American Psychological Association
American Psychology-Law Society

Law and Human Behavior

ISSN: 0147-7307

2021, Vol. 45, No. 4, 336–355
https://doi.org/10.1037/lhb0000455

The Impact of Minimal Versus Extended Voir Dire and Judicial Rehabilitation on Mock Jurors' Decisions in Civil Cases

Jessica M. Salerno[1], John C. Campbell[2], Hannah J. Phalen[1], Samantha R. Bean[1], Valerie P. Hans[3], Daphna Spivack[4], and Lee Ross[4]

[1] School of Social and Behavioral Sciences, Arizona State University
[2] Sturm College of Law, University of Denver
[3] Cornell Law School, Cornell University
[4] Department of Psychology, Stanford University

Objectives: Three experiments tested the utility of minimal versus extended voir dire questions in predicting mock jurors' verdicts and damage awards, and whether the biasing impact of their preexisting attitudes on case judgments could be reduced by judicial rehabilitation. ***Hypotheses*:** We hypothesized that extended voir dire questions would be more predictive of case judgments than minimal voir dire questions. We hypothesized that judicial rehabilitation would not reduce this impact of preexisting attitudes on case judgments. ***Method***: Across three experiments, each focusing on a different civil case (insurance bad faith, wrongful birth, medical malpractice misdiagnosis), online participants (N = 2,041; 62% female; 77% White, 9% African American, 6% Asian/Pacific Islander, 6% Hispanic or Latino/a; M_{age} = 40) were paid MTurk workers. They were randomly assigned to experience (a) no voir dire, minimal voir dire focusing on previous legal experience and self-identification of biases, or extended voir dire focusing on specific attitudes about civil litigation, parties, and laws, and (b) no judicial rehabilitation, or judicial rehabilitation, before judging the case. Participants read a civil case, made case judgments, and completed bias awareness measures. ***Results:*** Demographic information and minimal voir dire questions did not predict case judgments, but the majority of extended voir dire responses predicted verdicts and damage awards. Judicial rehabilitation did not reduce the biasing impact of their preexisting attitudes on case judgments—but did result in mock jurors reporting that they were less biased, despite judicial rehabilitation not *actually* reducing their bias. ***Conclusions:*** Attorneys need the opportunity during voir dire to ask jurors about specific attitudes that might bias their decisions because relying on jurors' self-identification of their own biases has little utility. Further, although judicial rehabilitation might make jurors think that they are less biased, it may not actually reduce the impact of their preexisting attitudes on their case decisions.

Bradley D. McAuliff served as Action Editor.

Jessica M. Salerno https://orcid.org/0000-0001-7394-6444
Hannah J. Phalen https://orcid.org/0000-0003-3165-3373
Samantha R. Bean https://orcid.org/0000-0001-9838-218X
Valerie P. Hans https://orcid.org/0000-0003-4470-5478
Daphna Spivack https://orcid.org/0000-0001-5524-1504

Samantha R. Bean is now at the Department of Psychology, New York University.

The authors would like to acknowledge funding for this project from the Civil Justice Research Initiative at the University of Denver Sturm College of Law. They would also like to acknowledge and thank Kevin Morrow for his assistance with legal research. These data were presented by Jessica M. Salerno at the 2020 Experimental Methods in Legal Scholarship Conference (EMLS) at Northwestern University. We thank Kenworthey Bilz and Avani Mehta Sood for their helpful commentary on a draft of this article as a part of the EMLS conference.

In memory of Lee Ross, the authors would like to acknowledge his inspiring contributions to this project and the field of psychology.

The data are available at osf.io/685qa

The experiment materials are available at osf.io/685qa

Correspondence concerning this article should be addressed to Jessica M. Salerno, School of Social and Behavioral Sciences, Arizona State University, 4701 West Thunderbird Road (MC 3051), Glendale, AZ 85036, Unites States. Email: jessica.salerno@asu.edu

Public Significance Statement
Contrary to popular belief, juror biases are not likely to be cured by judicial rehabilitation, which might backfire by creating the illusion in jurors that they are unbiased. Although very few mock jurors were able to self-identify things that might bias them when asked general questions in voir dire, they were willing to admit specific attitudes that biased their verdicts and damage awards when asked. Thus, we suggest reducing bias on juries by allowing attorneys to ask specific, detailed voir dire questions crafted by the parties to streamline the jury selection process and remove jurors for cause or via peremptory challenges, rather than relying on "quick fixes," such as general questions that ask jurors to self-identify their own bias or judicial rehabilitation.

Keywords: bias, civil juror decision making, jury instructions, jury selection, voir dire

Supplemental materials: https://doi.org/10.1037/lhb0000455.supp

The Seventh Amendment to the U.S. Constitution guarantees the right to a fair and impartial jury in civil cases. One of the major ways that the U.S. system attempts to achieve this lofty goal is through voir dire processes designed to identify jurors whose preexisting experiences or attitudes might bias their decisions and justify their exclusion from the jury. Although voir dire procedures differ across jurisdictions, they are based largely on two assumptions about human cognition and behavior. First, it is assumed that individuals are both aware of and willing to acknowledge their biases during voir dire. Second, it is assumed that jurors who acknowledge their own biases can be "rehabilitated" through a procedure whereby a judge informs prospective jurors that they must set aside their biases and asks them explicitly whether they can agree to do so. Most studies investigating the impact of voir dire have focused on the criminal system. The current experiments tasked mock jurors with judging one of three civil cases and investigated the effects of a minimal voir dire versus an extended voir dire and judicial rehabilitation on mock jurors' decisions.

Voir Dire Practices

The voir dire process in jury selection, in which potential jurors are questioned about potential biases, is an essential tool in the pursuit of an impartial jury (American Bar Association, 2005). Based on voir dire responses, the trial judge may excuse biased prospective jurors for cause. Attorneys may also exercise peremptory challenges to exclude jurors, which are a limited number of challenges attorneys can offer without a reason. Voir dire questioning varies considerably across jurisdictions in terms of length, who is conducting it (judges vs. attorneys), and the latitude attorneys are allowed in their questioning (Mize et al., 2007). It can range from a 10-min, judge-conducted voir dire to multiple days of attorney-conducted voir dire in which the judge is absent (Mize et al., 2007).

Voir dire is often very limited. For example, in Colorado, individual judges have sole discretion to limit or terminate voir dire questioning (Colorado Rule of Civil Procedure 47(a)(3)). Trial judges in Colorado commonly impose time limits of approximately twenty minutes or less per side at the outset of the trial (e.g., *Sanderson v. Advanced Eurology P.C.,* 2008). This means that if forty potential jurors are examined, the attorneys have—at most—30 seconds per juror to assess potential biases. In many federal courts, the judge conducts a substantial portion of the voir dire (Mize et al., 2007) and any attorney-led questioning is often limited to 15 minutes per side (Moore, 2019).

The consequence of such time limitations in voir dire is that attorneys are left to rely on stereotypes about jurors' demographic characteristics and previous experience with the legal system, and quick catchall questions that rely on jurors to identify predispositions or attitudes that might bias them. They must trust that prospective jurors can identify their own biases, when asked something as general as, "Do you know of any reason you may be prejudiced for or against the plaintiffs or defendants?" (Hans & Jehle, 2003; Hans & Vidmar, 2007). In contrast, when extensive questioning is allowed, attorneys are better able to probe for specific biases relevant to the case and identify jurors whose views are likely to compromise their impartiality.

Some judges advocate minimal voir dire because they believe that extensive pretrial questioning could "waste too much time and unduly invade jurors' privacy," (Mize et al., 2007, p. 28). Those judges and proponents of minimal voir dire assume that potential jurors can spontaneously self-identify their sources of bias, are willing to admit them and, when they do acknowledge biases, can set them aside and be impartial after undergoing judicial rehabilitation.

The Effectiveness of Voir Dire in Identifying Potentially Biased Jurors

Results from studies assessing attorneys' ability to accurately identify jurors who might be biased against their clients are mixed (e.g., Kerr et al., 1991; Morrison et al., 2016; Otis et al., 2014). Part of the difficulty might stem from jurors being unable to identify or admit their own biases when questioned. Awareness of bias has long been a topic of investigation by social and cognitive psychologists. Both conscious *and* unconscious (or "implicit") biases can affect judgments and behavior (e.g., Greenwald et al., 2009; Greenwald & Banaji, 1995; Stanley et al., 2011). People are often unaware of the biases that influence them (Gaertner & Dovidio, 1986; Nisbett & Wilson, 1977), and are unable to identify them when asked. People suffer from "bias blind spots," which make it particularly difficult for them to identify biases in their own judgments that they can and do recognize in other people's judgments (Pronin et al., 2002).

Even when jurors are aware of their potential biases, they might be reluctant to admit them because they want to be "good jurors" who adhere to the norms of impartiality in a courtroom. In interviews,

25%–30% of jurors in actual trials admitted to previous experiences with victimization or with police that they had not disclosed during voir dire (Seltzer et al., 1991). Anxiety about being evaluated can further result in juror dishonesty (Marshall & Smith, 2012).

Research demonstrates that there is utility in voir dire as long as the questions are about specific attitudes, rather than the more open-ended invitations to reveal bias (e.g., "Can you think of any factors that might prejudice your judgments?") of the sort posed in standard, minimal voir dire (e.g., Hans & Jehle, 2003; Kovera & Austin, 2016; Lieberman, 2011). Given the combination of unconscious bias blind spots and the normative pressure to be a "good juror," it is not surprising that jurors rarely answer these very general questions about bias affirmatively. In general, specific questions are more predictive of behavior than general questions (Ajzen & Fishbein, 1980). Studies in criminal contexts have identified many specific, case-relevant questions that effectively predict case judgments, such as asking about attitudes toward the insanity defense (e.g., Crocker & Kovera, 2010; Skeem et al., 2004) or the death penalty (e.g., Luginbuhl & Middendorf, 1988; Thompson et al., 1984).

Most studies in this area dealt with criminal cases where the jury's task is to offer a guilty or not guilty verdict and the criterion for conviction is the absence of reasonable doubt. It is important to investigate this in civil cases, where the criterion for reaching a verdict favoring the plaintiff is subtler (most often, a preponderance of evidence standard). Further, in civil cases, juries also decide on the size of financial damage awards, which might be particularly susceptible to the influence of preexisting biases about civil litigation and the parties involved. Yet, studies about voir dire in the civil realm are very rare. Those that do exist offer promising evidence of the predictive value of case-specific attitudes. One civil jury simulation study (Diamond et al., 1998) tested whether nine specific attitudes predicted civil case judgments in a products liability case regarding exposure to an airborne residue from fire-proofing material. More specifically, they had three measures of attitudes toward lawsuits (i.e., whether plaintiffs received too much/too little, legitimacy of lawsuits, and the ease/difficulty of winning lawsuits), three measures of attitudes toward business (i.e., attitudes toward business, how much regulation interferes with business, and regulation for public safety), two measures of attitudes toward experts (i.e., perceived expertise, trust), and one measure asking how blameworthy smokers were who developed health problems. All but one (perceived expertise of the expert) of those nine questions predicted verdicts—but not damage awards. While some traditional minimal voir dire questions (i.e., demographics, previous experience with legal system) were also predictive, they explained only half as much variance (5.4%) as mock jurors' case-specific attitudes (11.5%).

Although Diamond et al. (1998) found that neither minimal nor extensive voir dire responses predicted damage awards, other researchers have reported significant correlations between specific attitudes and the size of damage awards. For example, one study that surveyed actual jurors after their trials (Hans & Lofquist, 1992) found a negative correlation between the size of damage awards and posttrial beliefs regarding the existence of a civil litigation explosion. Vignette studies have demonstrated a similar negative correlation between damage awards and endorsement of so-called "tort reform" (Greene et al., 1991; Hans, 2000a; Moran et al., 1994). In summary, while traditional, very limited voir dire has been found to be largely ineffective (Fulero & Penrod, 1990; Hans & Jehle, 2003), it seems likely that more extensive voir dire could help attorneys identify potentially biased jurors in civil cases (Hans & Vidmar, 2007). This project compares the relative effectiveness of traditional minimal voir dire to extended voir dire.

The Effect of Judicial Rehabilitation

It is possible for people to manage potential sources of biases if they are aware of those biases, are motivated to change them, and have cognitive capacity to control them (Devine et al., 2012; Gawronski et al., 2006). The first two requirements could potentially be achieved by judicial rehabilitation. To our knowledge, there are no tests of this conjecture in the civil context.

This line of thinking is consistent with critics of extended voir dire, who argue that lengthy questioning is unnecessary because even if jurors do hold biasing attitudes, judges can effectively prevent juror bias from affecting their decisions via judicial rehabilitation. They argue that explicitly instructing jurors to control any biases they may have is more efficient than trying to identify and exclude all jurors that are biased. Extended voir dire advocates, however, maintain that there is no magic formula to prevent jurors from being influenced by their biases (Bennett, 2010; *Montgomery v. Kentucky*, 1991; *Utah v. Saunders*, 1999).

Judges have great latitude in conducting judicial rehabilitation (e.g., *People v. Griffin*, 1998; *Thomas ex rel. Thomas v. Mercy Hospitals East Communities*, 2017). Some argue that asking jurors yes/no questions about bias and securing a pledge to put them aside is inadequate to ensure a fair trial (*New York v. Biondo*, 1977; *Utah v. Saunders*, 1999). Yet, decisions about whether to remove jurors for cause hinge largely on whether they reveal a potential source of bias, and whether they pledge to put that bias aside (Hannaford-Agor & Waters, 2004).

Calling attention to biases and providing motivation to change are necessary for controlling bias (Devine et al., 2012), but they may not be sufficient. Attempts to control bias are often unsuccessful, or at best slightly successful—even when participants are highly motivated to be unbiased (Gawronski et al., 2006). Even when instructions to regulate bias are effective, they often only reduce, rather than eliminate, bias (Axt et al., 2018), result in small effect sizes (Forscher et al., 2019), and are short-lived (Lai et al., 2016).

Research in the *criminal* context has demonstrated that exposing mock jurors to a yes/no question during voir dire about whether their ability to be impartial would be affected by the fact that the defendant was Black did not reduce biases against Black (vs. White) defendants (Schuller et al., 2009). A similarly discouraging result involves the biasing effect of attitudes toward the insanity defense (Crocker & Kovera, 2010). As expected, mock jurors who acknowledged negative attitudes about the insanity defense were more likely to vote guilty than those who did not report holding such views. Relevant to our present concerns, however, being randomly assigned to undergo judicial rehabilitation did not reduce the degree to which mock jurors' preexisting attitudes about the insanity defense affected their dichotomous verdicts relative to those who did not experience judicial rehabilitation.

Posttrial interviews with actual jurors have cast further doubt on the utility of judicial rehabilitation. In one study, only 8% of Canadian jurors were willing to say they could not be impartial (Schuller et al., 2015). Although voir dire transcripts revealed that

the overwhelming majority of jurors asked for a public commitment to be impartial during voir dire agreed to do so (Johnson & Haney, 1994), many jurors interviewed afterward admitted that they did not even know what "impartial" meant or gave incorrect definitions. Moreover, several jurors admitted that they failed to keep their promise to be impartial. Indeed, jurors opined that "no person alive could do that," and that "it's a worthwhile goal, but cannot be achieved" (Johnson & Haney, 1994, p. 499). In short, minimal voir dire and judicial rehabilitation may exacerbate the problem of demand characteristics—the "right" answers to these types of questions are obvious and posing them may discourage jurors from acknowledging their biases.

The less-than-promising judicial rehabilitation findings are consistent with a larger literature demonstrating that judicial limiting instructions, in general, are ineffective (Kerr et al., 1991; Steblay et al., 2006). Further, instructing mock jurors to ignore their biases could increase, rather than decrease, the impact of their biases (Lieberman & Arndt, 2000). Telling people to ignore something often draws more attention to it and makes it more difficult to ignore (Wegner & Schneider, 2003) and telling people to suppress prejudice can actually increase the impact of that prejudice (Broeder, 1959; Macrae et al., 1994; Payne et al., 2002). Having jurors promise to set aside their biases might give them a false sense of security that they corrected their biases, which in turn could make them more comfortable expressing bias afterward (Effron et al., 2009). Thus, we also tested the impact of judicial rehabilitation on mock jurors' *perception* of their bias.

Research Overview and Hypotheses

The first goal of the present research was to determine whether more extensive voir dire procedures could identify specific preexisting attitudes likely to bias mock jurors' judgments in civil cases that standard, more minimal, voir dire procedures fail to reveal. Our second goal was to test the effectiveness of judicial rehabilitation procedures in "debiasing" potential jurors.

To pursue these goals, we conducted three separate mock juror experiments with similar methodologies, each presenting mock jurors with a different civil case, which we collapsed across for analyses (N = 2,041). In all three experiments, all mock jurors read a detailed summary of evidence and arguments and then indicated the verdict they favored, and what monetary damages, if any, they would award to the plaintiff. Before viewing and judging the case, however, mock jurors were randomly assigned to answer either (a) no voir dire questions, (b) traditional minimal voir dire questions, or (c) a combination of the minimal and extended voir dire questions before viewing the trial evidence. Mock jurors were also randomly assigned to either view or not view a judicial rehabilitation video in which a judge asked them whether they could set aside their biases and be impartial in rendering their judgments before they viewed the case materials. Finally, they reported their awareness of the potentially biasing effect of their views about litigation and litigants that they were asked about during voir dire on their case judgments.

Hypothesis 1: Effectiveness of Voir Dire

We predicted that extended voir dire questions that assessed specific attitudes relevant to civil litigation would provide substantially more significant predictors of case judgments than would demographic and minimal voir dire questions.

Hypothesis 2: Effectiveness of Judicial Rehabilitation

First, consistent with previous studies in the criminal realm, we predicted that very few mock jurors would report that they could not set aside their biases to be fair and impartial in response to a question posed by a judge (Hypothesis 2a). Second, based on the only previous study documenting the negligible effectiveness of judicial rehabilitation on verdicts (Crocker & Kovera, 2010) and previous research on the general ineffectiveness of judicial limiting instructions, we hypothesized that judicial rehabilitation would do little (if anything) to weaken the link between potentially biasing preexisting attitudes and case judgments (Hypothesis 2b). That is, we did not expect that judicial rehabilitation would moderate the effect of extended voir dire responses on case judgments.

Hypothesis 3: Variance In case Judgments Explained by Voir Dire

We hypothesized that when all voir dire responses were entered simultaneously in one stepwise regression, extended voir dire responses would explain significantly more variance in verdicts and damage awards than would demographic and minimal voir dire responses.

Hypothesis 4: Excludable Versus Nonexcludable Jurors

We predicted that extended voir dire responses would identify substantially more mock jurors who would be suitable candidates for exclusion (i.e., mock jurors whose attitudes and biases might make them unlikely to follow the law and judicial instructions, and to evaluate evidence open-mindedly), than would minimal voir dire questions. Further, we predicted that these "excludable" mock jurors' case judgments would differ significantly from those of mock jurors who do not express such extreme views.

Bias Awareness

Finally, we conducted an exploratory analysis to determine whether extended voir dire and/or judicial rehabilitation would increase mock jurors' awareness of factors that might have biased their decisions.

Method

We conducted three mock juror experiments simultaneously—each with the same procedure and design, differing only in the case that participants considered. All experiments employed a 3 voir dire (none, minimal, extended) × 2 judicial rehabilitation (absent, present) between-subjects design. The current research was approved by the Arizona State University IRB (#00001858). All materials and data can be found at https://osf.io/685qa/.

Participants

The participants for all three experiments (total N = 2,567) were recruited via Amazon's Mechanical Turk (MTurk), an online source commonly used in psychological research. MTurk samples are more demographically diverse than traditional "convenience" samples (e.g., college students or other online community member samples), generally yield the same results as nationally representative

samples (Coppock, 2019), and are considered to be a legitimate source of quality data (Buhrmester et al., 2011; Irvine et al., 2018). Nonoverlapping samples participated in the three experiments. We used Cloud Research to manage the MTurk workers, which allowed several avenues for quality assurance screening: We screened out participants who were not in the U.S. (including participants in other countries using VPNs), participants previously identified as using message boards to skirt attention checks, and those who did not have MTurk worker scores of at least 80%.

Data Exclusions

Data from a total of 526 participants were excluded (20%; an exclusion rate typical for MTurk samples, Goodman et al., 2013) for one or more of the following reasons. We excluded participants for failing any one of three general attention checks (AC) in which we requested participants choose a specific answer, such as telling them to choose *somewhat agree* on an agreement scale (AC 1: n = 48, 1.9%; AC 2: n = 166, 6.5%; AC 3: n = 32, 1.2%). Participants were asked to report demographic information twice during the survey and we excluded them if they failed to report the same gender (n = 25, 1.0%), age (n = 18, .7%), ethnicity (n = 33, 1.3%), or education level (n = 37, 1.5%) the second time the question was asked. We excluded participants who failed the judicial rehabilitation manipulation check (i.e., failing to correctly report whether they saw a video of the judge; n = 68, 2.6%). Participants were asked three comprehension checks right after the trial stimulus. We excluded participants if they got two or more of the three questions wrong (n = 181, 7%). Finally, as predicted, there were only a few participants who were randomly assigned to see the judicial rehabilitation video and when asked if they could be impartial said no (n = 11, .4%). Because the number of these participants was too small to be useful for comparison purposes, these participants were excluded. This finding speaks to our Hypothesis 2a that judicial rehabilitation would not be useful in identifying jurors who cannot be impartial because almost all mock jurors agreed.

Final Sample

The remaining 2,041 participants were 62% female; and 77% White, 9% African American, 6% Asian or Pacific Islander, 6% Hispanic or Latino/a, and 2% chose "Other;" with a mean age of 40 years (range = 18 to 83, SD = 12.6 years). Just over half (53%) of the sample were parents. Furthermore, 40.5% were married, 39.3% had never been married, 10.5% were divorced, 3.5% were widowed, 4.3% were partnered, and 2.1% were separated. With respect to education level, .4% had not finished high school, 8.7% had completed a high school degree or equivalent, 23.5% had finished some college, 12.3% had obtained an associate's degree, 3.0% had completed technical or vocational school, 37.7% had obtained a bachelor's degree, 12.7% had completed a graduate or professional degree, and 1.6% had earned a doctoral degree. The average annual household income of the participants was $59,377 ($SD$ = $43,181, median = $50,000). Demographics were similar across studies (see online supplemental materials, Table S1).

Power Analyses

For verdicts, we conducted a power simulation for the most complex potential interaction (i.e., a continuous extended voir dire response × 3 [voir dire] × 2 [*judicial rehabilitation*] three-way interaction) and all lower order interactions and main effects, given a dichotomous outcome variable at 90% power and based on a small effect size (n_p^2 = .01), which required 1,450 participants. Thus, after collapsing across the three studies (N = 2,041), we were well-powered to analyze verdicts. We had damage awards from mock jurors who voted liable (N = 1,246). A second similar power simulation that accounted for the nonnormal distribution of damage awards indicated that the 1,246 participants who made damage award judgments gave us 85% power to detect the small effect size (n_p^2 = .01). Thus, collapsing across the three cases gave us sufficient power to analyze both dichotomous verdicts and damage awards.

Materials

Trial Stimuli

In all three studies, the trial stimuli (which were based on actual civil cases that had been litigated in an American court) were quite detailed and took participants (who passed attention and manipulation checks) an average of 19–27 minutes to read, depending on the case (Medical Malpractice Misdiagnosis case: M = 19, SD = 9; Insurance Bad Faith case: M = 22, SD = 10; Wrongful Birth case: M = 27, SD = 12). Each trial stimulus included a summary of factual background, opening statements, direct- and cross-examination of witnesses, and closing arguments. Brief descriptions of each case follow. Our goal was to test whether our findings generalized across diverse cases that differed in content, evidence strength, and potential for large damage awards. Given that we did not expect that the effects of our experimental manipulations would differ by case, to reduce redundancy and to achieve sufficient statistical power, we collapsed response data across these studies.

Insurance Bad Faith Case. The plaintiff alleged that after a car accident, her insurance company did not pay her the amount owed and acted in bad faith. The defense argued that her attorney failed to cooperate in order to create and justify a bad faith claim. The trial stimulus presentation included 32 slides and 5,676 words.

Wrongful Birth Case. The plaintiff alleged that a genetic testing company (the defendant) was liable for her son's wrongful birth after they failed to identify the risk that her child would have cystic fibrosis. She alleged that if she had received the information from these tests, she would have terminated the pregnancy. The defense argued that the lab never received an order for the test from the doctor, that they were prohibited from notifying the doctor about the lack of that test, and that any harm suffered by the plaintiff was minimal because her son was functioning normally. The trial stimulus presentation included 40 slides and 6,860 words.

Medical Malpractice Misdiagnosis Case. The plaintiff alleged that the defendant misdiagnosed and improperly dismissed her husband's medical emergency, causing him to die at home. The plaintiff also alleged that the hospital was negligent by allowing the doctor to change and update notes in the medical charts weeks after seeing the patient. The defense argued that the defendant did not take care of himself and refused testing that would have saved his life. The trial stimulus presentation included 32 slides and 4,469 words.

In all studies, we provided participants with juror instructions regarding the burden of proof, case-relevant legal definitions, and case-specific requirements (depending on the case and claims) for reaching a verdict on liability and damages.

Participants (who passed attention and manipulation checks) took, on average, 54 min (*SD* = 20) to complete the studies (Medical Malpractice Misdiagnosis case: *M* = 49, *SD* = 17; Insurance Bad Faith case: *M* = 54, *SD* = 18; Wrongful Birth case: *M* = 61, *SD* = 21).

Judicial Rehabilitation Manipulation

Participants in the judicial rehabilitation condition saw a brief video of a judge in a courtroom who noted that they might have views that could bias their judgments in the case. He asked them whether they could "put aside" such views and biases and "apply the law as it is given"; participants replied *yes* or *no* (*Walls v. Kim*, 2001).

Voir Dire Manipulation

All participants across all studies ultimately answered all demographic and voir dire questions; however, we manipulated *when* they answered those questions. Prior to viewing and judging the case, participants in the no voir dire control condition answered only basic demographic questions in the prescreening phase. These participants completed the minimal voir dire measures and extended voir dire measures only after they had considered the case and made all of their case judgments. Participants randomly assigned to the minimal voir dire condition completed the prescreening demographic questions and the minimal voir dire measures before judging the case and answered the extended voir dire questions after judging the case. Participants in the extended voir dire condition completed the prescreening demographic questions, the minimal voir dire questions, and the extended voir dire questions before they viewed and judged the case.

Prescreening Demographic Items

All participants began the survey by indicating their gender, age, race and ethnicity, education, yearly income, marital status, and parental status.

Minimal Voir Dire Measures

The minimal voir dire questions were taken from an example set of voir dire questions for civil trials posted online by federal judge Susan Oki Mollway (Mollway, n.d.). These minimal voir dire measures were reviewed by a trial attorney and jury consultant and considered to be representative of common questions that attorneys or judges might ask when their time is highly limited (Bermant, 1982) and typical of the types of questions some judges list in their local courtroom rules. These questions included a set of yes/no questions that asked about prior experiences with the legal system and asked them to self-identify potential sources of bias or prejudice that might make them unable to be impartial (see Table 1 for verbatim wording of these questions and descriptive statistics). These questions are similar to those used in previous experiments that featured standard, minimal voir dire questions (e.g., Sommers, 2006) and are reflective of those commonly used in U.S. courts (Bermant, 1982; Hans & Jehle, 2003).

Extended Voir Dire Measures

The extended voir dire questions included those that an attorney might pose if given the opportunity to probe specific potential sources of bias rather than relying on jurors to self-identify any views that might compromise their objectivity (Hans, 2000b; Hans & Jehle, 2003). Many of the questions (described in detail below) have been used by attorneys and jury consultants when helping to select juries in actual cases and have been approved for use in voir dire (e.g., *In re: Bard IVC Filters Products Liability Litigation*, 2018). The others were derived from previous research on civil juries (Hans, 2000a, 2008; Hans & Vadino, 2007; Lazarus, 1999). Further, many research projects have used variations of the items tapping attitudes toward the civil justice system and tort reform (e.g., Greene et al., 1991; Hans & Lofquist, 1994; Moran et al., 1994). Trial tactics and jury selection manuals also commonly

Table 1
Minimal Voir Dire Responses Predicting Verdicts and Damage Awards

Voir dire questions	*n* (%) Yes (*N* = 2,014)	Effect on verdicts χ^2 (1, 2,037)	Effect on verdicts *p*	Effect on damages χ^2 (1, 1,241)	Effect on damages *p*
Have you previously served as a juror either in a criminal or civil case?	316 (15%)	0.22	.64	0.00	.88
Have you served as either a state or federal grand juror?	66 (3%)	0.54	.46	0.50	.46
Do you know of any reason you may be prejudiced for or against the plaintiffs or defendants because of the nature of the case, or otherwise?	40 (2%)	0.48	.49	0.20	.64
Are you a lawyer, married to a lawyer, or in a substantial relationship with a lawyer?	37 (2%)	0.03	.86	1.00	.32
Have you studied law or worked in a law office?	129 (6%)	0.50	.48	0.10	.71
Have you or a close family member sued or been sued by someone?	316 (15%)	0.55	.46	0.00	.96
Have you or a close family member ever testified in a lawsuit?	184 (9%)	0.43	.51	0.80	.38
You may be called upon in this case to decide liability and/or award money damages. Do any of you have any religious, philosophical, or other belief that prevents you from acting as an impartial juror in this case?	23 (1%)	2.41	.12	0.00	.96
Do you have any qualms about attempting to come to a verdict at the end of the case?	36 (2%)	1.07	.30	0.10	.79
Have you, any member of your family, or any very close personal friend ever engaged in investigating or otherwise acting upon claims for damages?	82 (4%)	0.06	.81	0.90	.34
Do you know of any reason that would prevent you from sitting in this case with complete fairness and impartiality and decide the case based only on the evidence presented in court and the law as given at the conclusion of the trial?	28 (1%)	3.92	.05	0.90	.33

Note. Questions are taken from Mollway (n.d.).

recommend similar items (e.g., Frederick, 2005; National Jury Project, 2000; Read, 2007).

We chose questions that attorneys could use in almost any civil case, comprising three general categories, described below. To help with readability of our results, we reverse scored some measures so that higher numbers (except where noted) indicate prodefense/antiplaintiff attitudes and lower numbers indicate antidefense/proplaintiff attitudes. See Table 2 for descriptive statistics and correlations among extended voir dire measures.

Preexisting Attitudes About Trial Participants

Burden of Convincing (BOC). Participants were asked, "In a case like this, do you think the plaintiff or defense is going to have a harder time convincing you?" They responded using a 5-point scale ranging from 1 = *the defense will have a much harder time* to 5 = *the plaintiff will have a much harder time*, with the midpoint labeled *neither*.

Trust. Similar to scales assessing trust in parties involved in the criminal justice system (e.g., police, attorneys, the DA) that successfully predict case judgments in criminal cases (e.g., Juror Bias Scale, Kassin & Wrightsman, 1983; Myers & Lecci, 1998; Pretrial Attitudes Questionnaire, Lecci & Myers, 2008), we assessed participants' trust in parties commonly involved in civil cases (i.e., attorneys, doctors, plaintiffs, insurance companies). Participants were asked a set of questions indicating the degree to which they trusted doctors, lawyers, people who sue others (i.e., plaintiffs), and insurance companies (Hans, 2008; Lazarus, 1999). They responded on 4-point scales, choosing from 1 = *none*, 2 = *only a little*, 3 = *some*, or 4 = *a great deal*. These four items were presented in a randomized order.

Likelihood of Fraudulent Claims Scale. This scale comprised three items designed to assess the degree to which participants were predisposed to think that civil cases are due to fraudulent claims or valid claims being denied (Hans & Vadino, 2007; Lazarus, 1999). Depending on whether they judged one of the medical malpractice cases (Wrongful Birth, Medical Malpractice Misdiagnosis) or the Insurance Bad Faith case, participants were asked about medical professionals or insurance companies, respectively. For example, one question asked which they thought was more likely: 1 = [medical professionals/insurance companies] *deny a valid claim*, 2 = *both are equally likely*, or 3 = *a person makes a fraudulent claim against a* [medical professional/insurance company]. Averaging responses to the three questions provided a scale with acceptable reliability ($\alpha = .69$).

Table 2
Descriptive Statistics and Correlations (p Values) Among Extended Voir Dire Responses

		Correlations														
Predictors	*M (SD)*	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Burden of convincing	2.94	**.10**	.02	.04	**–.12**	**.13**	**.06**	–.05	**.32**	**.25**	**.42**	**.17**	**.19**	**.12**	.02	–.03
(1)	(1.08)	**<.001**	.49	.04	**<.001**	**<.001**	**.01**	.03	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	.33	.15
Trust in doctors (2)	3.29		**.37**	**.27**	**.19**	.04	**–.19**	**–.22**	**.17**	**.16**	**.07**	**.08**	**.10**	.03	–.02	.04
	(0.69)		**<.001**	**<.001**	**<.001**	.18	**<.001**	**<.001**	**<.001**	**<.001**	**.002**	**<.001**	**<.001**	.13	.48	.05
Trust in lawyers (3)	2.52			**.47**	**.46**	–.04	**–.40**	**–.40**	**.07**	**.15**	**–.07**	**–.06**	**–.13**	**–.15**	–.03	**.07**
	(0.75)			**<.001**	**<.001**	.04	**<.001**	**<.001**	**.001**	**<.001**	**.001**	**.01**	**<.001**	**<.001**	.13	**.002**
Trust in insurance	2.16				**.33**	.005	**–.15**	**–.17**	**.14**	**.17**	.02	–.01	.01	.04	**.12**	**–.13**
companies (4)	(0.79)				**<.001**	.88	**<.001**	**<.001**	**<.001**	**<.001**	.34	.83	.69	.05	**<.001**	**<.001**
Trust in plaintiffs (5)	2.52					**–.15**	**–.29**	**–.25**	**–.10**	–.01	**–.29**	**–.14**	**–.22**	**–.24**	–.04	**.09**
	(0.66)					**<.001**	**<.001**	**<.001**	**<.001**	.76	**<.001**	**<.001**	**<.001**	**<.001**	.05	**<.001**
Unwilling to award	1.15						**.06**	.03	**.19**	**.12**	**.24**	**.17**	**.23**	**.24**	.04	–.04
noneconomic dam-	(0.36)						**.004**	.19	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	.07	.05
ages (6)																
Plaintiff attorneys are	2.91							**.77**	–.03	**–.14**	**.21**	.04	**.10**	**.16**	**.06**	**–.13**
liars (7)	(1.36)							**<.001**	.16	**<.001**	**<.001**	.06	**<.001**	**<.001**	**.008**	**<.001**
Defense attorneys are	3.05								**–.10**	**–.21**	**.07**	–.01	**.06**	**.14**	**.07**	**–.14**
liars (8)	(1.41)								**<.001**	**<.001**	**.003**	.66	**.007**	**<.001**	**.002**	**<.001**
Likelihood of	1.96									**.56**	**.51**	**.19**	**.28**	**.19**	**.05**	**–.08**
Fraudulent Claims	(0.51)									**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**.02**	**<.001**
Scale (9)																
Defendant Dishonesty	3.23										**.38**	**.13**	**.15**	**.08**	.01	–.01
Scale (10)	(0.95)										**<.001**	**<.001**	**<.001**	**.001**	.62	.74
Motives for Litigation	3.68											**.25**	**.41**	**.34**	**.10**	**–.14**
Scale (11)	(0.90)											**<.001**	**<.001**	**<.001**	**<.001**	**<.001**
Burden of proof is too	3.56												**.29**	**.22**	.04	**–.07**
low (12)	(0.80)												**<.001**	**<.001**	.05	**.002**
Negative Attitudes to-	3.63													**.45**	**.17**	**–.20**
ward Lawsuits	(0.78)													**<.001**	**<.001**	**<.001**
Scale (13)																
Limit Litigation Scale	2.49														**.14**	**–.18**
(14)	(0.61)														**<.001**	**<.001**
Trump Approval Scale	1.31															**–.67**
(15)	(0.45)															**<.001**
Liberalism (16)	3.83															
	(1.35)															

Note. Statistically significant differences at $p \leq .01$ are bolded.

Likelihood of Defendant Dishonesty Scale. This scale was designed to assess the degree to which participants might be predisposed to believe that defendants who are being sued are dishonest and prone to deny valid claims (Hans & Vadino, 2007; Lazarus, 1999). Depending on whether they judged one of the medical malpractice cases (Wrongful Birth, Medical Malpractice Misdiagnosis) or the Insurance Bad Faith case, they were asked about medical professionals or insurance companies, respectively. They were asked "How often do you think the following things happen? [medical professionals/insurance companies] decline valid claims for minor injuries because they think the person making the claim will drop it rather than pursue the issue and sue for the claim"; and (2) "[medical professionals/insurance companies] deny causing major injuries because they are trying to hold on to the money for as long as they can before they have to pay it." Participants were asked to rate the likelihood of these two rationales for claims denials on 6-point scales ranging from *always* to *never*. Averaging responses to the two items created a reliable scale (Spearman Brown Coefficient = .77). These two items were presented in a randomized order.

Motives for Litigation Scale. This scale assessed how much participants ascribed bad motives to plaintiffs and/or defendants (Hans, 2000a). They responded to five items (e.g., "If there is a trial like this it is much more likely that the plaintiffs are out for money than the [doctor making a mistake/insurance company is being dishonest]") using 7-point scales, with options ranging from 1 = *strongly disagree* to 7 = *strongly agree*. Averaging responses to these five questions provided a scale with acceptable reliability (α =.61) according to some (Murphy & Davidshofer, 1988), but falls short of the .70 threshold for "good" reliability (Nunnally, 1978).

Impressions of Attorneys. We measured how much participants agreed that "Most defense attorneys are liars" (i.e., higher numbers being a more proplaintiff attitude) and "Most plaintiff attorneys are liars" using two 7-point scales, with response options ranging from *strongly disagree* to *strongly agree*. These questions have been used on actual juror questionnaires.

The five items comprising the Motives for Litigation Scale and the two impressions of attorneys items were presented together and in a randomized order.

Support for Civil Litigation

Attitudes Toward Burden of Proof Threshold. This item assessed how much potential mock jurors agreed with the law regarding burden of proof. We gave participants a detailed explanation of the burden of proof criterion in civil cases, including the implication that "even if a plaintiff proves something is only 51% likely to be true, and it is 49% possible it is not true, the plaintiff has met their burden of proof and should win their case." They then indicated how they felt about this criterion using a 5-point scale, ranging from 1 = *way too high (the plaintiff should be required to reach a lower level of proof to win*) to 5 = *way too low (the plaintiff should be required to reach a higher level of proof to win)*, with a midpoint of 3 (*neither too low nor too high*). These questions have been used on actual juror questionnaires. Extreme values reflect a potentially problematic unwillingness to follow the law.

Unwillingness to Award Noneconomic Damages. We gave participants a detailed explanation of the difference between economic and noneconomic damages and asked them to indicate whether they were either *comfortable* (coded 0) or *not comfortable* (coded 1) awarding noneconomic damages. This question has been used on actual juror questionnaires. The *not comfortable* response suggests a bias that would disadvantage plaintiffs and flag potential unwillingness to follow the law regarding noneconomic damages.

Negative Attitudes Toward Lawsuits Scale. This scale was designed to assess the degree to which potential mock jurors hold preexisting negative attitudes toward lawsuits and damage awards in general (similar to measures in Hans, 2000a; Hans & Lofquist, 1994). We asked the degree to which they thought (a) the number of personal injury lawsuits, and (b) the amount of money awarded in such lawsuits, were 1 = *way too low* to 5 = *way too high*, with a midpoint of 3 = *just right*. Averaging responses to these two items provided a reliable scale (Spearman Brown Coefficient = .73). These items were presented in a randomized order.

Limit Litigation Scale. This scale assessed the degree to which participants supported efforts to curb litigation in general. They indicated their level of support for three efforts (e.g., placing a limit on how much a judge or a jury can award in a lawsuit resulting from a person being injured) on 4-point scales ranging from *strongly oppose* to *strongly support*. These questions have been used on actual juror questionnaires. Dropping one item (placing a limit on how much an attorney who represents an injured person in a lawsuit can charge for his or her services) increased the reliability of the scale (i.e., the Spearman Brown increased from .36 to .62). We averaged the two remaining items to create a Limit Litigation Scale. These items were presented in a randomized order.

Political Attitudes

Political Orientation. Participants rated their political orientation using a 7-point scale (1 = *extremely conservative*, 7 = *extremely liberal*). Higher numbers reflect more liberalism in general (as opposed to being more proplaintiff or prodefense).

Trump Approval Scale. Participants answered two yes/no questions regarding their support for President Trump ("Do you approve of the job President Trump is doing?" and "If the presidential election were today, would you vote for Trump?") by choosing either 1 = *no* or 2 = *yes*. Averaging responses to these two items provided a reliable scale (Spearman Brown Coefficient = .94). These questions have been used on actual juror questionnaires. Higher numbers reflected political attitudes more supportive of Trump in general (as opposed to being more proplaintiff or prodefense).

Coding for "Excludable" Mock Jurors

As an alternative way to test the utility of extended voir dire questions with even more practical importance, we used the extended voir dire responses to identify mock jurors who expressed extreme attitudes that could potentially get them challenged for cause (or excluded via a peremptory challenge) to see if they judged the case differently than those who did not. We coded each mock juror as "excludable" if their responses to extended voir dire questions revealed attitudes suggesting that the participant was relatively less likely to follow the law and judicial

instructions regarding impartiality—that is, someone who might give the judge or attorneys cause to strike them from the jury.

In jurisdictions that tend to consistently allow exclusions for cause, an extreme view on either (a) the rules that will govern the case, or (b) a clear bias related to the specific facts of the case could support exclusion from the jury. For example, Florida courts require that jurors impartiality is proven beyond a reasonable doubt (e.g., *Frogel v. Philip Morris U.S.A., Inc.*, 2020; *Jackson v. State*, 2017), Pennsylvania courts have said that "jurors should be above suspicion" (*Shinal v. Toms*, 2017), and Missouri's Supreme Court stated that jurors should be excluded if their views expressed in voir dire would "substantially impair the performance of his duties as a juror" (*State v. Kreutzer*, 1996). Thus, if courts embrace exclusions for cause, they would, at a minimum, exclude jurors who express extreme views that (a) could threaten their impartiality or (b) prevent them from following the law. To be conservative, we coded mock jurors as "excludable" if they met at least one of the following criteria that clearly indicated one of these two factors (described below). This operationalization is based on the idea that any one of these would give an attorney a strong argument for excluding the juror for cause or, if that failed, a reason to use a peremptory challenge.

Burden of Convincing

Mock jurors who chose the most extreme values of this scale and reported that either the *defense would have a much harder time convincing them* (9.6%) or that the *plaintiff would have a much harder time convincing them* (7.3%). This represents an extreme predisposition that might threaten their impartiality.

Attitudes Toward Burden of Proof Threshold

Mock jurors who chose the most extreme values of this item and reported that they thought that the burden of proof was either *way too high* (0.6%) or *way too low* (14.4%). This represents potential difficulty following the law (i.e., respecting the burden of proof threshold).

Support for Limiting Litigation

Mock jurors who chose the highest value of this scale reported that they *strongly support* making it harder to sue any person, business, or organization that injures another person either intentionally or through carelessness (3.6%) and/or placing a limit on how much a judge or jury can award in a lawsuit (10.8%). This represents an extreme predisposition against lawsuits and high damage awards that might threaten their impartiality and willingness to follow the law.

Willingness to Award Noneconomic Damages

Mock jurors who reported that they were *not willing* to award noneconomic damages (15.1%). This represents evidence of an extreme predisposition that might threaten their willingness to follow the law.

Dependent Measures

Case Judgments

After reading the case materials, participants chose a verdict (defendant *not liable* vs. *liable*). In the Wrongful Birth and Insurance Bad Faith studies, all participants were asked to determine the damage award they favored. In the Medical Malpractice Misdiagnosis study, only participants who judged the defendant as *liable* were asked to do so. Participants indicated their damage award by first writing it in numbers, and then (as a quality check) in words. Upon examining a few instances where words and numbers differed we found that most involved the adding/omitting of a zero. Accordingly, we used the written-out numbers.

Bias Awareness Scale

Bias awareness items were designed to assess the degree to which mock jurors were aware of and acknowledged the potential biasing factors that might have influenced their case judgments. Participants indicated how much they thought each factor might have influenced their decisions about liability and damage awards using 5-point scales ranging from 1 = *not at all* to 5 = *an extremely big impact*. The factors mirrored our extended voir dire items (i.e., attitudes toward doctors, plaintiffs, defense attorneys, plaintiff attorneys, the burden of proof standard, and the amount of money awarded in lawsuits; their political attitudes). Participants' level of awareness across these seven targets formed a reliable scale ($\alpha = .84$). We averaged these items to create Bias Awareness Scale scores.

Manipulation and Attention Checks

Participants completed three attention checks throughout the study that told them what response to choose, as well as three comprehension checks that assessed their attention to the trial stimulus (see online supplemental material for specific attention/comprehension check questions and item-specific failure rates). Participants were asked demographic questions (i.e., gender, age, ethnicity, education level) twice to confirm that they responded consistently (and presumably correctly). A judicial rehabilitation manipulation check asked participants to indicate whether they had viewed a video of a judge (*yes*, *no*, or *I do not remember*). The survey also included recall questions for bonusing participants (participants who answered > 50% of recall questions correctly received a $2.50 bonus).

Procedure

All potential participants completed a prescreening, which included basic demographic information (i.e., gender, age, ethnicity, income, education level) and attention checks. Those who failed these initial basic attention checks were not invited to complete the study. We asked participants who passed the prescreening to assume the role of a mock juror and, as is typical in civil cases, they were given a short overview of the case. Next, we randomly assigned them to one of the three voir dire conditions and they completed measures consistent with condition. Participants were also randomly assigned to either view or not view a brief judicial rehabilitation video before they considered the case evidence and rendered their judgments. Participants in all experimental conditions then read a detailed summary of evidence and arguments presented in one of three cases. After providing their verdict and damage award, they completed attention and comprehension checks, recall items, bias awareness questions, and any voir dire measures that they weren't assigned to

complete before the trial. All participants were compensated $2 and bonused when appropriate.

Results

Factor Analysis

Because of the relatively large number of extended voir dire questions posed to our participants, we conducted an exploratory factor analysis (EFA) to see if we could group and reduce the number of measures to form meaningful factors. Relying on a principle-axis factoring extraction and a Promax rotation, we identified two factors that were meaningful and usable for our purposes (for details, see online supplemental materials). The first factor comprised the Likelihood of Fraudulent Claims Scale (B = .81), the Defendant Dishonesty Scale (B = .66), and the Motives Scale (B = .65). We labeled this the *Skepticism Toward Litigants Factor*. The second factor, which comprised the political orientation item (B = −.85) and the Trump Approval Scale (B = .80), we labeled the *Political Attitudes Factor*. For these latent factors, we used factor scores as predictors in our regressions. Positive numbers reflect scores above the mean and negative numbers reflect scores below the mean, with each one-unit increase/decrease representing one standard deviation unit change.

Analysis Plan

We performed logistic regressions on binary verdict outcomes and report odds ratios (*OR*), which reflect the increase or decrease in likelihood that a mock juror would find for the plaintiff (for example, *OR* = 1 would mean no effect in either direction, *OR* = 2 would mean mock jurors would be twice as likely to find for the plaintiff for each step up the predictor scale, *OR* = .5 would mean mock jurors would be half as likely to find for the plaintiff for each step up the predictor scale).

Regarding damage awards, we analyzed damage awards only from mock jurors who voted liable (but include analyses with mock jurors who voted not liable in online supplemental materials, Table S4). QQ-plots (after outliers were replaced, see outlier analysis description below) indicated that, as is often the case with damage awards data (Greene et al., 2001), the data were non-normal and skewed (skewness = .85, kurtosis = −.07, Shapiro-Wilk W = .80, $p < .001$; See online supplemental materials, Figure S1 for QQ-plots). To deal with this skew, we analyzed damage awards data with generalized linear regressions using a gamma distribution, which is appropriate for data that are very right-skewed (Barber & Thompson, 2004). This analysis produces mean ratio coefficients (reported as exp(B) in the tables), which are interpreted similarly to *ORs* produced by logistic regressions. For example, if a means ratio is 2, that would mean that a one-unit increase in the predictor would produce twice as many dollars in damage awards; if a means ratio is .5, that would mean a one-unit increase in the predictor would produce half as many dollars in damage awards. Gamma distributions are unable to handle zeros, therefore we replaced the one participant who recommended $0's value with $1 (Fry et al., 2000).

Because we conducted a relatively large number of analyses to fully test our hypotheses and were predicting null moderator effects, we adopted a stringent level of $p \leq .01$, to evaluate our hypotheses (see Benjamin et al., 2018; e.g., Lai et al., 2016).

Preliminary Analyses

Outlier Analysis

We relied upon visual inspection of box plots created for both the combined data file, as well as for each of the three individual cases (see online supplemental materials, Figure S2). The box plots suggested the presence of at least two outliers in each case. We next conducted Rosner's (1983) tests on each of the three case-specific data files to confirm that these values were outliers. Rosner's test is a discordancy test that is capable of detecting multiple outliers and is suitable for samples with more than 20 observations. Although we conducted a generous test to catch outliers (i.e., Rosner test assumptions were set to allow for up to 10 outliers, which is the largest number the package creators recommend), across the three cases, the Rosner's tests identified only seven outliers (Insurance Bad Faith case: n = 2, Medical Malpractice Misdiagnosis case: n = 3, Wrongful Birth case: n = 2). We replaced these outliers with the next highest nonexcluded damage awards value from their specific case (Insurance Bad Faith case = $1,000,000, Medical Malpractice Misdiagnosis case = $51,370,000, Wrongful Birth case = $80,000,000) using winsorization (Tukey, 1962).

Case Differences in Legal Judgments

Collapsed across the three cases, 61% (n = 1,246) of participants found the defendant liable. As was expected, some plaintiff cases proved to be stronger than others. Relative to the Medical Malpractice Misdiagnosis case (38% liable verdicts), participants found the defendant liable significantly more often in both the Insurance Bad Faith case (71%), B = 1.40, OR = 4.08, $p < .001$, and in the Wrongful Birth case (76%), B = 1.68, OR = 5.36, $p < .001$. Verdicts in the Insurance Bad Faith and Wrongful Birth cases did not differ significantly, B = .27, OR = 1.31, p = .03.

Collapsed across the three cases after we replaced outliers, the average damage awards was $13,926,362 ($SD$ = $15,853,929; median = $5,000,000). However, damage awards differed significantly by case, $\chi^2(df = 2, N = 1{,}246) = 8{,}490.60$, $p < .001$. Bonferroni post hoc tests revealed that the Wrongful Birth case produced significantly higher damage awards (M = $20,478,826; SD = $16,860,678) than the Medical Malpractice Misdiagnosis case (M = $6,212,797; SD = $10,870,092), $p < .001$, which in turn led to significantly higher damage awards than the Insurance Bad Faith case (M = $355,326; SD = $190,700), $p < .001$.

Given the significant differences in overall verdicts and damage awards across cases, we statistically controlled for which case participants read in all of our analyses by including it as a categorical covariate. Although we did not predict that case would *moderate* any of our effects, it is clear that each case had very different baseline liability rates and average damage awards. Thus, in all of our main analyses, we statistically controlled for which case the participant read so that we could (a) account for the variance due to case (i.e., reduce noise); (b) ensure any effects we found existed above and beyond what case they read; and (c) account for a "Level 2" variable (i.e., data nested within case) that has less than 20 units and therefore is not appropriate for multilevel modeling (McNeish & Stapleton, 2016).

Hypothesis 1: Effectiveness of Minimal Voir Dire

We hypothesized that minimal voir dire would be less effective in predicting case judgments than extended voir dire. Supporting this hypothesis, a series of logistic regressions, each containing one of the minimal voir dire responses (e.g., previous jury service), revealed that *none* of the minimal voir dire responses significantly predicted verdicts or damage awards (see Table 1). Further, the first column of Table 1 reveals that most minimal voir dire questions elicited such a small number of affirmative answers that we were precluded from running analyses involving interactions with our experimental manipulations. In response to the three questions asking them to self-identify any personal prejudices, biases, or reasons they might not be impartial, 2% or less of the participants said that they could do so.

Hypothesis 1: Effectiveness of Extended Voir Dire

In contrast to minimal voir dire, we predicted that extended voir dire questions would be more effective in predicting mock jurors' liability verdicts and damage awards. To test this we ran a series of regressions, with each regression including: (a) one of the extended voir dire responses as a predictor (e.g., how much they trust plaintiffs); (b) the voir dire manipulation (Helmert coded); (c) the judicial rehabilitation manipulation; (d) all interactions among these variables; and (e) the case as a categorical covariate.

Table 3 displays the results of the logistic regressions predicting verdicts and Table 4 displays the results of the GLM regressions predicting damage awards. Each row represents a regression model that included the specified voir dire response as the focal predictor. The first two columns of statistics on the left side indicate the extent to which the specified voir dire response predicted verdicts (see Table 3) or damage awards (see Table 4); the right-hand column indicates whether there were any main effects and/or interaction effects involving experimental manipulations. For logistic regressions, confidence intervals including 1 are nonsignificant (see Table 3); for GLM regressions, confidence intervals including 0 are nonsignificant (see Table 4).

Verdicts

Table 3 reveals that almost all of the extended voir dire responses significantly predicted verdicts—except for trust in lawyers and the belief that defense attorneys are liars. Expressing relatively higher levels of trust in people who sue others (i.e., plaintiffs) was significantly associated with increased likelihood that mock jurors would side with the plaintiff (i.e., vote liable). All other significant negative *B* values reflected that higher values were significantly associated with decreased likelihood of siding with the plaintiff. Figure 1 depicts these relationships visually. This supports Hypothesis 1 in that extended voir dire questions provided substantially more significant predictors of verdicts than did minimal voir dire.

Damage Awards

Table 4 reveals that for eight of the 13 extended voir dire responses, mock jurors' responses significantly predicted damage awards. The exceptions were most of the trust variables (in doctors, lawyers, and insurance companies), the belief that defense attorneys are liars, and political attitudes. Trust in plaintiffs significantly predicted higher damage awards. For the other seven significant predictors, relatively higher scores were associated with significantly lower damage awards. Figure 2 depicts these relationships visually.

This supports Hypothesis 1 in that extended voir dire questions provided substantially more significant predictors of damage awards than did minimal voir dire.

Hypothesis 2: Effectiveness of Judicial Rehabilitation

Verdicts

We hypothesized that undergoing judicial rehabilitation would not reduce the biasing impact of preexisting attitudes on verdict decisions. That is, we predicted that judicial rehabilitation would not be a significant moderator of the relationships between

Table 3
The Prediction of Liability Verdicts From Extended Voir Dire Responses

Predictors	*B*(*SE*), *OR* [95% CI]	*p*	Experimental manipulation main effects & interactions
Burden of convincing	***B* = −0.75(0.06), *OR* = 0.47 [0.42, .53]**	**< .001**	Interaction with voir dire: **χ^2 (2, 2,017) = 28.73, *p* < .001** Others: χ^2s ≤ 3.61, *ps* ≥ .06
Trust in doctors	***B* = −0.35(0.07), *OR* = 0.70 [0.61, 0.81]**	**< .001**	None: χ^2s ≤ 2.40, *ps* ≥ .30
Trust in lawyers	*B* < −0.01(0.07), *OR* = 1.00 [0.88, 1.14]	.99	None: χ^2s ≤ 4.15, *ps* ≥ .12
Trust in insurance companies	***B* = −0.21(0.06), *OR* = 0.81 [0.71, 0.91]**	**<.001**	None: χ^2s ≤ 2.14, *ps* ≥ .34
Trust in plaintiffs	***B* = 0.46 (0.08), *OR* = 1.59 [1.37, 1.85]**	**< .001**	None: χ^2s ≤ 5.73, *ps* ≥ .06
Unwilling to award noneconomic damages	***B* = −1.12 (0.14), *OR* = 0.33 [0.25, 0.43]**	**< .001**	None: χ^2s ≤ 2.17, *ps* ≥ .34
Plaintiff attorneys are liars	***B* = −0.11(0.04), *OR* = 0.89 [0.83, 0.96]**	**.002**	None: χ^2s ≤ 3.05, *ps* ≥ .08
Defense attorneys are liars	*B* = 0.03(0.04), *OR* = 1.03 [0.96, 1.10]	.45	None: χ^2s ≤ 4.76, *ps* ≥ .09
Skepticism toward Litigants Factor	***B* = −1.28(0.08), *OR* = 0.28 [0.24, 0.32]**	**< .001**	None: χ^2s ≤ 7.60, *ps* ≥ .02
Burden of proof is too low	***B* = −0.46(0.06), *OR* = 0.63 [0.56, 0.71]**	**< .001**	None: χ^2 ≤ 5.62, *ps* ≥ .06
Negative Attitudes Toward Lawsuits Scale	***B* = −0.67(0.07), *OR* = 0.51 [0.45, 0.59]**	**< .001**	None: χ^2s ≤ 5.11, *ps* ≥ .08
Limit Litigation Scale	***B* = −0.59(0.08), *OR* = 0.55 [0.47, 0.65]**	**< .001**	None: χ^2s ≤ 3.07, *ps* ≥ .21
Political Attitudes Factor	***B* = −0.24(0.05), *OR* = 0.78 [0.70, 0.87]**	**< .001**	None: χ^2s ≤ 4.14, *ps* ≥ .13

Note. Statistically significant effects at $p \leq .01$ are bolded. For trust in plaintiffs and the belief that defense attorneys are liars, higher numbers represent more proplaintiff/antidefense views. For the Political Attitudes Factor, higher scores reflect greater conservatism. For all other predictors, higher numbers reflect more antiplaintiff/prodefense views.

Table 4
The Prediction of Damage Awards From Extended Voir Dire Responses

Predictors	*b*[95% CI]	Exp(*b*) [95% CI]	*p*	Experimental manipulation main effects and Interactions
Burden of convincing	**$b = -0.06$ [−0.09, −0.03]**	**$b = 0.94$ [0.91, 0.97]**	**$p = .001$**	None: χ^2s ≤ 5.06, all *p*s ≥ .08
Trust in doctors	$b = 0.01$ [−0.03, 0.05]	$b = 1.01$ [0.97, 1.06]	$p = .66$	None: χ^2s ≤ 6.30, all *p*s ≥ .04
Trust in lawyers	$b = 0.03$ [−0.02, 0.07]	$b = 1.03$ [0.98, 1.07]	$p = .21$	None: χ^2s ≤ 7.95, all *p*s ≥ .02
Trust in insurance companies	$b = 0.003$ [−0.04, 0.04]	$b = 1.00$ [0.96, 1.04]	$p = .90$	None: χ^2s ≤ 6.82, all *p*s ≥ .03
Trust in plaintiffs	**$b = 0.12$ [0.07, 0.17]**	**$b = 1.13$ [1.07, 1.19]**	**$p < .001$**	None: χ^2s ≤ 7.41, all *p*s ≥ .02
Award noneconomic damages	**$b = -0.29$ [−0.40, −0.18]**	**$b = 0.75$ [0.67, 0.84]**	**$p < .001$**	None: χ^2s ≤ 6.63, all *p*s ≥ .04
Plaintiff attorneys are liars	**$b = -0.04$ [−0.06, −0.02]**	**$b = 0.96$ [0.94, 0.98]**	**$p = .001$**	None: χ^2s ≤ 4.94, all *p*s ≥ .08
Defense attorneys are liars	$b = -0.02$ [−0.04, 0.001]	$b = 0.98$ [0.96, 1.00]	$p = .07$	None: χ^2s ≤ 5.57, all *p*s ≥ .06
Skepticism toward Litigants Factor	**$b = -0.11$ [−0.15, −0.07]**	**$b = 0.90$ [0.86, 0.94]**	**$p < .001$**	None: χ^2s ≤ 3.92, all *p*s ≥ .14
Burden of proof is too low	**$b = -0.07$ [−0.11, −0.03]**	**$b = 0.93$ [0.89, 0.98]**	**$p = .002$**	None: χ^2s ≤ 7.96, all *p*s ≥ .02
Negative Attitudes Toward Lawsuits Scale	**$b = -0.17$ [−0.21, −0.12]**	**$b = 0.85$ [0.81, 0.88]**	**$p < .001$**	None: χ^2s ≤ 6.08, all *p*s ≥ .05
Limit Litigation Scale	**$b = -0.12$ [−0.18, −0.07]**	**$b = 0.89$ [0.84, 0.93]**	**$p < .001$**	None: χ^2s ≤ 8.70, all *p*s ≥ .01
Political Attitudes Factor	$b = -0.01$ [−0.04, 0.03]	$b = 0.99$ [0.96, 1.03]	$p = .77$	None: χ^2s ≤ 5.60, all *p*s ≥ .06

Note. Statistically effects at at $p \leq .01$ are bolded. For trust in plaintiffs and the belief that defense attorneys are liars, higher numbers represent more proplaintiff/antidefense views. For the Political Attitudes Factor, higher scores reflect greater conservatism. For all other predictors, higher numbers reflect more antiplaintiff/prodefense views. Exp(*b*) represents the means ratio.

extended voir dire responses and case judgments. As is evident in the right-hand column of Table 3, there were no significant interactions between any of the preexisting attitudes and judicial rehabilitation on mock jurors' verdicts. Eleven of mock jurors' preexisting attitudes significantly influenced their verdict decisions, *regardless* of whether they underwent judicial rehabilitation. Thus, Hypothesis 2 is supported regarding verdict decisions: Many of mock jurors' preexisting attitudes identified by extended voir dire significantly influenced whether they found the defendant liable—to the same degree whether they got judicial rehabilitation or not.

There was one unexpected significant interaction between one extended voir dire response (burden of convincing) and which voir dire condition they were in—this effect was not relevant to our hypotheses but is followed up and described in the online supplemental materials.

Damage Awards

Analyses of damage awards also revealed no interaction effects with judicial rehabilitation (see Table 4). That is, the predictive power of preexisting attitudes on mock jurors' damage awards was not reduced by judicial rehabilitation. Thus, Hypothesis 2 is supported regarding damage awards as well: many of mock jurors' preexisting attitudes identified by extended voir dire significantly influenced how much money they awarded—to the same degree whether they got judicial rehabilitation or not.

Alternative Models

Our pattern of results was robust to many alternative models detailed in online supplemental materials. We found no substantial differences in the pattern of results, but detail minor deviations that we did find in online supplemental materials.

Hypothesis 3: Unique Variance in Case Judgments Explained by Extended Voir Dire Versus Minimal Voir Dire Responses

To determine what percentage of variation in mock jurors' verdicts can be explained by the minimal versus extended voir dire questions, we conducted stepwise regressions on verdicts and damage awards. The first step of the model included all basic demographic information: age, gender, ethnicity (Hispanic, Asian, Black, with White as the reference group), parental status, marital status (married, divorced, widowed, partnered, with never married as the reference group), yearly income, and education. The second step of the model included the minimal voir dire responses in addition to demographics. The third step added the extended voir dire responses. This model further allowed us to determine how much *unique* variance in verdicts or damage awards each predictor explains beyond all the other predictors in the model.

Verdicts

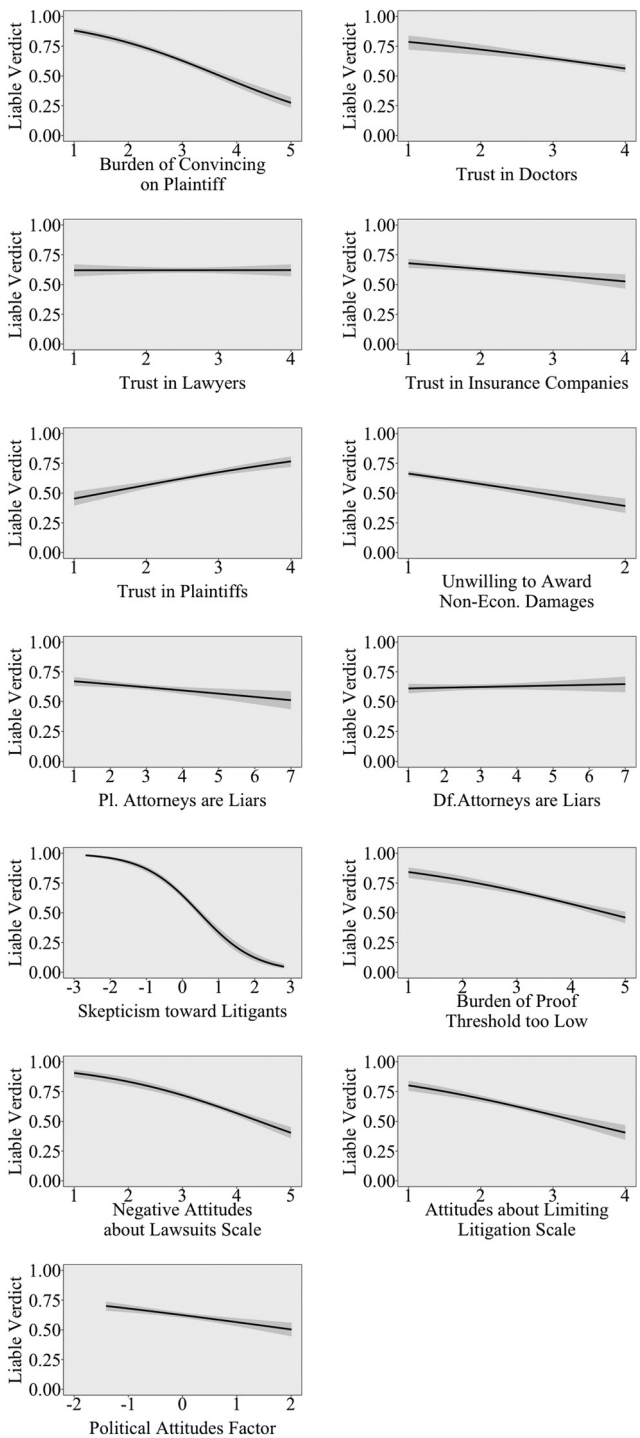
In Step 1, none of the demographic variables predicted verdicts, all *B*s ≤ |.49|, all *p*s ≥ .08. Furthermore, the combined demographic variables explained only .7% of the variation in verdicts, χ^2 ($df = 12$, $N = 2{,}041$) = 14.74, $p = .26$.

In the second step of the model, the minimal voir dire responses and demographic variables explained only 1% of the variation in verdicts. This step still was not statistically different from zero, χ^2 ($df = 11$, $N = 2{,}041$) = 11.75, $p = .38$. Further, none of the individual minimal voir dire responses predicted verdicts, *B*s ≤ |.89|, *p*s ≥ .05.

By contrast, the third step of the model, which added the extended voir dire responses, explained 31% of the variance in verdicts, an amount that was statistically significant, χ^2 (13, N = 2,041) = 696.75, $p < .0001$.

Four extended voir dire responses predicted a significant amount of unique variance in verdicts; that is, variance above and beyond that explained by all other demographic, minimal voir dire, and extended voir dire variables. Those items were (a) burden of convincing: $B = -.65$, $OR = .52$, 95% CI [.46, .59], $p < .001$; (b) unwillingness to awards noneconomic damages: $B = -.67$, OR = .51, 95% CI [.37, .70], $p < .001$; (c) Skepticism toward Litigants Factor: $B = -1.14$, $OR = .32$, 95% CI [.27, .38], $p < .001$; and (d) Negative Attitudes Toward Lawsuits Scale: $B = -.29$, $OR = .75$, 95% CI [.63, .90], $p = .002$.

Figure 1
The Relationships Between Extended Voir Dire Responses and Liability Verdicts



Note. For trust in plaintiffs and the belief that defense attorneys are liars, higher numbers represent more proplaintiff/antidefense views. For the Political Attitudes Factor, higher scores reflect greater conservatism. For all other predictors, higher numbers reflect more antiplaintiff/prodefense views.

Damage Awards

For damage awards, the stepwise regression was again based on a gamma distribution. In Step 1, the demographic variables did not explain a significant amount of the variance in damage awards, $F(12, N = 1{,}246) = 2.06$, $p = .02$, pseudo-$R^2 = .02$. None of the demographic variables were significant predictors of damages, all bs $\leq .38$, all ps $\geq .01$.

In the second step of the model, the minimal voir dire responses did not explain significantly more variation in damage awards than the demographic variables, $F(11, N = 1{,}246) = .95$, $p = .49$, pseudo-$R^2 = .02$. None of the minimal voir dire responses were significant predictors of damages, all bs $\leq .64$, all ps $\geq .02$.

By contrast, the third step of the model, which added the extended voir dire responses, explained significantly more variance in damage awards, beyond what the demographic and minimal voir dire questions explained, $F(13, N = 1{,}246) = 8.88$, $p < .001$, pseudo-$R^2 = .09$.

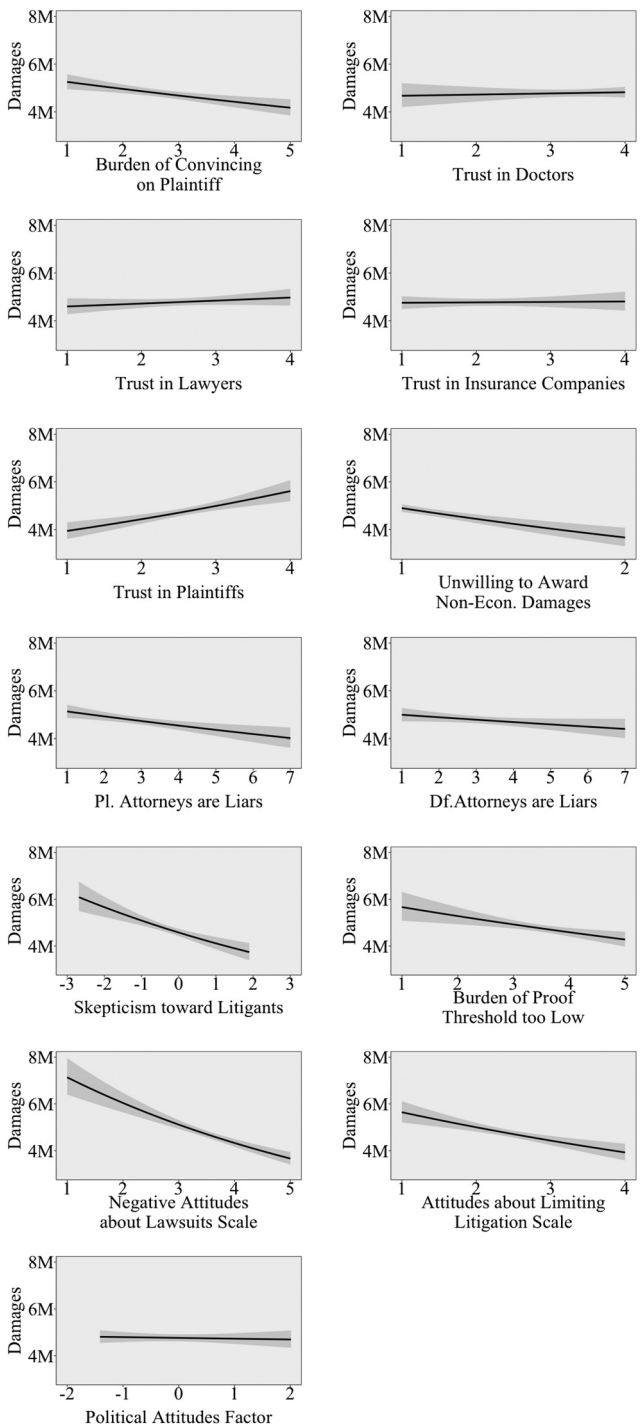
Eight extended voir dire responses accounted for a significant amount of unique variance in damages beyond what all of the other demographics and voir dire questions explained. Those items were: (a) burden of convincing: $b = -.11$, 95% CI [−.18, −.03], exp(b) = .90, 95% CI [.83, .97], $p = .008$; (b) trust in doctors: $b = -.17$, 95% CI [−.29, −.05], exp(b) = .85, 95% CI [.75, .96], $p = .006$; (c) trust in insurance companies: $b = -.34$, 95% CI [−.46, −.23], exp(b) = .71, 95% CI [.63, .80], $p < .001$; (d) trust in plaintiffs, $b = .30$, 95% CI [.15, .44], exp(b) = 1.35, 95% CI [1.16, 1.55], $p < .001$; (e) unwillingness to awards noneconomic damages, $b = -.45$, 95% CI [−.72, −.17], exp(b) = .64, 95% CI [.49, .85], $p < .001$; (f) belief that plaintiff attorneys are liars: $b = -.17$, 95% CI [−.26, −.09], exp(b) = .84, 95% CI [.77, .92], $p < .001$; (g) Skepticism toward Litigants Factor: $b = .40$, 95% CI [.28, .52], exp(b) = 1.49, 95% CI [1.32, 1.68], $p < .001$; and (h) Negative Attitudes Toward Lawsuits Scale, $b = -.20$, 95% CI [−.32, −.08], exp(b) = .82, 95% CI [.73, .92], $p = .001$.

Thus, we found support for Hypothesis 3 regarding the relative effectiveness of minimal versus extended voir dire in explaining case judgments. The inclusion of the extended voir dire questions proved to be valuable in providing significantly more accurate predictions of mock jurors' verdicts and damage awards than the combination of demographics and minimal voir dire responses alone. Further, although related, there were many extended voir dire questions that explained unique variance in mock jurors' decisions—they were not redundant with each other.

Hypothesis 4: "Excludable" Versus "Nonexcludable" Jurors

We next tested our hypothesis that extended (vs. minimal) voir dire questions would be better able to identify mock jurors who hold extreme attitudes that could threaten their impartiality or willingness to follow the law (i.e., mock jurors who gave a response that could be the basis for a judge or attorney striking them from the jury)—and that these "excludable" jurors would render different judgments. In contrast to the very small percentage of mock jurors able to self-identify and willing to admit bias in response to minimal voir dire questions (< 2%), 42% acknowledged relatively extreme views in response to at least one of the extended voir dire items. Supporting Hypothesis 4, these potentially biased mock

Figure 2
The Relationships Between Extended Voir Dire Responses and Damage Awards



Note. For trust in plaintiffs and the belief that defense attorneys are liars, higher numbers represent more proplaintiff/antidefense views. For the Political Attitudes Factor, higher scores reflect greater conservatism. For all other predictors, higher numbers reflect more antiplaintiff/prodefense views.

jurors would not have been identifiable to attorneys without extended voir dire. Next, we tested whether these "excludable" mock jurors offered systematically different verdicts and damage awards than the rest of our sample. We found mixed support for this hypothesis.

Verdicts

Supporting Hypothesis 4, we found that excludable mock jurors were significantly less likely to judge the defendant liable than other mock jurors, $B = -.60$, $OR = .55$, 95% CI [.45, .67], $p < .001$. The odds ratio reveals that excludable mock jurors were roughly half as likely to judge the defendant liable (51%) than the nonexcludable mock jurors (68%).

Damage Awards

Not supporting Hypothesis 4, however, we found that excludable mock jurors did not awards significantly lower damage awards than the other mock jurors, $b = -.06$, 95% CI [−.13, .002], exp(b) = .94, 95% CI [.88, 1.00] $p = .06$. On average, excludable mock jurors awarded approximately $5,780,496 to the plaintiff, while the rest of the mock jurors who read the same case awarded approximately $5,443,866—a difference of $336,630.

Exploratory Analysis: Bias Awareness

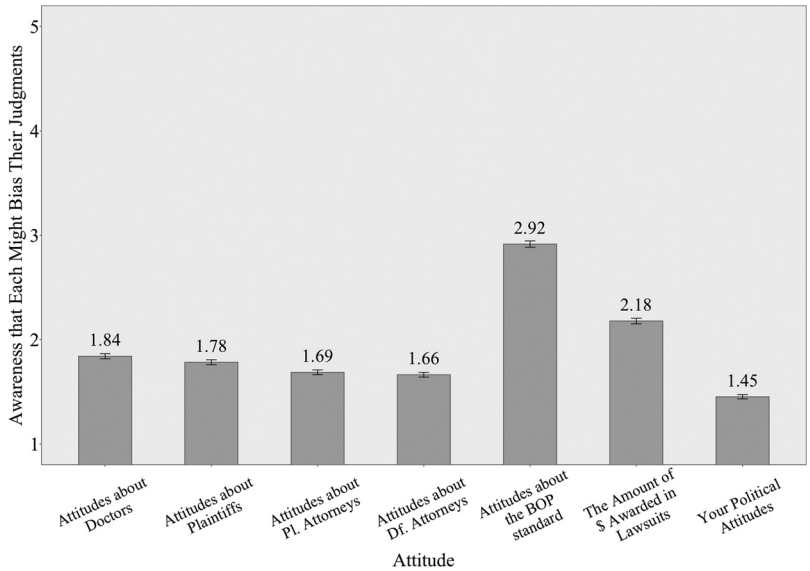
Having established that undergoing judicial rehabilitation did little to weaken the link between potential sources of bias and mock jurors' case judgments, we tested whether it nevertheless changed mock jurors' ability or willingness to acknowledge such biases. Figure 3 provides the answer to this question. Mock jurors thought that most factors had somewhere between (a) *no impact at all* and (b) and *a tiny impact* on their decision making. Only the item asking about their agreement with the burden of proof threshold provided an average level of acknowledgment approaching (c) *a moderate impact*.

An additional and provocative finding regarding awareness of bias is worth noting. To test whether the manipulations in our study affected self-awareness and acknowledgment of bias, we tested their impact on the Bias Awareness Scale that averaged ratings across potential bias sources. The presence versus absence of judicial rehabilitation was a significant predictor of bias awareness, $\chi^2(1, 2{,}033) = 8.44$, $p = .004$ (no other effects were significant, χ^2s ≤ 7.92, all ps $\geq .02$). Despite judicial rehabilitation not *actually* reducing the impact of mock jurors' preexisting views on their judgments, mock jurors who heard and agreed to judicial rehabilitation nevertheless *reported* being significantly less impacted by these biasing factors than did mock jurors who did not undergo judicial rehabilitation, $b = -.10$, 95% CI [−.17, −.03], $z = -2.91$, $p = .004$. In other words, judicial rehabilitation had a perverse and perhaps undesirable effect. It gave mock jurors the illusion that they were putting their biases aside when the impact of those biases remained undiminished.

Discussion

The voir dire process allows attorneys to question jurors before trial in order to uncover biases and extreme attitudes that they think would be likely to influence jurors' verdicts and damage

Figure 3
Bias Awareness as a Function of Each Potential Bias Source



Note. The bias awareness responses were measured on 5-point scales ranging from 1 = *not at all* to 5 = *an extremely big impact.*

awards. The findings from the present research suggest that generic questions requiring jurors to spontaneously and explicitly acknowledge that they cannot be impartial are unlikely to aid attorneys or presiding judges in that task. The opportunity to probe mock jurors' attitudes and biases in a more extended voir dire that asked about specific attitudes regarding civil litigation and the parties involved was necessary to accomplish that task in the realm of two medical malpractice cases and a bad faith case. Mock jurors were able and willing to express extreme attitudes—but they had to be asked about them directly.

Our present findings, furthermore, refute a key assumption often cited in support of limiting voir dire: Including judicial rehabilitation did not reduce the biasing impact of mock jurors' preexisting attitudes on their decisions. In fact, judicial rehabilitation had an ironic backfire effect: Despite not *actually* reducing mock jurors' biases, it did make them *think* they were less biased.

Hypothesis 1: The Effectiveness of Minimal Voir Dire in Predicting Case Judgments

The minimal voir dire generic questions invited mock jurors to identify and acknowledge potential biases that might impact their decisions, but without offering or specifying examples of such biases. No demographic or minimal voir dire items proved useful in identifying mock jurors likely to be biased in their judgments. Although these findings may be disconcerting to many attorneys and trial consultants who regularly rely on these factors to choose juries, these findings are consistent with those of many other studies and reviews of the literature (e.g., Hans & Jehle, 2003; Kovera & Austin, 2016; Lieberman, 2011).

It is also noteworthy that very few mock jurors (< 2%) volunteered information that they harbored potentially biasing views in response to minimal voir dire questions that asked them to come up with potential sources of bias themselves. Expecting naïve jurors, most of whom have never been in a courtroom, to spontaneously realize and acknowledge biases arising from, for example, negative attitudes toward litigation or the awarding of noneconomic damages is not reasonable. Moreover, the small minority who responded to minimal voir dire questions asking them if they could think of anything that might prevent them from being impartial (< 2%) made similar judgments to those who did not. The combination of these questions *and* demographic variables explained less than 1% of variance in verdicts and only 2% of variance in damages. Limiting voir dire to these questions thus compromises the constitutional right of plaintiffs and defendants alike to an impartial jury.

Hypothesis 1: The Effectiveness of Extended Voir Dire in Predicting Case Judgments

In contrast, our findings suggest that extended voir dire questions that go beyond asking jurors to identify their own biases and, instead, ask about preexisting views about civil litigation might be more helpful in identifying jurors who should be excluded for cause. Almost every preexisting view assessed in our extended voir dire questioning, including those about plaintiffs, attorneys, doctors, insurance companies, burden of proof, and limitations on litigation, significantly influenced mock jurors' verdicts, and the majority of them influenced damages.

The magnitude of the effect sizes revealed in our studies merit serious consideration. In the case of several of our measures (e.g., thinking the plaintiff would have a harder time convincing them than the defense, unwillingness to awards noneconomic damages, and displaying skepticism toward litigants), just a one-unit increase on these measures was associated with a drop in the likelihood of a verdict favoring the plaintiff that ranged from half as

likely to vote liable to two-thirds as likely to vote liable. The extended voir dire questions' predictive power regarding damage awards were similarly impressive. They ranged from over a $19,000 decrease in damage awards for a one-unit increase in thinking plaintiff attorneys are liars, to over a $140,000 decrease in damage awards as participants move from willing to unwilling to awards noneconomic damages.

Hypothesis 2: Effectiveness of Judicial Rehabilitation

Basic social psychological research (Devine et al., 2012; Gawronski et al., 2006) suggests that questioning jurors about their biases and asking them if they can remain impartial might enhance their awareness of those biases and motivate them to control them —thereby reducing the impact of their predispositions on their ultimate case judgments. Our findings indicate otherwise. Across three different cases, two decisions (verdict, damages), and 13 different preexisting attitudes, we found that these attitudes influenced verdicts and damage awards, regardless of whether mock jurors underwent judicial rehabilitation.

As expected, very few mock jurors were able or willing to admit that they would not be able to set aside their biases and be impartial when asked by a judge via video (Hypothesis 2a). In fact, even fewer jurors in our study (n = 11, 0.4%) did not think they could be impartial than actual potential jurors in a previous study of real criminal trials (8%, Schuller et al., 2015). The fact that mock jurors said they could be fair at such stunningly high rates in our studies is even more surprising because there was little of the typical pressure real jurors might feel in a real courtroom with a real judge—there were no real consequences to refusing in these studies. Consistent with bias blind spots (Pronin et al., 2002), it is possible that most potential jurors might really think that they can put their biases aside and be impartial—rendering this question during judicial rehabilitation not useful in identifying jurors who cannot be impartial.

Further, it is clear that experiencing the version of judicial rehabilitation employed in our three experiments and agreeing to be impartial did not reduce the strength of the relationship between any of the jurors' preexisting attitudes and biases on their verdicts or damage awards (Hypothesis 2b). Although we did our best to create realism by recording a video of a judge in a courtroom delivering a common instruction about controlling bias and also making jurors commit to being impartial, we acknowledge that our procedures could not obviously fully capture the experience of being in front of a real judge in a real courtroom. However, our null results do converge with those of a prior experiment (Crocker & Kovera, 2010), in which mock jurors experienced actual questioning in a courtroom by someone they believed to be a judge. This research suggests the legal system needs a more effective way of rehabilitating jurors.

The judicial rehabilitation manipulation *did* affect how biased mock jurors *thought* that they had been. Mock jurors in all conditions in our study claimed that all potential sources of bias had *no impact at all* or *a tiny impact*, on average. Ironically, however, mock jurors who had gone through judicial rehabilitation and agreed to be impartial *claimed* that they were significantly less influenced by these potential sources of biases than mock jurors who had not experienced such judicial rehabilitation. Consistent with social psychological research on "credentialing" (Effron et al., 2009) and "backfire effects" of judicial limiting instructions (Lieberman & Arndt, 2000), the rehabilitation procedure seems to have given mock jurors a false sense of security that their judgments would no longer be biased by their preexisting attitudes. Judicial rehabilitation might have increased mock jurors' "bias blindspot" (Pronin et al., 2002).

Hypothesis 3: Unique Variance in Case Judgments Explained by Extended Voir Dire Responses

In contrast to the combination of demographic variables and limited voir dire responses explaining less than 1% of variance in verdicts and only 2% of variance in damages, extended voir dire responses explained substantial unique variance in case judgments. Extended voir dire questions increased the percentage of variance explained in verdicts to 31% and in damage awards to 9%, across cases. Further, there were many unique predictors of verdicts and damage awards—meaning many extended voir dire responses explained unique variance above and beyond the other questions. This supports the utility of extending how many questions attorneys can ask—many of the questions in our study were not redundant in their explanatory value.

The verdict findings are consistent with Diamond et al.'s (1998) jury simulation experiment; both their study and ours found a number of voir dire questions about attitudes toward litigation that predicted verdicts—both finding that attitudes toward lawsuits, for example, predicted verdicts. Similar to our study, they found that significantly more variance in verdicts was explained by specific attitudes than by demographics and minimal voir dire—our study found even higher percentages of variance was explained by extended voir dire responses (31% in our studies vs. 11.5% in Diamond et al., 1998). Further, extended voir dire questions also predicted damage awards in our study, while those in the Diamond et al. (1998) study did not. The predictive value of specific extended voir dire questions on civil case judgments that our study found also converges with results in Greene et al. (1991) and Hans (2000a).

Hypothesis 4: The Effectiveness of Extended Voir Dire in Identifying "Excludable" Jurors

It is further worth noting how prevalent many of the biases revealed in the extended voir dire questionnaire were in our sample of mock jurors. Almost half of the sample (n = 806, 42%) could likely be struck from the jury for admitting in extended voir dire questioning to not being impartial and able to follow the law —despite the fact that fewer than 2% spontaneously acknowledged such biases during minimal voir dire. The good news from these results is that mock jurors *were* willing and able to admit potentially biasing predispositions when asked—they just might not be aware enough to self-identify them on their own.

An obvious implication of the findings we have reviewed pertains to attorneys' potential use of challenges for cause and/or peremptory challenges. Overall, mock jurors whose responses placed them in the potentially "excludable" category were only about half as likely to find the defendant liable as the jurors who did not admit to extreme attitudes. In contrast to previous studies casting doubt on the ability of attorneys to identify potentially biased jurors during voir dire (e.g., Otis et al., 2014), our findings suggest that extended voir dire questioning could enable attorneys to

identify jurors with strong predispositions that might bias them and make better use of their challenges in civil cases. Similarly, courts could more effectively exclude jurors who may struggle to follow the law, such as jurors who might resist the existing burden of proof or who believe noneconomic damages are inappropriate.

The Effect of Exposure to Extended Voir Dire

Critics of extended voir dire have raised the problematic possibility that attorneys might sway or "taint" jurors through the questions they ask before the trial even starts (e.g., Mize et al., 2007). It is noteworthy that—contrary to the claims and fears of some judges and attorneys—experiencing voir dire before judging the case did not appear to have a "tainting" effect on mock jurors. It did not systematically tilt the playing field to advantage plaintiffs or defendants. Similar tests of exposure to voir dire on criminal verdicts after reviewing trial evidence provided the same reassurance (Greathouse et al., 2011; Vitriol & Kovera, 2018). Nevertheless we cannot rule out the possibility that voir dire questioning conducted by an authority figure, like an attorney or judge in court (Jones, 1987) could have such a tainting effect. Further, the extended voir dire responses were not less predictive when they were asked before the trial relative to after the trial—providing no support for the idea that calling awareness to potential biases during voir dire might reduce their impact.

Limitations and Future Directions

Despite our efforts to make our mock jurors' task as realistic as possible by providing a rich set of case facts, commonly used voir dire questions, and a video of a judge delivering rehabilitation, we obviously could not create a true juror experience. It is possible, for example, that the judicial rehabilitation manipulation would have been more impactful if it had been more interactive, or perhaps more mock jurors would have said that they could not be impartial if the judge's question had not been made so salient by being the only videotaped portion of the case materials. It is also possible that (as is a common concern with online samples) that mock jurors might have reported that they could set aside their biases so as to "keep the studying moving" and finish more quickly—though this motivation might also exist for real jurors during voir dire. These are important limitations. However, we are reassured that our judicial rehabilitation results are not simply an artifact of our methodology given that they converge with information from posttrial interviews with real jurors who report that they were reticent during voir dire to admit biases or to say they could not be impartial (Johnson & Haney, 1994; Schuller et al., 2015). Further, our results conceptually replicate the other manipulation of judicial rehabilitation on verdicts that was an in-person interaction (Crocker & Kovera, 2010).

There are limits to the case materials we presented. We employed two very different medical malpractice cases and an insurance bad faith case, but of course, a vast array of other types of civil lawsuits in torts, property, and contracts cases exist that merit exploration in future research. We could not duplicate the setting and witnessing of the live events of a real trial, with more extensive presentation of facts, expert testimony, and arguments offered by attorneys. A review of jury simulation literature revealed relatively minimal impact of less versus more ecologically valid designs (Bornstein et al., 2017), but we recognize and appreciate the differences between our experimental approach and civil trials in the real world.

In particular, mock jurors did not deliberate. Both civil jury simulations that include deliberation and interviews with real jurors reveal how deliberation can lead people to shift from their initial damage awards suggestion (Diamond et al., 1998; Hans & Lofquist, 1992). Research suggests that deliberation can diminish or exacerbate the impact of biases (Salerno et al., 2013) and pretrial publicity (Lieberman & Arndt, 2000). It thus would be valuable to test whether our findings would replicate if the mock jurors were asked the same questions in a courtroom by an attorney or judge and then allowed to deliberate before making case judgments.

Notwithstanding the limitations of our studies, our findings are important in that they suggest the potential value of voir dire questionnaires that probe specific predispositions and biases directly relevant to civil litigation regarding medical malpractice and bad faith cases. Such questionnaires provide privacy, reduce potential embarrassment (Hans & Jehle, 2003), and decrease the demand and temptation to give the court the answers it hopes to hear. Although having potential jurors complete an extensive voir dire questionnaire takes some time, if the questionnaires were administered in advance of the trial, or just prior to voir dire questioning, it might actually save time. Thus, these questionnaires could eliminate the need for a tradeoff between efficiency and effectiveness in the conduct of voir dire.

Finally, the goal of these studies was to demonstrate the benefits of more extended voir dire relative to minimal voir dire, rather than to develop any specific measure to recommend to attorneys for use during voir dire. That being said, some of our measures could be improved psychometrically in future research. For example, several of our scales had Cronbach's alpha levels in the .61–.69 range, which reached "acceptable" (Murphy & Davidshofer, 1988), but not "good" (Nunnally, 1978) reliability (Peterson, 1994). One of our measures of political attitudes is two dichotomous items, which could be better measured on a continuous scale, and there are surely other measures that could better capture political conservatism.

Along similar lines, one of our goals was to show the relative utility of minimal versus extended voir dire for identifying "excludable" jurors, who attorneys could argue should be stricken for cause. However, our coding of who could be "excludable" was based on one of many potential operationalizations of what responses could justify exclusion. We tried to be conservative by coding mock jurors as potentially excludable only when they expressed clearly extreme predispositions or an unwillingness to follow the law with respect to specific issues such as opposition to awarding noneconomic damages. Even using that relatively conservative approach, fully 42% of the sample expressed views that would potentially merit exclusion. Whether judges would remove such potential jurors for cause, or whether lawyers would strike them peremptorily, are separate issues that we cannot address with our methodology. However, it is important to note that if we had used a more liberal cutoff and included participants who had less extreme but still potentially troublesome attitudes, the percentage of mock jurors who could be counted as excludable would be even higher. In short, we think it is clear that specific extended voir dire questions are much more helpful in identifying excludable jurors than minimal self-identification of bias questions—but the degree

of that difference depends on the specific questions asked, the cutoffs drawn, and the behavior of judges and lawyers.

Conclusion

Allowing extensive voir dire is time consuming. Some courts believe it is unnecessary, and that most jurors can set aside their biases and be fair—especially after judicial rehabilitation. Accordingly, they favor a very limited voir dire, one that does not probe potential biases to the degree that our research suggests would be necessary to detect and excuse jurors whose biases should preclude participation in a particular trial. Our data suggest that extended voir dire is worthwhile. Almost half of the mock jurors in our studies (42%) acknowledged views and biases during extensive voir dire that, if revealed and discovered, would potentially have resulted in their exclusion from the jury. By contrast, less than 2% of the mock jurors in our studies would have been identified through standard, minimal voir dire questioning that requires jurors to identify and report such biases without the court further stipulating the specific biases that might be problematic. Most of the questions asked in extended voir dire predicted verdicts and damage awards. Neither experiencing voir dire or judicial rehabilitation before the trial reduced this biasing impact of these preexisting attitudes on their case judgments.

Attorneys need the opportunity to pose the kinds of questions to jurors necessary to guarantee their clients a fair trial. More extensive voir dire could provide that opportunity, and if conducted via questionnaire it could be provided without unduly taxing the court's time and patience.

References

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Prentice Hall.

American Bar Association. (2005). Principles for juries and jury trials. https://www.uscourts.gov/sites/default/files/aba_principles_for_juries_and_jury_trials_2005.pdf

Axt, J. R., Nguyen, H., & Nosek, B. A. (2018). The Judgment Bias Task: A flexible method for assessing individual differences in social judgment biases. *Journal of Experimental Social Psychology*, *76*, 337–355. https://doi.org/10.1016/j.jesp.2018.02.011

Barber, J., & Thompson, S. (2004). Multiple regression of cost data: use of generalised linear models. *Journal of Health Services Research & Policy*, *9*(4), 197–204. https://doi.org/10.1258/1355819042250249

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., Cesarini, D., Chambers, C. D., Clyde, M., Cook, T. D., De Boeck, P., Dienes, Z., Dreber, A., Easwaran, K., Efferson, C., . . . Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*(1), 6–10. https://doi.org/10.1038/s41562-017-0189-z

Bennett, M. W. (2010). Unraveling the Gordian knot of implicit bias in jury selection: The problems of judge-dominated voir dire, the failed promise of Batson, and proposed solutions. *Harvard Law & Policy Review*, *4*(1), 149–171.

Bermant, G. (1982). *Jury selection in United States District Courts* (Education and Training Series). Federal Judicial Center. https://www.ojp.gov/pdffiles1/Digitization/86809NCJRS.pdf

Bornstein, B. H., Golding, J. M., Neuschatz, J., Kimbrough, C., Reed, K., Magyarics, C., & Luecht, K. (2017). Mock juror sampling issues in jury simulation research: A meta-analysis. *Law and Human Behavior*, *41*(1), 13–28. https://doi.org/10.1037/lhb0000223

Broeder, D. W. (1959). The University of Chicago jury project. *Nebraska Law Review*, *38*(3), 744–760. https://digitalcommons.unl.edu/nlr/vol38/iss3/9/

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*(1), 3–5. https://doi.org/10.1177/1745691610393980

Colo. R. Civ. P. 47(a)(3).

Coppock, A. (2019). Generalizing from survey experiments conducted on Mechanical Turk: A replication approach. *Political Science Research and Methods*, *7*(3), 613–628. https://doi.org/10.1017/psrm.2018.10

Crocker, C. B., & Kovera, M. B. (2010). The effects of rehabilitative voir dire on juror bias and decision making. *Law and Human Behavior*, *34*(3), 212–226. https://doi.org/10.1007/s10979-009-9193-9

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, *48*(6), 1267–1278. https://doi.org/10.1016/j.jesp.2012.06.003

Diamond, S. S., Saks, M. J., & Landsman, S. (1998). Juror judgments about liability and damages: Sources of variability and ways to increase consistency. *De Paul Law Review*, *48*(2), 301–325.

Effron, D. A., Cameron, J. S., & Monin, B. (2009). Endorsing Obama licenses favoring Whites. *Journal of Experimental Social Psychology*, *45*(3), 590–593. https://doi.org/10.1016/j.jesp.2009.02.001

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*, *117*(3), 522–559. https://doi.org/10.1037/pspa0000160

Frederick, J. T. (2005). *Mastering voir dire and jury selection* (2nd ed.). American Bar Association.

Frogel v. Philip Morris USA, Inc., 305 So. 3d 793, 798 (Fla. Dist. Ct. App. 2020).

Fry, J. M., Fry, T. R., & McLaren, K. R. (2000). Compositional data analysis and zeros in micro data. *Applied Economics*, *32*(8), 953–959. https://doi.org/10.1080/000368400322002

Fulero, S. M., & Penrod, S. D. (1990). The myths and realities of attorney jury selection folklore and scientific jury selection: What works. *Ohio Northern University Law Review*, *17*, 229–254.

Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61–89). Academic Press.

Gawronski, B., Hofmann, W., & Wilbur, C. J. (2006). Are "implicit" attitudes unconscious? *Consciousness and Cognition*, *15*(3), 485–499. https://doi.org/10.1016/j.concog.2005.11.007

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making*, *26*(3), 213–224. https://doi.org/10.1002/bdm.1753

Greathouse, S. M., Sothmann, F. C., Levett, L. M., & Kovera, M. B. (2011). The potentially biasing effects of voir dire in juvenile waiver cases. *Law and Human Behavior*, *35*(6), 427–439. https://doi.org/10.1007/s10979-010-9247-z

Greene, E., Goodman, J., & Loftus, E. F. (1991). Jurors' attitudes about civil litigation and the size of damage awards. *The American University Law Review*, *40*(2), 805–820.

Greene, E., Johns, M., & Smith, A. (2001). The effects of defendant conduct on jury damage awards. *Journal of Applied Psychology*, *86*(2), 228–237. https://doi.org/10.1037/0021-9010.86.2.228

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*(1), 4–27. https://doi.org/10.1037/0033-295x.102.1.4

Greenwald, A. G., Smith, C. T., Sriram, N., Bar-Anan, Y., & Nosek, B. A. (2009). Implicit race attitudes predicted vote in the 2008 U.S.

presidential election. *Analyses of Social Issues and Public Policy*, *9*(1), 241–253. https://doi.org/10.1111/j.1530-2415.2009.01195.x

Hannaford-Agor, P. L., & Waters, N. L. (2004). *Examining voir dire in California*. The Judicial Council of California, Administrative Office of the Courts. https://www.courts.ca.gov/documents/voir_dire_report.pdf

Hans, V. P. (2000a). *Business on trial: The civil jury and corporate responsibility*. Yale University Press.

Hans, V. P. (2000b). Juries. In A. E. Kazdin (Ed.), *Encyclopedia of psychology* (Vol. 4, pp. 419–421). American Psychological Association.

Hans, V. P. (2008). Faking it? Citizen perceptions of whiplash injuries. In B. H. Bornstein, R. L. Wiener, R. Schopp, & S. Willborn (Eds.), *Civil juries and civil justice: Psychological and legal perspectives* (pp. 131–150). Springer. https://doi.org/10.1007/978-0-387-74490-2_8

Hans, V. P., & Jehle, A. (2003). Avoid bald men and people with green socks? Other ways to improve the voir dire process in jury selection. *Chicago-Kent Law Review*, *78*(3), 1179–1201.

Hans, V. P., & Lofquist, W. S. (1992). Jurors' judgments of business liability in tort cases: Implications for the litigation explosion debate. *Law & Society Review*, *26*(1), 85–115. https://doi.org/10.2307/3053837

Hans, V. P., & Lofquist, W. S. (1994). Perceptions of civil justice: The litigation crisis attitudes of civil jurors. *Behavioral Sciences & the Law*, *12*(2), 181–196. https://doi.org/10.1002/bsl.2370120207

Hans, V. P., & Vidmar, N. (2007). The verdict on juries. *Judicature*, *91*(5), 226–230.

Hans, V. P., & Vadino, N. (2007). *After the crash: Citizens' perceptions of connective tissue injury lawsuits*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1012392

In re: Bard IVC Filters Products Liability Litigation, No. MDL 15-02641-PHX-DGC (D. Ariz. March 12, 2018), ECF No. 10402 (Juror Questionnaire).

Irvine, K., Hoffman, D. A., & Wilkinson-Ryan, T. (2018). Law and psychology grows up, goes online, and replicates. *Journal of Empirical Legal Studies*, *15*(2), 320–355. https://doi.org/10.1111/jels.12180

Jackson v. State, 213 So. 3d 754, 770 (Fla. 2017).

Johnson, C., & Haney, C. (1994). Felony voir dire. *Law and Human Behavior*, *18*(5), 487–506. https://doi.org/10.1007/BF01499170

Jones, S. E. (1987). Judge- versus attorney-conducted voir dire: An empirical investigation of juror candor. *Law and Human Behavior*, *11*(2), 131–146. https://doi.org/10.1007/BF01040446

Kassin, S. M., & Wrightsman, L. S. (1983). The construction and validation of a juror bias scale. *Journal of Research in Personality*, *17*(4), 423–442. https://doi.org/10.1016/0092-6566(83)90070-3

Kerr, N. L., Kramer, G. P., Carroll, J. S., & Alfini, J. J. (1991). On the effectiveness of voir dire in criminal cases with prejudicial pretrial publicity: An empirical study. *The American University Law Review*, *40*(2), 665–701.

Kovera, M. B., & Austin, J. L. (2016). Identifying juror bias: Moving from assessment and prediction to a new generation of jury selection research. In C. Willis-Esqueda & B. H. Bornstein (Eds.), *The witness stand and Lawrence S. Wrightsman, Jr.* (pp. 75–94). Springer Science. https://doi.org/10.1007/978-1-4939-2077-8_6

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., Marshburn, C. K., Simon, S., Blanchar, J. C., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhaas, F. M. H., Burns, M., . . . Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, *145*(8), 1001–1016. https://doi.org/10.1037/xge0000179

Lazarus, E. (1999). *Public opinion: Soft tissue injury lawsuits* [Report]. Lazarus Strategic Services. https://www.lazaruslaw.us/2020/08/soft-tissue-neck-and-shoulder-injuries-due-to-rear-end-collisions-lazarus-law-firm.html

Lecci, L., & Myers, B. (2008). Individual differences in attitudes relevant to juror decision making: Development and validation of the pretrial juror attitude questionnaire (PJAQ). *Journal of Applied Social Psychology*, *38*(8), 2010–2038. https://doi.org/10.1111/j.1559-1816.2008.00378.x

Lieberman, J. D. (2011). The utility of scientific jury selection: Still murky after 30 years. *Current Directions in Psychological Science*, *20*(1), 48–52. https://doi.org/10.1177/0963721410396628

Lieberman, J. D., & Arndt, J. (2000). Understanding the limits of limiting instructions: Social psychological explanations for the failures of instructions to disregard pretrial publicity and other inadmissible evidence. *Psychology, Public Policy, and Law*, *6*(3), 677–711. https://doi.org/10.1037/1076-8971.6.3.677

Luginbuhl, J., & Middendorf, K. (1988). Death penalty beliefs and jurors' responses to aggravating and mitigating circumstances in capital trials. *Law and Human Behavior*, *12*(3), 263–281. https://doi.org/10.1007/BF01044384

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology*, *67*(5), 808–817. https://doi.org/10.1037/0022-3514.67.5.808

Marshall, L. L., & Smith, A. (2012). The effects of demand characteristics, evaluation anxiety, and expectancy on juror honesty during voir dire. *The Journal of Psychology*, *120*(3), 205–217. https://doi.org/10.1080/00223980.1986.10545246

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review*, *28*(2), 295–314. https://doi.org/10.1007/s10648-014-9287-x

Mize, G. E., Hannaford-Agor, P., & Waters, N. L. (2007). *The state-of-the-states survey of jury improvement efforts: A compendium report*. National Center for State Courts. https://www.ncsc-jurystudies.org/__data/assets/pdf_file/0016/5623/soscompendiumfinal.pdf

Mollway, S. O. (n.d.). *Sample voir dire subjects covered by Judge Susan Oki Mollway in civil trials*. Retrieved from https://www.hid.uscourts.gov/reqrmts/SOM/SOM_standard_civil_voir_dire.pdf?pid=19&mid=63

Montgomery v. Commonwealth of Kentucky, 819 S.W.2d 713 (Ky. 1991).

Moore, R. (2019). *Practice standards (civil)*. United States District Court, District of Colorado. http://www.cod.uscourts.gov/Portals/0/Documents/Judges/RM/RM_Civil_Practice_Standards_Aug19.pdf

Moran, G., Cutler, B. L., & De Lisa, A. (1994). Attitudes toward tort reform, scientific jury selection, and jury bias: Verdict inclination in criminal and civil trials. *Law and Psychology Review*, *18*, 309–328.

Morrison, M., DeVaul-Fetters, A., & Gawronski, B. (2016). Stacking the jury: Legal professionals' peremptory challenges reflect jurors' levels of implicit race bias. *Personality and Social Psychology Bulletin*, *42*(8), 1129–1141. https://doi.org/10.1177/0146167216651853

Murphy, K. R., & Davidshofer, C. O. (1988). *Psychological testing: Principles and applications*. Prentice Hall.

Myers, B., & Lecci, L. (1998). Revising the factor structure of the Juror Bias Scale: A method for the empirical validation of theoretical constructs. *Law and Human Behavior*, *22*(2), 239–256. https://doi.org/10.1023/A:1025798204956

National Jury Project. (2000). *Jurywork: Systematic techniques (Release 21)*. West Group.

New York v. Biondo, 41 N.Y.2d 483 (1977).

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259. https://doi.org/10.1037/0033-295X.84.3.231

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

Otis, C. C., Greathouse, S. M., Kennard, J. B., & Kovera, M. B. (2014). Hypothesis testing in attorney-conducted voir dire. *Law and Human Behavior*, *38*(4), 392–404. https://doi.org/10.1037/lhb0000092

Payne, B. K., Lambert, A. J., & Jacoby, L. L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology*, *38*(4), 384–396. https://doi.org/10.1016/S0022-1031(02)00006-9

People v. Griffin, 985 P.2d 15 (Colo. App. 1998).

Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *The Journal of Consumer Research*, *21*(2), 381–391. https://doi.org/10.1086/209405

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, *28*(3), 369–381. https://doi.org/10.1177/0146167202286008

Read, D. S. (2007). *Winning at trial*. National Institute for Trial Advocacy.

Rosner, B. (1983). Percentage points for a generalized ESD many-outlier procedure. *Technometrics*, *25*(2), 165–172. https://doi.org/10.1080/00401706.1983.10487848

Salerno, J. M., Haegerich, T. M., & Bottoms, B. L. (2013). Are the effects of juvenile offender stereotypes maximized or minimized by jury deliberation? *Psychology, Public Policy, and Law*, *19*(1), 81–97. https://doi.org/10.1037/a0027808

Sanderson v. Advanced Eurology P.C., Case No: 06CA2089, 11-13 (Colo. App. Ct. 2008) (unpublished decision).

Schuller, R. A., Erentzen, C., Vo, A., & Li, D. (2015). Challenge for cause: Bias screening procedures and their application in a Canadian courtroom. *Psychology, Public Policy, and Law*, *21*(4), 407–419. https://doi.org/10.1037/law0000056

Schuller, R. A., Kazoleas, V., & Kawakami, K. (2009). The impact of prejudice screening procedures on racial bias in the courtroom. *Law and Human Behavior*, *33*(4), 320–328. https://doi.org/10.1007/s10979-008-9153-9

Seltzer, R., Venuti, M. A., & Lopes, G. M. (1991). Juror honesty during the voir dire. *Journal of Criminal Justice*, *19*(5), 451–462. https://doi.org/10.1016/0047-2352(91)90019-R

Shinal v. Toms, 640 Pa. 295, 319, 162 A.3d 429, 443 (2017).

Skeem, J. L., Louden, J. E., & Evans, J. (2004). Venirepersons's attitudes toward the insanity defense: Developing, refining, and validating a scale. *Law and Human Behavior*, *28*(6), 623–648. https://doi.org/10.1007/s10979-004-0487-7

Sommers, S. R. (2006). On racial diversity and group decision making: Identifying multiple effects of racial composition on jury deliberations. *Journal of Personality and Social Psychology*, *90*(4), 597–612. https://doi.org/10.1037/0022-3514.90.4.597

Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., & Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(19), 7710–7715. https://doi.org/10.1073/pnas.1014345108

State v. Kreutzer, 928 S.W.2d 854, 866 (Mo. 1996).

Steblay, N., Hosch, H. M., Culhane, S. E., & McWethy, A. (2006). The impact on juror verdicts of judicial instruction to disregard inadmissible evidence: A meta-analysis. *Law and Human Behavior*, *30*(4), 469–492. https://doi.org/10.1007/s10979-006-9039-7

Thomas ex rel. Thomas v. Mercy Hospitals East Communities, 525 S.W.3d 114 (Mo. 2017).

Thompson, W. C., Cowan, C. L., Ellsworth, P. C., & Harrington, J. C. (1984). Death penalty attitudes and conviction proneness: The translation of attitudes into verdicts. *Law and Human Behavior*, *8*(1–2), 95–113. https://doi.org/10.1007/BF01044353

Tukey, J. W. (1962). The future of data analysis. *Annals of Mathematical Statistics*, *33*(1), 1–67. https://doi.org/10.1214/aoms/1177704711

Utah v. Saunders, 992 P. 2d 951 (Utah 1999).

Vitriol, J. A., & Kovera, M. B. (2018). Exposure to capital voir dire may not increase convictions despite increasing pretrial presumption of guilt. *Law and Human Behavior*, *42*(5), 472–483. https://doi.org/10.1037/lhb0000304

Walls v. Kim, 250 Ga. App. 259, 260–61 (2001); *aff'd*, 275 Ga. 177 (2001).

Wegner, D. M., & Schneider, D. J. (2003). The white bear story. Psychological Inquiry, *14*(3–4), 326–329. https://doi.org/10.1207/S15327965PLI1403

Received October 9, 2020
Revision received June 26, 2021
Accepted June 28, 2021 ■