that because people with intellectual disabilities are at a higher risk of wrongful conviction and death sentences.” It appears that participants on both sides of the issue had a difficult time separating the sentencing decision from the intellectual disability decision—a consideration that likely would be much more salient in a real case.

There were fewer explicit mentions of race in the responses given by participants in this study compared to the first iteration, likely due to the less prominent nature of the defendant’s race in the stimuli. Nevertheless, as with the first study, we found a greater likelihood that participants judging defendants of color would describe the defendant as stupid, yet conclude he was not intellectually disabled. One participant in the Latino condition said, “[J]ust because J.M. has the life skills of a child doesn’t necessarily mean J.M. is actually disabled. By the fact that J.M. failed and dropped out of multiple grades suggests that J.M. just hasn’t had intellectual opportunities.” Another participant said of an African American defendant, “The man is not disabled in his intellect. J.M. simply had a terrible education in his childhood years.”

Taken together, the participants’ open-ended responses in this and the prior study revealed some influence of bias and external factors on participants’

decisions about intellectual disability. In order to further understand how this decision is made, in the next study, we altered the manipulation of race, and we also asked the participants two additional questions: (1) to predict the defendant's actual IQ; and (2) to rate what outside factors (like neighborhood or genetics) influenced the defendant's intelligence.

C. Experiment 3

1. Methodology

In Experiment 3, we included the same ambiguous and clear evidence conditions. We increased the salience of the race manipulation by including a photograph of the defendant's face, rather than just a name or label. We presented participants with one of twelve possible photographs of the defendant. The photographs depicted twelve men of four different races: white, black, Latino, and Asian. We selected photographs from the Chicago Face Database that had the highest levels of prototypicality within their races.²⁹ After reading the vignette and viewing the photograph of the defendant, participants determined whether he was a person with intellectual disability and explained how they made their decision.

We added additional questions to help understand the mechanisms that may cause people to reach different conclusions based on race. We asked participants to make other judgments about the defendant, including likelihood of committing a similar violent act in prison, blameworthiness for the crime, and the defendant's true IQ. We also asked participants to rank how different influences—such as neighborhood and genes—shaped the defendant's IQ scores and adaptive functioning, how they made their intellectual disability determinations, and what facts were relevant to their decision. We expected to see bias in the different race conditions based on stereotypes about criminality, blameworthiness, or intelligence. Participants were predicted to be less willing to find the claimant to be a person with intellectual disability when shown a photograph of a Latino or African American defendant compared to a white defendant.

We recruited 831 participants from Amazon's MechanicalTurk, but 105 had to be dropped from analysis based on nonsensical answers to the open-ended

29. See Debbie S. Ma, Joshua Correll, & Bernd Wittenbrink, *The Chicago Face Database: A Free Stimulus Set of Faces and Norming Data*, 47 BEHAV. RES. METHODS 1122, 1131–32 (2015). To obtain measures of racial prototypicality for each face in the database, raters were asked to rate how closely a face's physical features resemble the features of people of the same race. *Id.* at 1126–27 (describing the norming process in more detail).

questions.³⁰ We increased our sample size because the previous experiments were underpowered. The final 726 participants were 56.3 percent male, on average 35.56 years old ($SD = 11.10$) and 49 percent Republican. The racial composition of participants in this sample was 75.2 percent white, 15.3 percent black, 4.5 percent Asian, 2.8 percent Latino, and 2.2 percent other.

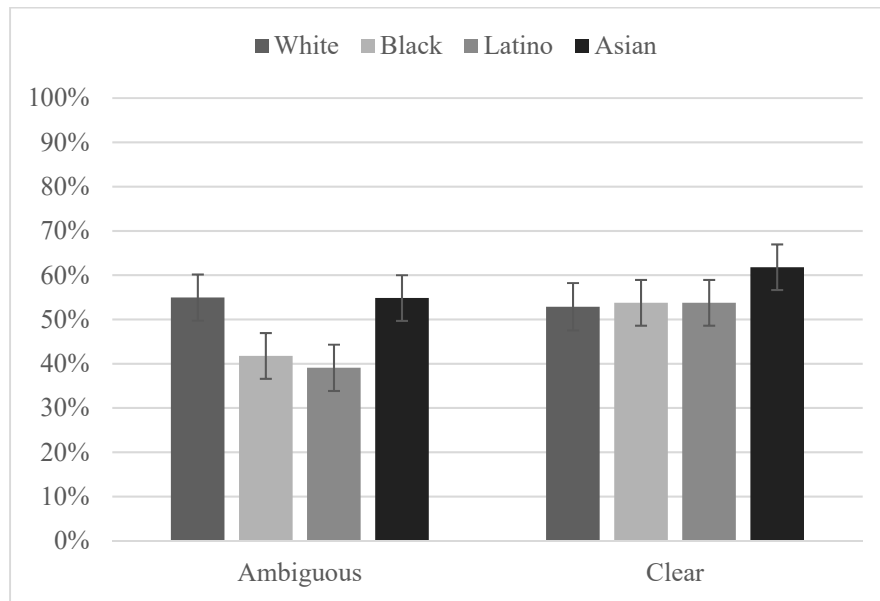
2. Results

With more power and a more salient manipulation of race than we employed in Experiment 2, we found a significant—and sizable—influence of race. When the evidence was ambiguous, participants were significantly less likely to find that a Latino defendant was a person with intellectual disability, as compared to a white or Asian defendant.³¹ Only 39 percent of participants shown a Latino defendant and ambiguous evidence found that the defendant was a person with intellectual disability, as compared to 55 percent of participants shown white defendants and 55 percent shown Asian defendants (see Figure 1). Forty-two percent of African American defendants were found to be a person with intellectual disability.

30. For example, we dropped 19 participants for responding “good” when we asked them to describe how they made their decision. We collected data for Experiment 3 around the same time other researchers were noticing a decrease in quality of responses on MTurk. See, e.g., Emily Dreyfuss, *A Bot Panic Hits Amazon’s Mechanical Turk*, WIRED (Aug. 17, 2018, 11:38 AM), <https://www.wired.com/story/amazon-mechanical-turk-bot-panic> [<https://perma.cc/3TJA-MGUD>].

31. As we did in Experiment 2, we estimated logistic regression to predict when participants would determine that the defendant had an intellectual disability based on the evidence condition (ambiguous or clear), the defendant’s race, and the interaction between race and evidence type. Results of the regression are displayed in the Appendix.

Figure 1: Percent Finding the Claimant to be a Person With Intellectual Disability Based on Evidence Type and Race³²



a. White Participants

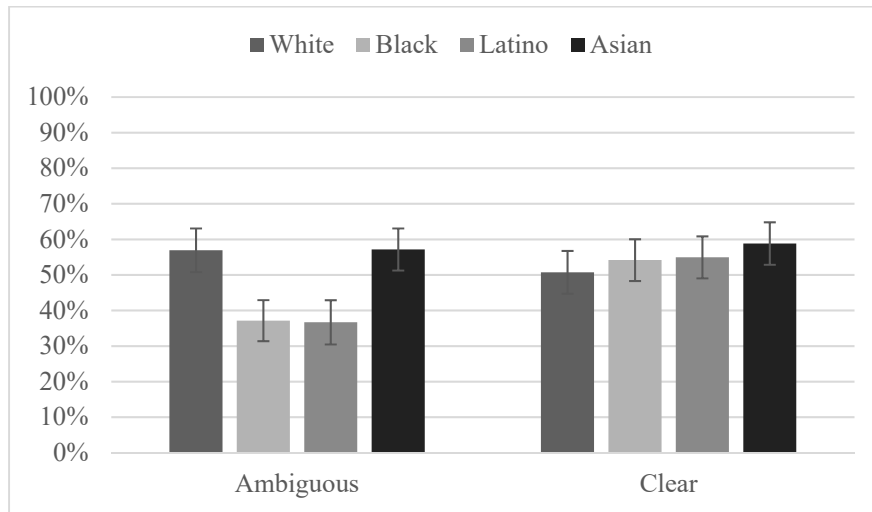
Our sample was 75 percent white and 15 percent African American, but less than 3 percent Latino. To investigate the possibility that the less unfavorable treatment of African American defendants (as compared to Latino defendants) was the result of greater representation of African Americans in our participant pool, we limited our sample to just white participants ($n = 546$). When we did so, two things happened: the race effects became stronger, and African American and Latino defendants were treated equally (badly). Thus, in the ambiguous evidence condition, white participants were significantly less likely to find that either a Latino or an African American defendant was a person with intellectual disability, compared to a white or Asian defendant.³³ White participants found that a defendant was a person with intellectual disability only 37 percent of the time when shown a Latino defendant *or* an African American defendant—

32. Error bars display the standard error.

33. We estimated fourth logistic regression to predict when white participants would determine that the defendant had an intellectual disability based on the evidence condition (ambiguous or clear), the defendant's race, and the interaction between race and evidence type. Results of the regression are displayed in the Appendix.

compared to 57 percent of the time when shown a white defendant *or* an Asian defendant (see Figure 2).

Figure 2: Percent of White Participants Finding the Claimant to be a Person With Intellectual Disability Based On Evidence Type and Race³⁴



b. Explanation for the Participants' Decisions

Linear regressions were conducted to test whether race and strength of evidence conditions predicted how participants rated the influence of different factors—genetics, family, neighborhood, culture, choice, unexplained—on the defendant's IQ score. The influence of neighborhood was significantly different depending on condition, with participants in the clear evidence condition ranking defendant's neighborhood as less influential than participants in the ambiguous evidence condition.³⁵ No other factors showed significant differences between either the race or evidence conditions.³⁶ This suggests that some participants may be responding to perceived socioeconomic status. When

34. Error bars display the standard error.

35. $F(2, 705) = 4.534, p = 0.010, R^2 = 0.103, \beta = -0.291$.

36. We do not understand how the data we gathered regarding estimated IQ scores should be reconciled with the data on intellectual disability determinations. Subjects found higher IQs for Latino and Asian defendants than for white and black defendants, but lower rates of intellectual disability for Latino and African American defendants. However, when outlier estimates of IQ were eliminated, the only remaining difference was between black and white defendants, and in the direction stereotypes would predict. We hope to continue to explore the questions raised by these results.

given less clear information about the defendant's intellectual functioning, some participants looked to outside information, such as the type of neighborhood the defendant grew up in, to make the intellectual disability determination.³⁷

The participants' open-ended responses also reflect this weighing of factors and struggle to understand the defendant's mind. Even as participants weigh the facts of the case, they often were aware that their decision about intellectual disability would have consequences, and sometimes struggled to separate the facts of the case from the decision about the defendant's eligibility. One participant in the strong evidence condition who decided the defendant was not a person with intellectual disability wrote this:

This was very difficult. While his IQ scores (65, 72, 70, and finally 64) do show some form of intellectual "lack," to me the greatest pieces of evidence against him having intellectual disability were given at the end: he was married, employed, and provided basic care for himself.

A participant in the ambiguous evidence condition wrote this:

Based on the facts provided, including the fact that he did obtain a drivers' [sic] license, has been married, and can appear to communicate and understand somewhat complex situations (like the one he is in now). The fact that he understands the death penalty indicates that he understands the concept of death, and suggests that he understands the death he brought to others.

Another participant said, "He may not be the smartest person in the world, but he has been able to do things in life that show he is capable of deciding whether to murder two people." One participant who decided the defendant was a person with intellectual disability also speculated about the defendant's mindset and said:

I think that early in life his drive to succeed was killed by something in his environment and could lead to intellectual and emotional disabilities, and I don't think it's a permanent state for him, but I understand there is a spectrum and when certain attributes are found, like the inability to interact socially, it can be one of the main contributing factors to meeting the requirement of the classification.

37. The reader may wonder if environmental factors such as neighborhood are a legitimate consideration in evaluating intellectual disability. In other words, if environment is the cause of low performance, can that rule out intellectual disability? The answer is no, because the diagnosis of intellectual disability does not depend upon causation; causation may be attributable to genetics, but it may be caused by accident or other environmental factors or be indeterminable. MARC J. TASSE & JOHN H. BLUME, *INTELLECTUAL DISABILITY AND THE DEATH PENALTY: CURRENT ISSUES AND CONTROVERSIES* 1 (2018).

All of these answers show the participants' efforts to establish intellectual disability from the facts, but also through speculation about what the defendant thought or felt.

As with the previous two studies, participants were very aware of the consequences of their decision. One participant shown an African American defendant said that the defendant "is a danger to society either way and a drain on the taxpayers" and determined the defendant was not a person with intellectual disability. Another participant in the African American defendant condition explained their decision against intellectual disability by writing, "I show no mercy. Death is something you don't play with." A third participant wrote, "I don't believe in the death penalty under any circumstances. For J.M., I would want him to be in a rehabilitation-type prison for a long time." One mock juror in the white defendant condition was particularly explicit about his consideration of the consequences of his decision:

I guess I took the decision into my own hands. With my power as a juror, in the end I don't have to respect the fact that if you are legally considered mentally retarded, you can be exempt from the death penalty Also, as a juror, who am I to say another life should or shouldn't be taken because they took someone else's life? I think this man is going to spend the rest of his life in jail anyways, if he does get the death penalty it's going to take 40 years to actually happen, and if he doesn't he just stays in jail for a super long time. His life is done. That's what factored in my decision.

CONCLUSION

Our discussion of the participants' open-ended responses across all three studies focused on how participants deviated from legally relevant factors of the decision about intellectual disability. This is not to say that all participants did so. Many were very conscientious in their use of the three-prong test for intellectual disability (significant limitations in intellectual functioning, significant limitations in adaptive behavior, and onset prior to the age of eighteen) and gave apparently unbiased, reasoned responses.³⁸ However, the fact

38. The following response is an example:

I think he fulfills the first criteria because his first three IQ tests coupled with failing first, sixth and seventh grade demonstrate well below average intellectual functioning *and* that the onset of these difficulties occurred before the age of 18. It also seems quite clear to me that the defendant has significant difficulty meeting even the most basic of his own needs on his own. He clearly does not function as a normal adult person—he can't cook for himself, he's never held a

that some participants were able to perform the task without being influenced by extraneous or invidious factors neither diminishes the significance of those participants who were influenced by those factors nor provides any guidance as to how to eradicate those influences.

This study's finding that an individual's attitudes towards the death penalty influence intellectual disability determinations in ambiguous cases is disturbing, but it is consistent with previous research on motivated cognition finding significant influence of bias in ambiguous situations.³⁹ Because most people with intellectual disability fall within the mild range, and because errors in measurement and variations in adaptive skill levels can make diagnosis challenging in some cases, the facts in litigated *Atkins* cases will be ambiguous, and thus, vulnerable to the influence of motivated cognition. Our research sheds light on why claims of intellectual disability are almost never successful before juries: jurors—who are death qualified—may not make the decision based on facts, but instead may rely upon their own (pro-death) belief systems.

Most importantly, we found that Latino and African American defendants are disadvantaged by their race in the ultimate determination of intellectual disability. Even with a relatively weak manipulation of race—here, a single photo—the size of the race effect is substantial. Because we found that increasing the salience of race increases the effect, the effect of race is almost certainly greater at trial where race is far more salient. Finally, we note that race effects are strongest with white participants, which also suggests an enhanced influence of race at trial, given that capital juries are almost always predominantly—and often exclusively—white.⁴⁰

Although our results strongly suggest that race influences the determination of intellectual disability in a way that harms Latino and African American defendants, they do not reveal the precise mechanism by which it does so. Stereotypes about lesser intelligence may lead decisionmakers to dismiss evidence of impaired functioning as normal for African American and Latino defendants, and thus decrease the likelihood of a finding of intellectual disability

driver's license, he's never had a bank account. Clearly he fits the second criteria as well. Thus, he meets all three provided criteria and is not eligible to receive the death penalty in my opinion.

39. See Sood, *supra* note 13, at 309–10 (explaining the theory of motivated cognition).

40. Shamena Anwar, Patrick Bayer & Randi Hjalmarsson, *The Impact of Jury Race in Criminal Trials*, 127 Q.J. ECON. 1017, 1017–18 (2012); see also, e.g., Catherine M. Grosso & Barbara O'Brien, *A Stubborn Legacy: The Overwhelming Importance of Race in Jury Selection in 173 Post-Batson North Carolina Capital Trials*, 97 IOWA L. REV. 1531, 1538–40 (2012) (describing studies of racial bias in juror selection and finding that prosecutors disproportionately strike black venire members).

for such defendants. Alternatively, it may be that white animosity toward those two groups determines the ultimate decision, and that the negative intellectual disability finding is arrived at by reasoning backwards from the desired result. Quite possibly both mechanisms, in some measure, contribute to the biased results.

APPENDIX

Logistic Regressions Predicting a Diagnosis of Intellectual Disability

Experiment 1

Table 1: Prediction Based on Claimant's Name and Context of Benefits or Death Penalty Case (n = 392)

	β	SE	95% CI	Odds	Odds 95% CI
Intercept	0.13	0.25	-0.37, 0.62	1.13	0.69, 1.86
Latino	-0.40	0.35	-1.10, 0.29	0.67	0.33, 1.33
African American	0.02	0.35	-0.66, 0.71	1.02	0.51, 2.04
Benefit Case	1.02*	0.41	0.25, 1.84	2.78	1.28, 6.31
Latino x Benefit	0.73	0.58	-0.41, 1.88	2.07	0.66, 6.55
African American x Benefit	0.52	0.60	-0.64, 1.71	1.68	0.53, 5.52

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Experiment 2

Table 2: Prediction Based on Description of Claimant's Race, and Clear or Ambiguous Evidence of Intellectual Disability (n = 248)

	β	SE	95% CI	Odds	Odds 95% CI
Intercept	-0.24	0.40	-1.05, 0.55	0.79	0.35, 1.73
Latino	-0.55	0.55	-1.65, 0.54	0.58	0.19, 1.71
African American	-0.27	0.54	-1.35, 0.80	0.76	0.26, 2.23
Asian	-0.74	0.56	-1.86, 0.35	0.48	0.16, 1.42
Clear Evidence	0.12	0.54	-0.94, 1.18	1.12	0.39, 3.25

	β	SE	95% CI	Odds	Odds 95% CI
Latino x Clear	0.74	0.75	-0.73, 2.22	2.09	0.48, 9.20
African American x Clear	0.14	0.74	-1.31, 1.60	1.15	0.27, 4.95
Asian x Clear	1.13	0.76	-0.35, 2.64	3.11	0.71, 14.00

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Experiment 3

Table 3: Prediction Based on Picture of Claimant, and Clear or Ambiguous Evidence of Intellectual Disability (n = 726)

	β	SE	95% CI	Odds	Odds 95% CI
Intercept	0.20	0.21	-0.21, 0.62	1.22	0.81, 1.85
Latino	-0.64*	0.30	-1.25, -0.05	0.53	0.29, 0.95
African American	-0.53+	0.30	-1.12, 0.05	0.59	0.33, 1.05
Asian	-0.00	0.30	-0.59, 0.58	1.00	0.56, 1.78
Clear Evidence	-0.08	0.30	-0.67, 0.51	0.92	0.51, 1.66
Latino x Clear	0.68	0.43	-0.16, 1.52	1.97	0.86, 4.56
African American x Clear	0.57	0.42	-0.26, 1.40	1.76	0.77, 4.05
Asian x Clear	0.37	0.43	-0.46, 1.21	1.45	0.63, 3.35

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Table 4: Regression With Only White Participants (n = 546)

	β	SE	95% CI	Odds	Odds 95% CI
Intercept	0.28	0.25	-0.21, 0.78	1.32	0.81, 2.18
Latino	-0.83*	0.37	-1.56, -0.11	0.44	0.21, 0.89
African American	-0.80*	0.35	-1.50, -0.12	0.45	0.22, 0.89
Asian	0.01	0.35	-0.68, 0.69	1.01	0.51, 2.00
Clear Evidence	-0.25	0.35	-0.94, 0.43	0.78	0.39, 1.54
Latino x Clear	0.99*	0.50	0.020, 1.98	2.70	1.02, 7.24
African American x Clear	0.94+	0.49	-0.01, 1.90	2.57	0.99, 6.72
Asian x Clear	0.32	0.49	-0.64, 1.28	1.38	0.53, 3.60

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.