CAN JURORS SELF-DIAGNOSE BIAS?
TWO RANDOMIZED CONTROLLED TRIALS

Christopher Robertson, JD, PhD, David Yokum, MA, and Matt Palmer*

Accepted Via Peer Review for the 7th Annual Conference on Empirical Legal Studies
Arizona Legal Studies Discussion Paper No. 12-35

Draft as of Feb 14, 2013

* Robertson is associate professor at James E. Rogers College of Law, University of Arizona, and Research Associate with the Edmond J. Safra Center for Ethics at Harvard University. He may be reached at chris.robertson@law.arizona.edu. Yokum and Palmer are students at the College of Law, and Yokum is also a doctoral candidate in the Department of Psychology at University of Arizona. Disclosure: Robertson has represented defendants who have litigated issues of pretrial publicity and juror bias. No such defendant was involved in the design or funding of this study. We thank the Conference on Empirical Legal Studies, the James E. Rogers College of Law Faculty Enrichment Program, and the Rocky Mountain Junior Scholars Forum for hosting works in progress sessions on this paper, and thank Lawrence Solan, Ed Cheng, Valerie Hans, Sarah Haan, Leslie Francis, Christopher Lasch, Derek Bambauer, Jane Bambauer, Brent White, and Jamie Robertson for helpful comments on drafts.
ABSTRACT

The Constitution guarantees litigants an “impartial” jury, one that bases its judgment on the evidence presented in the courtroom, untainted by affiliations with the parties, racial animus, or media coverage that may include inadmissible facts, a one-sided portrayal, and naked opinion. Problems of juror bias arise in almost every trial – state and federal, civil and criminal – and the problem is most severe in the highest profile cases, where the need for accuracy and legitimacy in outcomes is most salient.

The Supreme Court has instructed courts to use a simple method to determine whether jurors are biased: ask them. Studies have shown that the juror’s self-diagnosis is the most important factor for the court’s decision about whether to seat the juror.

To test the reliability of these self-diagnoses, we fielded two randomized controlled trials, in which we exposed 248 mock jurors to news articles that were either prejudicial to the defendant (in one condition) or irrelevant (in the other condition). We then gave jurors the admonitions and questions endorsed by the Supreme Court for the purpose of identifying biased jurors, prior to all of them watching a 32-minute condensed video of a civil trial, rendering binary judgments, and awarding damages.

We found that jurors were simply unable to diagnose their own biases. Even after we excluded those jurors who said that they would be unable to be fair and impartial (or were unsure), the remaining jurors were significantly more likely (odds ratio 2.4, p = .004) to rule against the defendant and those that did so also awarded eight-times larger damages on the median (p=.01), than those in the control condition.

Thus, juror self-assessments were not related to actual bias. Nor were the self-diagnoses associated with demographic variables or the cognitive scales that we used in regression models. The jurors’ self-diagnoses appear to be random, and thus about as useful to the courts as coin-flipping.

These experiments show that the courts’ current method of policing jury bias is ineffective, since the underlying data lack diagnosticity. These findings call into doubt the validity of prior trial outcomes, and even pleas and settlements that were made in fear of a biased jury.

There is an alternative way to satisfy the Constitutional mandate. Courts should adopt the test that they currently use for judge disqualification. Simply exclude all jurors whose impartiality can be reasonably questioned.
Contents

I. BACKGROUND .............................................................................................................. 4

A. The Courts’ Approach to Jury Bias ........................................................................... 5
   1. The Guarantee of an Impartial Jury ................................................................. 5
   2. The Reliance on Jurors’ Self-Diagnoses ......................................................... 6
   3. The Doctrine on Publicity and Self-Diagnosis ............................................... 7

B. The Science on Jury Bias ......................................................................................... 13
   1. The Biasing Effects of Pretrial Publicity ....................................................... 13
   2. The Efficacy of Voir Dire for Removing that Bias ......................................... 13

II. EXPERIMENTS .......................................................................................................... 16

A. Design and Materials .............................................................................................. 16
   1. Analytical Framework for Hypothesis Tests ................................................... 16
   2. Biasing Stimuli ................................................................................................. 17
   3. Operationalizing the Supreme Court’s Voir Dire Protocol ............................. 18
   4. The Trial Stimulus and Dependent Variables ............................................... 19

B. Experiment 1: Convenience Sample ....................................................................... 20
   1. The Effect of Pretrial Publicity ....................................................................... 20
   2. The Accuracy of Jurors’ Self-Diagnoses ....................................................... 21

C. Experiment 2: National Sample .............................................................................. 24
   1. The Sample Demographics ............................................................................ 24
   2. The Effect of Pretrial Publicity ....................................................................... 25
   3. The Accuracy of Jurors’ Self-Diagnoses ....................................................... 26
   4. Does Juror Self-Diagnosis Help At All? ......................................................... 27
   5. Opinion-Screening as an Alternative to the Skilling Protocol ....................... 28
   6. The Unexplained Behavior of Fair Jurors Disqualifying Themselves .......... 29

D. Overall Analysis ..................................................................................................... 30

E. Limitations .............................................................................................................. 33

III. DISCUSSION ............................................................................................................. 35

A. The Cognitive Limitations of Self-Diagnosis ....................................................... 35

B. A Way Forward: Borrowing the Method Used for Judge Bias ............................. 37

APPENDIX A. REGRESSION ANALYSES .................................................................. 40
CAN JURORS SELF-DIAGNOSE BIAS?

TWO RANDOMIZED CONTROLLED TRIALS

I. BACKGROUND

In late August 2012, a jury in Silicon Valley, California was deliberating on the outcome of a billion dollar patent dispute between two industry giants, Apple and Samsung. The foreman of that jury, Velvin Hogan, was himself an inventor and holder of a patent, which he had litigated. During voir dire, Mr. Hogan had revealed some of these facts in response to the trial judge’s questioning, and Mr. Hogan later told the media that he then “expected to be dismissed from the jury.” Nonetheless, the trial judge also asked Mr. Hogan what trial attorneys know as ‘the magic question:’ “will you be able to decide this case based solely on the evidence that’s admitted during the trial?” Mr. Hogan answered affirmatively, was thus seated in the jury, and later became its foreman. A few weeks later, Mr. Hogan led the jury towards a billion dollar judgment against Samsung, one of the largest in the history of patent law cases.

Likewise in the extensively-publicized criminal prosecution of Gerry Sandusky, a former football coach for Pennsylvania State University, the prosecution moved to change the venue of the trial to avoid the particularly extensive attention the case was receiving in Centre County, where the University is based. The trial court judge denied the motion, siding with the defendant, holding that “the answer to whether a juror can be fair and impartial, despite the myriad of influences to which he or she may be exposed, cannot be known until the juror is actually asked.”

This paper investigates the accuracy and effectiveness of this ubiquitous procedure of asking jurors whether they can be fair and impartial, and then using their responses to decide which jurors should be dismissed for cause. Does this colloquy provide useful information to the litigants and judge, who are together tasked with impaneling an impartial jury? Are the trial outcomes that result from this procedure accurate and legitimate?

---

1 Case5:11-cv-01846-LHK Document2013-9, Filed 10/02/12 (Page 3).
2 Case5:11-cv-01846-LHK Document1991-1, Filed 09/21/12 (Page 22).
3 In a series of media interviews, Mr. Hogan said that he wanted the verdict “to send a message to the industry at large that patent infringing is not the right thing to do” and “make sure the message we sent was not just a slap on the wrist.” Case5:11-cv-01846-LHK Document2013, Filed 10/02/12 (Page14) (quoting multiple exhibits).
A. The Courts’ Approach to Jury Bias

1. The Guarantee of an Impartial Jury

The Sixth Amendment of the United States Constitution provides that “the accused shall enjoy the right to... trial by an impartial jury,” and the guarantee of impartiality applies equally in the civil context.5 “The theory of our [trial] system is that the conclusions to be reached in a case will be induced only by evidence and argument in open court, and not by any outside influence, whether of private talk or public print.”6 The presence of even one biased juror on the jury is a structural error requiring a new trial.7 Still, the U.S. Supreme Court has held that “a litigant is entitled to a fair trial but not a perfect one, for there are no perfect trials.”8 This axiom has been the rule for jury selection in criminal, as well as civil, trials, as “the process of voir dire is designed to help impanel a fair and impartial jury, not a favorable one.”9

Jury bias can arise from all sorts of causes, such as the juror seeing pretrial publicity, the juror having an affiliation with one of the parties or attorneys, the juror having personal experience with the type of crime alleged, or the juror being motivated by racial animus.10 “A court must excuse a juror for cause if the juror is related to one of the parties in the case, or if the juror has even a tiny financial interest in the case.”11

---

5 Several criminal cases are reviewed below. For civil contexts, see e.g., McDonald Power Equipment, Inc. v. Greenwood, 464 U.S. 548 (1984), Sawyer v. Southwest Airlines Co., 145 Fed.Appx. 238, 240-241 (10th Cir. 2005) (“During voir dire, these prospective jurors stated that they could follow the court's instructions and render a fair verdict.”); Moran v. Clarke, 443 F.3d 646, 650-651 (8th Cir. 2006); Kelley v. Wegman's Food Markets, Inc., 98 Fed.Appx. 102 (3d Cir. 2004).
7 Ross v. Oklahoma, 487 US 81, 85 (1980) (“It is well settled that the Sixth and Fourteenth Amendments guarantee a defendant on trial for his life the right to an impartial jury. Had Huling (the biased juror) sat on the jury that ultimately sent petitioner to death, and had petitioner properly preserved his right to challenge the trial court’s failure to remove Huling for cause, the sentence would have to be overturned.”) (citations removed); United States v. Hendrix, 549 F.2d 1225, 1227 (9th Cir. 1977) (“If only one juror is unduly biased or prejudiced or improperly influenced, the criminal defendant is denied his Sixth Amendment right to an impartial panel.”) (citing Tillman v. United States, 406 F.2d 930, 937 (5th Cir.), vacated on other grounds, 395 U.S. 830, 89 S.Ct. 2143, 23 L.Ed.2d 742 (1969) and Stone v. United States, 113 F.2d 70, 77 (6th Cir. 1940)).
9 Id.
10 See generally, PAULA L. HANNAFORD-AGOR, AND NICOLE L. WATERS, ADMINISTRATIVE OFFICE OF THE COURTS, JUDICIAL COUNCIL OF CALIFORNIA, EXAMINING VOIR DIRE IN CALIFORNIA, 27 (2004) (listing eight bases for exclusion observed in California trials). The Supreme Court has held that due process requires inquiry into potential racial bias if “under all the circumstances presented there [is] a constitutionally significant likelihood, absent questioning about racial prejudice, the jurors would be as indifferent as they stand unsworne [sic].” Ristaino v. Ross, 424 U.S. 589, 596 (1976). However, it is unclear exactly when this protection becomes necessary. See e.g., Rosales-Lopez v. U.S., 452 U.S. 182, 189 (1981) (“Only when there are more substantial indications of the likelihood of racial or ethnic prejudice affecting the jurors in a particular case does the trial court’s denial of a defendant’s request to examine the jurors’ ability to deal impartially with this subject amount to an unconstitutional abuse of discretion. Absent such circumstances, the Constitution leaves it to the trial court, and the judicial system within which that court operates, to determine the need for such questions.”) U.S. v. Hosseini, 679 F.3d 544 (7th Cir. 2012). (“Where racial or ethnic bias may be an issue in a case and the defendant requests voir dire on the subject, it is an abuse of discretion to refuse the request.”)
11 U.S. v. Polichemi, 219 F.3d 698, 704 (7th Cir. 2000).
Concerns about juror bias also arise in criminal cases where prosecutors seek the death penalty and seek to remove (for cause) any jurors who express moral reservations about imposing the death penalty. Jurors are instructed that, they should not be swayed by “mere sentiment, conjecture, sympathy, passion, prejudice, public opinion or public feeling.” Even if the juror expresses doubts about her ability to do so, she can and usually will be seated as long as she says that she “will consider and decide the facts impartially and conscientiously apply the law as charged by the court.”

2. The Reliance on Jurors’ Self-Diagnoses

To implement the impartial jury guarantee in both criminal and civil trials, in both federal and state courts, potential jurors are selected through a certain colloquy. The judge or attorneys ask the potential jurors whether they have any feelings or opinions about the litigants, attorneys, facts, or law of the case. As Suggs and Sales explain, “if the juror admits that he has formed an opinion about the case, it is standard procedure to ask if he can set aside that opinion and decide the case on the basis of the evidence to be presented” and the law as instructed.

Courts, and litigants, appear to rely heavily upon the answer to that question. The Ninth Circuit has stated that “[a]ctual bias is found where a prospective juror states that he cannot be impartial, or expresses a view adverse to one party’s position and responds equivocally as to whether he could be fair and impartial despite that view.” Scholars explain that, “it is clear that the juror’s self-assessment about fairness is the strongest factor in judicial decision-making in challenges for cause.” In one recent study of California cases, once a
potential source of bias was identified, if a juror said that he or she could be fair, it made her 71% less likely to be dismissed for cause, all other things being equal. And once the trial judge has made the determination that a juror can be fair, it is virtually unreviewable. Of course, some courts have derided this “magic question,” emphasizing that the juror’s self-professions of fairness should not be determinative, though the trial courts use them nonetheless. “A trial court’s findings of juror impartiality may be overturned only for manifest error. ... [in this domain, the] deference due to district courts is at its pinnacle ...”

3. The Doctrine on Publicity and Self-Diagnosis

Of the various potential sources of jury bias, we focus on bias due to publicity for methodological reasons – i.e., that it can be manipulated in a randomized controlled trial. Pretrial publicity is also, naturally, most problematic in the most highly publicized trials, which means that this particular problem can disproportionately impact perceptions of judicial legitimacy. Furthermore, social media and the 24-hour news cycle may now exacerbate this problem.

Still, the right to challenge jurors for bias has a long history, going back to Chief Justice John Marshall and the trial of Aaron Burr. “The media of the day described the feud between Jefferson and Burr in detail, the citizenry chose sides, and the difficulties in selecting an impartial jury increased.” Marshall analogized the problem of a juror having preconceived notions about the case to the problem of a juror being related to a party.

The relationship may be remote; the person may never have seen the party; he may declare that he feels no prejudice in the case; and yet the law cautiously incapacitates him from serving on the jury because it suspects prejudice, because in general persons in a similar situation would feel attorneys, finding that when a juror expressed that she could not be fair, she was more than eight times likely to be challenged for cause, compared to jurors who expressed sureness or mere doubt about fairness. See Hannaford-Agor and Waters supra note 10 at 37, Table 4.11 (reporting the 71% likelihood figure). More particularly, the authors show in Table 4.3 that in a dataset of 58 such jurors who reported that they were crime victims, the six jurors who said that they could not be fair were all excused, while of the 40 jurors who said they could be fair, only 1 was excused (2.5%). In Table 4.7 (p. 34), for jurors who had views about the case facts, all nine of those who said that they could not be fair were excused, while none of the 10 who said they could be fair or likely could be fair were excused. At p. 31 the authors note that, “at least in the category of previous victimization, this [juror self-assessment of fairness] appears to be a fairly strong basis for the judge’s decision.”

Juror Self-Diagnoses

prejudice. He will listen with more favor to that testimony which confirms, than to that which would change his opinion; it is not to be expected that he will weigh evidence or argument as fairly as a man whose judgment is not made up in the case.25

Notably, Justice Marshall recognized that the pretrial publicity created a “suspicion” of prejudice or “bias”, an empirical claim that must be resolved by the judge.26 Marshall said that, the trial court should question such jurors to decide whether they are “capable of hearing fairly, and of deciding impartially, on the testimony which may be offered to them, or as possessing minds in a situation to struggle against the conviction which that testimony might be calculated to produce.”27 Nonetheless, Marshall himself expressed some doubt about the value of such a colloquy, since there may be prejudice even where the juror “declares that he feels” none. Marshall’s opinion was extremely influential for both the state and federal courts, who adopted the practice of questioning jurors, a practice that did not and still does not exist across the Atlantic.28

The Supreme Court ruled on the pretrial publicity issue in the 1878 case of Reynolds v. United States.29 Two seated jurors admitted to having formed an opinion on the guilt of the accused from newspaper accounts they had read.30 Both jurors, however, stated that they believed they could still be impartial when assessing the facts of the case.31 The trial judge took these jurors at their words, and the Supreme Court found that these jurors were seated properly.32 The Court stated that jurors ought to be disqualified only when their partiality is so strong that it “leave[s] nothing to the ‘conscience or discretion’ of the triers.”33 The Court also put a very high standard for overturning a trial judge who allows a juror with preexisting notions to be seated, stating that to warrant reversal, the trial judge must have committed “manifest error”.34

The Reynolds court noted that trial judges are wise to be suspicious of juror self-assessment of bias. But the Court reasoned that this is because jurors are likely to overstate

25 Burr, 25 F. Cas. at 50.
26 Further, see id. ("It is admitted that where there are strong personal prejudices, the person entertaining them is incapacitated as a juror ... Why do personal prejudices constitute a just cause of challenge? Solely because the individual who is under their influence is presumed to have a bias on [h]is mind which will prevent an impartial decision of the case according to the testimony.")
27 Id., at 51.
28 Dyke supra note 23 at 69.
29 Reynolds v. United States, 98 U.S. 145 (1878). Generally, there are two lines of cases about trial publicity: those that focus on publicity that occurred prior to trial and those that focus on publicity during the trial itself, including disruptions in the trial atmosphere. For our purposes, they can be treated together.
30 Id. at 148-50.
31 Id.
32 Id. at 156.
33 Id.
34 Id.
their biases in order to dodge jury duty. The errors in jury self-diagnoses – if any – are supposed to be of the harmless sort.

In *Marshall v. United States*, the trial judge refused a motion from the prosecution that would have admitted the defendant’s prior criminal record into evidence. However, that same information was published in the newspapers, and a “substantial number” of jurors read that information in the newspapers while the case was at trial. The U.S. Supreme Court overturned the conviction.

In *Irvin v. Dowd*, the Court addressed potential juror bias derived from pervasive pretrial publicity. The case was subject to a deluge of media about the upcoming trial, overwhelmingly negative about the defendant. In the end, the community was so biased against the defendant that 268 of 430 potential jurors were dismissed for cause because of their fixed belief in the guilt of the accused. Nonetheless 8 of the 12 jurors seated believed the defendant was guilty before the trial even began. Irvin was convicted and sentenced to death. On review, the Supreme Court overturned the conviction, citing jury bias, but offered little criteria by which to judge future cases. The Court wrote that on these facts, the jurors’ "statement of impartiality can be given little weight."

Likewise in *Rideau v. Louisiana*, the defendant’s confession was broadcast over television, and rerun over several days, eventually reaching 150,000 residents. Rideau was tried and convicted three weeks later, with three of the seated jurors admitting that they had seen the broadcast confession. The Court overturned Rideau’s conviction, referring to the proceedings as a "kangaroo court," and stating that the case should have been transferred to an unbiased venue.

In *Sheppard v. Maxwell*, the defendant was subject to an "editorial artillery" of unfavorable publicity, before and during trial. The Court said that, “where there is a reasonable likelihood that prejudicial news prior to trial will prevent a fair trial, the judge should continue the case until the threat abates, or transfer it to another county not so

35 Id. at 156 ("In considering such questions in a reviewing court, we ought not to be unmindful of the fact we have so often observed in our experience, that jurors not unfrequently [sic] seek to excuse themselves on the ground of having formed an opinion, when, on examination, it turns out that no real disqualification exists").
37 Id. at 311.
38 Id. at 312.
40 Id. at 725-7.
41 Id. (103 potential jurors were excused for conscientious objection to the death penalty, 20 were peremptorily challenged, and the rest were excused on personal grounds).
42 Id. at 727.
43 Id. at 728-9.
44 Id., at 728.
46 Id. at 726.
47 Id.
Juror Self-Diagnoses

permeated with publicity.”49 Sheppard is the apex of cases that display a serious concern about the potential impact of publicity, and the Court’s demand that trial courts take appropriate precautions, when necessary to ensure a fair trial.

In 1975, the Court affirmed the conviction in Murphy v. Florida.50 Murphy had gained notoriety for a jewel heist in 1964, years before being arrested for robbery in Dade County, Florida, in 1970.51 Murphy had also been convicted of one count of murder and had pled guilty to a federal conspiracy charge, all of it receiving extensive media coverage.52 At voir dire for the robbery charge, Murphy unsuccessfully moved to dismiss jurors for having learned of any of his previous convictions.53 The Supreme Court upheld his conviction and cited the passage of time since the most extensive news coverage of the earlier trials.54 Since the trial court did not have to dismiss more than 20 out of 78 jurors for having prejudged the defendant, the Court found that juror bias did not rise to the level required for change of venue.55 In dissent, Justice Brennan questioned the ability of jurors to self-assess bias derived from pretrial publicity: “[n]o doubt each juror was sincere when he said that he would be fair and impartial to petitioner, but the psychological impact requiring such a declaration before one’s fellows is often its father. Where so many, so many times, admitted prejudice, such a statement of impartiality can be given little weight.”56

In 1984, the Court upheld a conviction for murder in Patton v. Yount.57 The Pennsylvania Supreme Court had overturned the defendant’s conviction over a violation of Miranda v. Arizona.58 The confessions were reported widely in the local newspapers. The defendant was tried again and convicted.59 Although the venire panel at the second trial appeared to be highly biased against the defendant,60 the Court focused on the passage of four years time since the media exposure.61 The Court expressed trust in the ability of voir dire to produce neutral juries: “[i]t is fair to assume that the method we have relied on since the beginning ... usually identifies bias.”62


49 Id. at 363.
51 Id. at 795.
52 Id. at 796.
53 Id.
54 Id. at 802-3, and fn. 1. (the press coverage for Murphy’s earlier convictions occurred mostly from May 1968 to March 1969. Jury selection on the robbery charge began August 1970).
55 Id.
56 Id. at 808. (quoting Irvin, 366 U.S., at 728).
58 Id. at 1027.
59 Id.
60 Id. at 1029 (161 out of 163 potential jurors had heard of the defendant’s previous case, 126 out of 163 admitted preexisting opinions of the case, 8 out of 14 seated jurors admitted they had formed an opinion of the defendant’s guilt).
61 Id. at 1032-5 (“it is true that a number of jurors and veniremen testified that at one time they had held opinions, for many, time had weakened or eliminated any conviction they had had”).
62 Id. at 1038. See also id. at 1038-40 for the deference given to trial judges to determine juror credibility (“[j]urors thus cannot be expected invariably to express themselves carefully or even consistently. Every trial judge understands this, and under our system it is that judge who is best situated to determine competency to
In 1991, the Supreme Court upheld the conviction of Dawud Mu’Min for a murder he committed as an escapee from a prison work detail. Local media had covered the events leading up to the trial, including Mu’Min’s criminal history. Mu’Min moved for a change of venue, was denied, and was ultimately convicted. On appeal, the Court rejected his jury bias claim, again holding that jurors’ professed neutrality, when found credible by the trial judge, could be a basis for overruling the defendant’s objection.

Recently, the Court took up this question again in the case of former Enron executive Jeffrey Skilling, who was tried in Houston, the city where his company was headquartered and where several thousand people had lost their jobs and fortunes, as a result of its collapse. “Pointing to ‘the community passion aroused by Enron’s collapse and the vitriolic media treatment’ aimed at him, Skilling argue[d] that his trial ‘never should have proceeded in Houston.’” And even if it had been possible to select impartial jurors in Houston, “the truncated voir dire ... did almost nothing to weed out prejudices,” he contend[ed], so “far from rebutting the presumption of prejudice, the record below affirmatively confirmed it.”

The voir dire in Skilling’s trial was extensive, but led by the judge himself, who refused to allow the attorneys to question the jurors themselves. The trial judge had solemnly instructed the potential jurors that, “The bottom line is that we want ... jurors who ... will faithfully, conscientiously and impartially serve if selected.” Further, “each of you,” the court explained, “needs to be absolutely sure that your decisions concerning the facts will be based only on the evidence that you hear and read in this courtroom.” In all, two potential jurors were excused after stating they could not be impartial, one out of five was removed for cause at the government’s request, and three out of nine were removed for cause at the defendant’s request. Skilling was convicted of 19 counts and acquitted of 9 counts.

The Fifth Circuit reversed the conviction, holding that the “magnitude and negative tone of media attention directed at Enron” created a presumption of bias. The Supreme Court

64 Id. at 418.
65 Id. at 418-21.
66 Id. at 431.
68 Id. at 2912.
69 Id. at 2912.
70 Id. at 2910. The attorneys were permitted to ask follow-up questions.
71 Id. at 2910.
72 Id. at 2911.
73 Id.
74 Id. at 2911.
75 Id. at 2916.
reinstated the convictions, holding that Skilling was not denied a fair trial and that he did not prove that the jury was biased.\(^76\)

On the question of whether the jury was in fact biased by pretrial publicity, the Court expressed great deference for the trial court’s determination, which was itself based on the jurors’ own self-assessments and assurances of impartiality. As the Supreme Court emphasized, “in response to the question whether ‘any opinion [they] may have formed regarding Enron or [Skilling] [would] prevent’ their impartial consideration of the evidence at trial, every juror—despite options to mark ‘yes’ or ‘unsure’—instead checked ‘no.’”\(^77\) When rebutting Justice Sotomayor’s dissent, the Skilling majority again cited back to these juror self-diagnoses.\(^78\) A third time, the majority relied on the idea that, “all of Skilling’s jurors had already affirmed on their questionnaires that they would have no trouble basing a verdict only on the evidence at trial,” and emphasized that the trial court had nonetheless removed one such juror, who said he could ‘abide by the law.’\(^79\)

One particular juror “stated that ‘greed on Enron’s part’ triggered the company’s bankruptcy and that corporate executives, driven by avarice, ‘walk a line that stretches sometimes the legality of something.’”\(^80\) The Supreme Court nonetheless found it appropriate to seat the juror because “he also asserted that he could be fair and require the government to prove its case” and because the trial judge had “looked [Juror 11] in the eye and ... heard all his [answers], [and the trial judge] found his assertions of impartiality credible.”\(^81\)

The Skilling case makes clear that, going forward, courts should and will give considerable weight to the self-professed neutrality of jurors, but are also told to sometimes disregard professions if a stern “look in the eye” suggests that they are not credible. State courts likewise follow the procedure endorsed by Skilling, and some have statutes that require that “a judge must inquire whether a prospective juror has expressed or formed an opinion on a case or is aware of any bias or prejudice.”\(^82\)

\(^{76}\) The majority opinion cited four criteria when a change of motion should be granted to ensure a fair trial: (1) the size of the community in which the trial takes place; (2) the content of the trial coverage, whether or not any confessions or other “blatantly prejudicial information” is in the news; (3) the amount of time between the trial and the initial news coverage of the crime; and (4) whether or not the jury convicted the defendant of all counts against him. Id. at 2915-6.

\(^{77}\) Id. at 2921.

\(^{78}\) Id. at 2922 (“See supra, at 2919–2921 (noting, inter alia, that none of Skilling’s jurors answered ‘yes’ when asked if they ‘ha[d] an opinion about — Skilling’”).

\(^{79}\) Id. at 2922-2923, and n.30.

\(^{80}\) Id. at 2924.

\(^{81}\) Id. at 2924.


B. The Science on Jury Bias

1. The Biasing Effects of Pretrial Publicity

There is an extensive literature linking pretrial publicity to juror prejudice. Some studies suggest that greater exposure to media coverage produces greater prejudice among jurors. Costantini and King surveyed potential jurors in a single jurisdiction in which three different crimes were committed, publicized, and prosecuted. They found that, among various factors that predicted a juror’s likeliness to convict, exposure to pretrial publicity was by far the strongest indicator of a likelihood of conviction. Their study also found that higher levels of media consumption correlated with stronger feelings of defendant guilt.

Moran and Cutler surveyed potential jurors about actual criminal cases about to be tried within their jurisdictions. They found that potential jurors who had consumed more media coverage believed there to be more actual evidence against the defendant, conflating news coverage with admissible evidence.

Furthermore, Ruva et al. found that pretrial publicity that paints the defendant in a positive light tends to bias jurors in favor of the defendant when compared to neutral publicity. Their study also found that pretrial publicity that included prejudicial information about the defendant biased jurors against the defendant.

2. The Efficacy of Voir Dire for Removing that Bias

There is another body of research that investigates the efficacy of procedural safeguards against juror bias. Several studies cast doubt on the effectiveness of voir dire in producing a neutral jury.
Dexter et al. tested the effects of an extensive voir dire against a brief voir dire in their ability to persuade jurors to disregard prejudicial information they obtained through pretrial media consumption.92 The Dexter study did not dismiss jurors for bias, but only tested for a reduction in bias after exposure to the voir dire process.93 The extensive voir dire was no more effective in reducing the propensity of jurors to convict than the brief voir dire.94

A study by Kramer et al suggests that neither limiting instructions, juror deliberation, nor continuance succeeds in mitigating the effect of juror bias derived from pretrial publicity.95 There is one exception, however. Their study measured for the effects of pretrial publicity that was biasing for its factual content and pretrial publicity that was biasing for its emotional content. Continuance proved effective in reducing the bias from factual pretrial publicity but not emotional pretrial publicity.96 The authors found that jury deliberation might have actually enhanced the bias from pretrial publicity.97

Little work has been done on jury bias derived from pretrial publicity in civil cases. Kline and Jess exposed student mock juries to pretrial publicity then instructed to disregard the prejudicial information when deliberating.98 Of the four juries in the study, one of them referred to the prejudicial information contrary to the judge’s orders.99 Tanford and Cox exposed mock jurors to evidence that was supposed to be limited to the purpose of impeaching the defendant, but found that they were biased by that evidence despite the judge’s limiting instruction.100 In an unpublished study, Otto et al. exposed jurors to pretrial publicity in a civil case.101 Scholars discussing the Otto study noted that “the pretrial publicity influenced not only the judgments of negligence, but also impressions of the parties, memory, and inferences from the trial.”102 Other than a few studies, however, the bulk of the research on pretrial publicity has been focused on the criminal context.
Sue et al. conducted a brief study of university students that suggested that mock jurors were not capable of self-assessing their bias after exposure to pretrial publicity.\textsuperscript{103} The students were exposed to prejudicial publicity prior to reading a short written vignette about the trial evidence, and they convicted at significantly higher rates than control jurors, regardless of the juror’s professions that they could be unbiased in their judgment of the case.\textsuperscript{104}

One of the strongest studies was Kerr et al., which used a realistic videotape of a trial, and manipulated several types of pretrial publicity, prior to asking jurors questions including: “Can you put out of your mind any information you might have received from the newspapers or television and decide this case solely upon the evidence to be presented in court?”\textsuperscript{105} The scholars found that the pretrial publicity had no effect on individual verdict preferences immediately after the trial video, but did impact juror votes and jury verdicts after deliberation.\textsuperscript{106} Kerr and colleagues found that jurors “who admitted having formed an opinion from pretrial publicity that they could not set aside were no more or less likely to convict the defendant than those who reported no such fixed bias.”\textsuperscript{107} Kerr and colleagues also explored the ability of judges and attorneys to detect bias more holistically, after reviewing videotapes of colloquies with jurors that included that question. They were unable to select a jury with significantly lower conviction rates than a jury not subjected to voir dire.\textsuperscript{108} In other words, voir dire failed to detect and eliminate biased jurors.

Vidmar performed an extensive survey of various jurisdictions in preparation for the John Walker Lindh terrorism trial.\textsuperscript{109} Vidmar found that potential jurors with more exposure to pretrial publicity were more predisposed to finding guilt in the defendant. But he also found that about 35% of jurors who claimed they were able to be impartial had also expressed a presumption of Lindh’s guilt in the survey, raising doubts about the jurors’ ability to self-assess their own bias.\textsuperscript{110}

A study by Rose and Diamond showed that trial judges are more likely to believe a juror is capable of putting pretrial publicity aside when the juror expresses that belief with more

\textsuperscript{103} Stanley Sue, Ronald E. Smith, and George Pedroza, \textit{Authoritarianism, Pretrial Publicity, and Awareness of Bias in Simulated Jurors: 37 PSYCHOLOGICAL REPORTS 1299 (1975).}
\textsuperscript{104} \textit{Id.} at 1301.
\textsuperscript{105} See Kerr et al supra note 17 at 647(“[T]here were four levels of exposure to pretrial publicity in the Phase One experiment: exposure to both the emotional and the factual publicity, exposure to the emotional but not the factual publicity, exposure to the factual but not the emotional publicity, and exposure to neither form of publicity.”)
\textsuperscript{106} \textit{Id.} at 675.
\textsuperscript{107} \textit{Id.} at 690.
\textsuperscript{108} \textit{Id.} at 700.
\textsuperscript{109} Neil Vidmar, \textit{When All of Us are Victims: Juror Prejudice and ‘Terrorist’ Trials}, 78 CHICAGO-KENT LAW REVIEW 1143 (2003).
\textsuperscript{110} \textit{Id.} at 1163.
confidencenot. Although judges tended to give juror self-assessments of bias great weight, attorneys and laypersons were not as likely to believe the self-assessments.

In summary, the research has shown that exposure to pretrial publicity prejudices jury pools. However, the research also shows that the mechanisms in place to ensure a fair trial for defendants may be ineffective at eliminating, or even reducing, juror bias. As other scholars explain, “a review of the literature shows a paucity of recent, systematic scientific research on the mechanics of voir dire.” Nonetheless, judges are still relying heavily on juror self-assessments of bias when trying to construct impartial juries, apparently just assuming that the assessments have epistemic value. Indeed, the Supreme Court has simply ignored the foregoing scientific literature, and instead proceeded blithely, relying on jurors’ self-diagnoses to affirm convictions and the imposition of civil liabilities, without requiring evidence about whether those self-diagnoses are reliable. The studies described below attempt to either provide an evidentiary basis for this continuing and ubiquitous practice, or failing that, suggest a better way forward.

II. EXPERIMENTS

A. Design and Materials

We conducted two experiments, one with a sample of law students and a second with a larger and more diverse national sample online. Both experiments used the same design, stimuli, and instrument. In particular, we used a 2 x 1 between-subjects experimental design, wherein subjects were exposed to either irrelevant or prejudicial pretrial publicity concerning the defendant, respectively, in the control or treatment conditions. We then screened jurors out, following an operationalized version of the voir dire questioning and exclusions for cause endorsed by the Supreme Court.

1. Analytical Framework for Hypothesis Tests

The courts have long understood that the Constitutional guarantee of an “impartial” jury includes a guarantee that no juror is infected by “actual prejudice,” a term that they use interchangeably with “bias.” As Justice Marshall said, “the great value of the trial by jury certainly consists in its fairness and impartiality. Those who most prize the institution, prize it because it furnishes a tribunal which may be expected to be uninfluenced by an undue

---

111 See Rose and Diamond, supra note 17.
112 Id. at 542.
113 Hannaford-Agor and Waters, supra note 15 at 4.
114 For example, the 20-year old and relatively rigorous Kerr study described supra note 17 has never even been cited by the Supreme Court or a Court of Appeals. (Westlaw Keycite performed Dec. 8, 2012.)
115 See e.g., Skilling, 130 S.Ct. at 2917 (referring to “actual prejudice”) and id., at 2919 (referring to “bias”).
bias of the mind.” The Court has recognized that this is an empirical question: whether “the method we have relied on since the beginning . . . usually identifies bias.”

To test that assumption, we must specify a counterfactual for the null hypothesis. We ask whether the juror whose impartiality has been questioned due to some allegedly biasing factor can nonetheless decide the case the same as she would have decided the case without such a biasing factor. We do not require the juror to be ignorant of potentially biasing factors, but ask whether she can set aside those factors and decide the case in an unbiased way, similarly to those who are ignorant.

Thus, we first create a bias (the treatment group), to compare against a non-biased control group. We conceived the control condition as the gold standard of an “uninfluenced” fair trial, which the Skilling protocol seeks to replicate after jurors have been exposed to a biasing factor. Thus, if 30% of jurors impose liability in the control condition, and 50% impose liability in the treatment condition, then the Skilling protocol will succeed in its goal of providing a fair trial if the screened treatment condition replicates the 30% conviction rate, having thus neutralized the biasing influence. Likewise, if the Skilling protocol succeeds, damages awards should be indistinguishable in the control condition and the Skilling-screened treatment condition. In that way, we try to answer Justice Marshall’s question of whether the potentially biasing factor will “have a real influence on the verdict to be rendered.” We can observe empirically whether we have achieved “this mental attitude of appropriate indifference,” that the Skilling court says is “impartiality.”

We presume that the biasing factor will not infect 100% of the human subjects, but rather impact some on the margins. The remainder will decide the case just the same as if they had not been so exposed. Thus, amongst those exposed, we conceive the task as one of “diagnosis” – i.e., a sorting function to determine which ones were causally impacted by the biasing factor. If the diagnosis succeeds, we can remove those biased jurors. If the sorting function is instead indiscriminate (or perverse) then the treatment condition will remain worse for the defendant than the control condition. That is to say that the defendant is prejudiced by the pretrial publicity even after the Skilling protocol.

2. Biasing Stimuli

The pretrial publicity stimuli were based on either of two abridged articles from the Kansas City Star, each about 1300 words in length. The control article discussed

---

116 Burr, 25 F. Cas. at 50 (emphasis added).
117 Patton, 467 U.S. at 1038 (citing Burr).
118 Skilling, 130 S.Ct. at 2902 (“juror impartiality does not require ignorance.”). Id., at 2925 (jurors “need not enter the box with empty heads in order to determine the facts impartially.”)
119 Burr, 25 F. Cas. at 51.
120 Skilling, 130 S. Ct. at 2917 (quoting United States v. Wood, 299 U.S. 123, 145–146, (1936)).
121 The study materials are available upon request. The original articles are: Alan Bavley, Bad Medicine: Doctors with Many Malpractice Payments Keep Clean Licenses, KC STAR, Sept. 4, 2011 available at: http://www.kansascity.com/2011/09/04/3362970/bad-
employer incentive programs for preventative health maintenance, information which was irrelevant to the trial. The treatment article was modified to actually name the defendant in the trial, Dr. John Dennis, and discussed a prior case of medical malpractice, a history of prior malpractice claims and settlements much higher than the national average, and the effects of the malpractice on the other injured patients – all thereby painting him in a negative light. The article suggested that most medical malpractice is due to a few bad doctors, who need to be taken out of the system.\footnote{See Bavley, supra note 121 at \_\_ ("Nationwide, fewer than 2 percent of doctors have accounted for half the reported $67 billion paid out for malpractice claims in the United States since 1990, according to a study by Robert Oshel, a former analyst for the National Practitioner Data Bank. 'Taking that 2 percent of physicians out of practice would certainly make quite a difference,' Oshel said.")} The purpose of the Rules of Evidence is, of course, to carefully delimit the information available to jurors, so as to ensure that they base their decision only on proper evidence. The article used as “pretrial publicity” in the treatment condition included numerous things that the Rules of Evidence would preclude in a trial, making it thus prejudicial to the defendant. The information regarding Dr. Dennis’ prior incidents of malpractice would be inadmissible pursuant to Rule 404.\footnote{Rule 404(b) provides that, "Evidence of a crime, wrong, or other act is not admissible to prove a person’s character in order to show that on a particular occasion the person acted in accordance with the character."} The statements made by those who had previously sued Dr. Dennis would most likely be irrelevant under Rule 401 and are hearsay under Rule 801, and also excludable in a criminal trial under the Confrontation Clause. The statistics quoted in the article are hearsay. Thus, since this stimulus would be properly excluded from trial, if it nonetheless effects jurors’ decisions, we label that behavior “biased.”

3. Operationalizing the Supreme Court’s Voir Dire Protocol

Subjects in both conditions were given written admonitions from a mock judge about the need to be impartial, and they were asked a series of questions about whether they could be impartial, based on the judge-juror colloquy affirmed in \textit{Skilling v. U.S}. Specifically, the stimulus provided:

You are called into jury duty. After waiting in the jury commissioner’s office, you are ushered into the courtroom. The judge calls you to the bench individually, and he says:

You have been called to potentially be a juror in a medical malpractice case involving Mr. Andrew Stevens, as the plaintiff, suing Dr. John Dennis as the defendant. It is important for Mr. Stevens, Dr. Dennis, and for our legal system that the jurors be fair and impartial. Jurors must decide the case based only on the evidence presented during the trial, and not based on any prejudices, biases, preconceived ideas, or extraneous information.
The bottom line is that we want jurors who will faithfully, conscientiously, and impartially serve if selected. Each of you needs to be absolutely sure that your decisions concerning the facts will be based only on the evidence that you hear and read in this courtroom.

Unfortunately, I understand that some of you may have seen some news items about one or more of the parties in this case, or may have negative opinions about doctors, patients who sue, or the healthcare system at large. This fact alone does not automatically disqualify you from hearing this case however. You have a duty to perform your civic duty as a juror, if you can be fair and impartial in doing so.

Therefore, I am going to ask you a few questions. And, there are no right or wrong answers to the questions.”

Subjects then answered the following questions, with potential answers of “yes,” “no,” or “unsure”:

1. “Did you read a news article about Dr. John Dennis?”
2. “Do you have an opinion about Dr. John Dennis?”
3. “Would any opinion you have prevent your impartial consideration of the evidence at trial?”; and
4. “Could you base a verdict only on the evidence at trial?”

The first question was just a manipulation check. To simulate the screening process endorsed by the Supreme Court in *Skilling*, we used questions #3 and #4 for analysis. A subject answering “yes” or “unsure” to #3, saying that her opinions would or may prevent impartial consideration of the evidence, would be disqualified. And a juror answering “no” or “unsure” to #4, saying she could not base a verdict only on the evidence at trial, would be excluded from the jury. We understand that this is the test endorsed by the Supreme Court in *Skilling*, in accordance with prior cases going back to Justice Marshall’s opinion in *Burr*.

4. **The Trial Stimulus and Dependent Variables**

The foregoing screens were done merely for analytical purposes. Regardless of how they answered these questions, all subjects watched a 32-minute medical malpractice trial video that included opening statements from the plaintiff’s and defendant’s attorneys, testimony from expert witnesses about the standard of care in the case, cross-examination of both experts, and closing statements from the plaintiff’s and defendant’s attorneys. This video was developed by real physicians serving as writers of the medical scenario and serving as actors for the expert witnesses, along with an experienced arbitrator consulting

---

124 For these purposes, we also treated “unsure” as a disqualifying answer. See text accompanying note 77, supra (quoting the *Skilling* decision on this point). Robustness checks revealed that allowing “unsure” respondents to remain does not improve the accuracy of jury-self diagnosis.
on the jury instructions and serving as the judge. Two of the co-authors served as attorneys. Thus, although condensed, the video had a high degree of verisimilitude.

The scenario in the video concerned the failure of a primary care physician to diagnose a possible case of lumbar radiculopathy and refer the patient to imaging, which allegedly would have allowed timely surgery and avoidance of the permanent disability that the patient now suffers. The primary dispute concerned whether the physician-defendant met the standard of care when, instead of ordering imaging, he simply instructed the patient to take over-the-counter medications and return if the pain got worse. An actor posing as a judge provided jury instructions, based on the standard templates used in Arizona medical malpractice cases.

Subjects rendered individual judgments, responding “yes” or “no” to the prompt: “Based on the instructions provided by the judge in the video, do you believe that the Plaintiff has proved, by the greater weight of the evidence, that the Defendant committed medical negligence?” They also made Likert ratings on a one to six scale of whether this case was “certainly not medical negligence” (1) or “certainly medical negligence” (6). The jurors who found negligence also awarded non-economic damages for “pain and suffering,” which had been defined by the judge’s instructions. In Experiment 2, conducted with a national sample of 174 subjects, we also administered a demographic questionnaire, the “cognitive reflection test,” and the “need for cognition” instrument.

B. Experiment 1: Convenience Sample

The first of our two experiments involved a convenience sample of first year law students at the University of Arizona. Subjects were randomly assigned to unequal control ($n = 30$) and treatment ($n = 44$) groups. Demographic variables were not collected.

1. The Effect of Pretrial Publicity

Referring to the first major row of Table 1 (all jurors), only 13% (4 of 30) of the jurors in the control group voted against the defendant. Of the jurors who received negative pretrial publicity in the treatment condition, on the other hand, 32% (14 of 44) voted against the defendant. This 19% increase in verdicts against the defendant indicates that, as hypothesized, exposure to pretrial publicity biases jurors. The effect is substantial; with an odds ratio of 3.03, exposure to prejudicial publicity more than doubled the odds of a verdict.

---

125 Economic damages were not awarded, because the abridged trial did not include evidence thereof. We assumed they might be stipulated by the parties. Finally, we asked jurors to “in a sentence or two explain your answers.”

against the defendant. Statistically speaking the result is marginally significant $\chi^2 (1) = 3.31$, $p = .07$, likely the result of insufficient sample size.\textsuperscript{127}

The effect on awarded damages would be the real variable of interest to players in civil litigation. On this point, we find a clear, significant effect: jurors in the treatment ("exposed") condition imposed nearly nine times as much damages for pain and suffering as the jurors in the control condition (mean and 5% trimmed mean scores, respectively, of $98,500 and $81,286 versus $10,600 and $10,600; $U = 832, z = 2.42, p = .015$). Defense verdicts were counted as having damage awards of zero dollars, and since a preponderance of verdicts were for the defense, the median award in both conditions was zero. However, even when defense verdicts are excluded, we still see a robust shift in median damages awards, $33,000 in the control condition versus $300,000 in the treatment condition ($p = .05$). Thus, based on these findings, we concluded that our intervention succeeded in creating a bias.

2. The Accuracy of Jurors’ Self-Diagnoses

The more interesting question is whether the courts can use juror self-assessments to eliminate that bias, and thus secure the defendant’s right to a fair trial. Of the 44 respondents in the treatment group, five (11%) thought that they could not be impartial, 15 (34%) indicated that they were unsure, and 24 (55%) indicated that they could be impartial. Suppose that all those who thought the pretrial information would prevent them from being impartial, or were unsure about that point, were removed from the jury, which is the criterion suggested by Skilling.\textsuperscript{128} If the Skilling procedure is effective, then the remaining jurors (which we label “self-screened”) should vote similarly as do those in the control condition.

Of this allegedly impartial subgroup of the treatment condition ($n = 24$), seven (29%) voted against the defendant. Compared with the control group ($n = 24$), where only three (13%) voted against the defendant, it seems that the Skilling protocol failed – subjects exposed to prejudicial information, despite claiming not to be affected by it, were 2.88 times as likely to adjudicate that malpractice occurred than were subjects exposed only to irrelevant information. As before a problem of small sample size appears likely – indeed it

\textsuperscript{127} Textbook statistics generally recommend that all cell sizes within a chi-square test contain at least 5 observations. The reason is that estimates of the chi-square distribution rely on large-sample theory, an assumption which is possibly violated when cells are smaller than 5. There is, however, debate about whether and when small cell sizes undermine the chi-square test. See e.g. G. D. Ruxton, and M. Neuhauser, \textit{Good Practice in Testing for an Association in Contingency Tables}, \textbf{BEHAV. ECOL. SOCIOBIOL.} 64: 1505-1513 (2010). For our purposes with this initial pilot study, the $p = .069$ statement is suggestive enough to motivate the second experiment, without need to delve into this chi-square debate.

\textsuperscript{128} For our experimental purposes, we will exclude such jurors from both the control and treatment conditions, since jurors may have other bases for doubting their own partiality (e.g., a personal experience with malpractice), which we did not explore in our truncated voir dire. In a real trial, jurors may not be asked such a question with particularity, unless there were prima facie concerns about bias (which do not arise in our control condition). See e.g., \textit{Mu’Min}, 500 U.S. at 420 (discussing trial judge’s questioning procedure).
is exacerbated by the *Skilling* exclusions of subjects – with this difference failing to achieve traditional statistical difference, $\chi^2 (1) = 2.02, p = .16$.

Examination of pain and suffering awards, where mean and 5% trimmed mean scores awards of $87,542$ versus $15,000$ were observed for the treatment and control conditions, respectively, now shows a failure to reach traditional significance, $U = 341.5, z = 1.55, p = .12$. But this too likely reflects a problem with sample size, given that the mean award in the treatment condition was over six times as large as that within the control condition. Even excluding zeros, we see a quintupling of median damages awarded, from $60,000$ to $300,000$ ($p = .27$), though well short of statistical significance on this small sample.

Results on the 6-point scale as to the “certainty” of negligence (not shown) mimicked the above results. Treatment subjects exposed to prejudicial publicity were significantly more certain than those in the control condition exposed to irrelevant publicity (3.3 vs. 2.8, respectively), $t(72) = -2.47, p = .02$. And this remained true (albeit marginally, by traditional significance standards) even when excluding those who admitted or were unsure of bias as *Skilling* prescribes (for that subsample, scores were 3.2 vs. 2.8, respectively, $t(46) = -1.93, p = .06$). These suggestive results motivated a second study with greater statistical power, and a more representative sample.
Table 1: Convenience Sample. In a randomized experiment with 74 law students, those exposed to a news article that negatively portrayed the defendant voted against him more often ("liability votes"), and awarded higher damages (even after excluding zeros). Prior to viewing the trial, jurors were asked, "Would any opinion you have prevent your impartial consideration of the evidence at trial?" and "Could you base a verdict only on the evidence at trial?" Exclusion of jurors who said 'no' or 'unsure' (the "self-screen") failed to remove this bias. Statistical power was limited however.

<table>
<thead>
<tr>
<th>Jurors</th>
<th>N</th>
<th>Condition</th>
<th>Can be Impartial</th>
<th>Screened Out</th>
<th>Liability Votes</th>
<th>Bias Gap</th>
<th>p</th>
<th>Odds Ratio (95CI)</th>
<th>Mean Damages</th>
<th>5% Trim</th>
<th>SD</th>
<th>Nonzero Median</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>74</td>
<td>Unexposed</td>
<td>30, 24 (80%)</td>
<td>6 (20%)</td>
<td>0 (0%)</td>
<td>4 (13%)</td>
<td>19%</td>
<td>0.07</td>
<td>3.03 (0.9–10)</td>
<td>$11k</td>
<td>$11k</td>
<td>$47k</td>
<td>$33k</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exposed</td>
<td>44, 24 (55%)</td>
<td>15 (34%)</td>
<td>5 (11%)</td>
<td>14 (32%)</td>
<td></td>
<td></td>
<td>$99k</td>
<td>$82k</td>
<td>$175k</td>
<td>$300k</td>
<td></td>
</tr>
<tr>
<td>Self-</td>
<td>48</td>
<td>Unexposed</td>
<td>24 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>3 (13%)</td>
<td>16%</td>
<td>0.16</td>
<td>2.88 (0.7–12)</td>
<td>$15k</td>
<td>$15k</td>
<td>$54k</td>
<td>$60k</td>
</tr>
<tr>
<td>screened</td>
<td></td>
<td>Exposed</td>
<td>24 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>7 (29%)</td>
<td></td>
<td></td>
<td>$88k</td>
<td>$88k</td>
<td>$172k</td>
<td>$300k</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: National Sample. In a randomized experiment 174 jury-eligible persons recruited online, those exposed to a news article that negatively portrayed the defendant voted against him more often ("liability votes"), and awarded higher damages (even after excluding zeros). Prior to viewing the trial, jurors were asked, "Would any opinion you have prevent your impartial consideration of the evidence at trial?" and "Could you base a verdict only on the evidence at trial?" Exclusion of jurors who said 'no' or 'unsure' (the "self-screen") failed to remove this bias.

<table>
<thead>
<tr>
<th>Jurors</th>
<th>N</th>
<th>Condition</th>
<th>Can be Impartial</th>
<th>Screened Out</th>
<th>Liability Votes</th>
<th>Bias Gap</th>
<th>p</th>
<th>Odds Ratio (95CI)</th>
<th>Mean Damages</th>
<th>5% Trim</th>
<th>SD</th>
<th>Nonzero Median</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>174</td>
<td>unexposed</td>
<td>65, 59 (91%)</td>
<td>6 (9%)</td>
<td>0 (0%)</td>
<td>23 (35%)</td>
<td>17%</td>
<td>0.03</td>
<td>2.00 (1.1–3.8)</td>
<td>$293k</td>
<td>$72k</td>
<td>$1,303k</td>
<td>$150k</td>
</tr>
<tr>
<td></td>
<td></td>
<td>exposed</td>
<td>109, 95 (87%)</td>
<td>9 (8%)</td>
<td>5 (5%)</td>
<td>57 (52%)</td>
<td></td>
<td></td>
<td>$438k</td>
<td>$345k</td>
<td>$711k</td>
<td>$500k</td>
<td></td>
</tr>
<tr>
<td>Self-</td>
<td>154</td>
<td>unexposed</td>
<td>59 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>20 (34%)</td>
<td>19%</td>
<td>0.02</td>
<td>2.17 (1.1–4.3)</td>
<td>$309k</td>
<td>$76k</td>
<td>$1,366k</td>
<td>$100k</td>
</tr>
<tr>
<td>screened</td>
<td></td>
<td>exposed</td>
<td>95 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>50 (53%)</td>
<td></td>
<td></td>
<td>$442k</td>
<td>$442k</td>
<td>$729k</td>
<td>$500k</td>
<td></td>
</tr>
</tbody>
</table>
C.  **Experiment 2: National Sample**

We recruited subjects from the population of workers on Amazon Mechanical Turk ("Mturk") in May and July 2012, screening for those that were “jury eligible,” meaning residents of the United States, over age 18, who could read, write, and speak English. Subjects were paid four dollars each to complete the experiment online, with an opportunity of bonus pay for respondents who scored highest on tests of recall (to incentivize attention and effort within the experimental task, given that subjects might otherwise cheat). The experiment was identical to that described above for the law school sample, although demographic information and cognitive scales were also collected for this Mturk sample. All subjects consented in accordance with IRB requirements.

1.  **The Sample Demographics**

We recruited 264 persons who proceeded past the informed consent webpage into the experiment. Sixty-four persons exited the study before completion, constituting an attrition rate of 24%. An additional 26 persons were excluded for failure to comply with the task – they finished the experiment, which entails a 32 minute video and several pages of questions, in the impossibly fast time of 34 minutes or less.\(^{129}\) The final sample thus includes 174 subjects. Demographic variables of sex, age, education, and gender were examined to explore whether characteristics of the person predicted whether he or she dropped out or cheated. None had predictive power; the excluded group demographically resembled the final sample.

As can be seen in Table 3, the overall sample roughly resembled U.S. Census data, although our subjects were on average more educated and younger, regression analyses displayed in the Appendix showed that neither demographic significantly affected verdict. Especially relative to jury research using convenience college samples, our sample constitutes a respectable level of representativeness, and allows reasonable inferences about the jury pool at large.

Demographic variations were fairly well distributed across the experimental conditions, showing that randomization succeeded. Subsequent analyses thus proceed using condition-splits, proportions, and central tendencies for ease of communication. Regression tables are presented in the Appendix.\(^{130}\)

---

\(^{129}\) The length of the video – 32 minutes – is an obvious and objective threshold for exclusion. Setting a threshold of greater length requires an estimate of how quickly it would be possible to read and answer the task materials, which in turn requires a judgment that risks the possibility of a false positive (wrongly excluding a subject who happens to work quickly). Examination of the data revealed a sharp gap, with two distributions. Those labeled a “cheater” almost all finished under the 32-minute mark; only two persons took longer, and each were below 34 minutes. The tail of the other distribution, which represents the quickest “non-cheating” score, was 44 minutes.

\(^{130}\) Appendix A, Table A1, reports the results of regression analyses of models testing condition as a predictor, as well as nested models testing for any additional variance accounted for by self-diagnosis,
The overall sample from Amazon Mechanical Turk roughly resembled U.S. census data, although it was on average younger and more educated. The control and treatment conditions were relatively similar in their demographic distributions.

<table>
<thead>
<tr>
<th>Education</th>
<th>Control (n = 65)</th>
<th>Treatment (n = 109)</th>
<th>Total (N = 174)</th>
<th>U.S. Census</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; HS Diploma/GED</td>
<td>3%</td>
<td>1%</td>
<td>2%</td>
<td>18%</td>
</tr>
<tr>
<td>HS Diploma/GED</td>
<td>15%</td>
<td>9%</td>
<td>12%</td>
<td>30%</td>
</tr>
<tr>
<td>Some College/Assoc.</td>
<td>49%</td>
<td>48%</td>
<td>48%</td>
<td>27%</td>
</tr>
<tr>
<td>College Grad</td>
<td>25%</td>
<td>30%</td>
<td>28%</td>
<td>17%</td>
</tr>
<tr>
<td>Graduate Degree</td>
<td>8%</td>
<td>12%</td>
<td>10%</td>
<td>10%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>42%</td>
<td>45%</td>
<td>44%</td>
<td>49%</td>
</tr>
<tr>
<td>Female</td>
<td>58%</td>
<td>55%</td>
<td>56%</td>
<td>51%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age Groups</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>18-24</td>
<td>17%</td>
<td>28%</td>
<td>23%</td>
<td>13%</td>
</tr>
<tr>
<td>25-34</td>
<td>54%</td>
<td>31%</td>
<td>40%</td>
<td>18%</td>
</tr>
<tr>
<td>35-44</td>
<td>17%</td>
<td>20%</td>
<td>19%</td>
<td>19%</td>
</tr>
<tr>
<td>45-59</td>
<td>11%</td>
<td>18%</td>
<td>15%</td>
<td>27%</td>
</tr>
<tr>
<td>60+</td>
<td>1%</td>
<td>3%</td>
<td>3%</td>
<td>23%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Race</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>80%</td>
<td>77%</td>
<td>78%</td>
<td>74%</td>
</tr>
<tr>
<td>Non-White</td>
<td>20%</td>
<td>23%</td>
<td>22%</td>
<td>26%</td>
</tr>
</tbody>
</table>

2. **The Effect of Pretrial Publicity**

As with the pilot experiment, our analyses proceeded in two steps. We first asked whether pretrial publicity has a biasing effect by examining if, as hypothesized, subjects exposed to relevant publicity were more likely to find that medical malpractice occurred, to award larger monetary damages, and to be more certain of their verdict than those subjects exposed to irrelevant publicity. Second, we examined whether, as hypothesized, the screen proposed by the Supreme Court, most recently in *Skilling*, was sufficient to remove the biasing effect of prejudicial publicity observed in the first step.

Exposure to prejudicial pretrial publicity did, as hypothesized, significantly bias jurors. As seen in Table 2 ("all jurors" row), only 35% of jurors in the control condition found psychological, and demographic constructs. The regression analyses reveal conclusions that are equivalent to the chi-square analyses discussed in the body. In particular, a model regressing verdict on condition is significant, $\chi^2 (1) = 4.74, p = .029$, and remains significant after exclusion of jurors who fail the *Skilling* screening, $\chi^2 (1) = 5.22, p = .022$. Nested model comparisons revealed that self-diagnosis, NFC, and CRT added no significant predictive power, $\chi^2 (3) = 1.06, p = .787$, nor did any of the demographic variables (race, sex, education, age), $\chi^2 (4) = 0.34, p = .987$. 

---

25
medical negligence, but this percentage significantly increased to 52% amongst those jurors exposed to prejudicial publicity, $\chi^2 (1) = 4.687$, $p = .030$, $\phi = .16$. This amounts to an odds ratio of 2.00, meaning that exposure to prejudicial pretrial publicity doubled the odds of a verdict against the defendant. It is worth emphasizing that the trial stimulus itself was exactly identical across conditions, and thus a strong inference can be made that the prejudicial publicity was the casual source of this change in verdict rates.

The effect on pain and suffering damages, the ultimate variable of interest to practicing litigators, was more dramatic. The median and 5% trimmed mean awards in the control condition were $0$ and $71,943$, respectively ($SD = 1,302,768$). But when the taint of prejudicial pretrial publicity was introduced (in the treatment condition), median and 5% trimmed mean awards increased to $50,000$ and $345,178$, respectively ($SD = 710,960$). This is highly significant with a medium effect size, $U = 4,505$, $Z = 3.127$, $p = .002$, $r = .24$. Even when defense verdicts are excluded, the median damages award increased, more than tripling, from $150,000$ to $500,000$, $p = .02$.

3. The Accuracy of Jurors’ Self-Diagnoses

If, as Skilling prescribes, jurors who admit bias or are unsure of bias are removed, do the effects of pretrial publicity disappear? No, the Skilling protocol failed; very few people admitted bias in the first place, and those who did were equally likely to impose liability. Referencing Table 2, the vast majority of jurors denied bias and instead expressed a sureness that they would be able to impartially consider only the evidence presented at trial (91% and 87% in the control and treatment conditions, respectively). Notably, comparing those who deny bias against those who are either unsure or admit bias across conditions, we find no significant difference, $\chi^2 (1) = 0.523$, $p = .470$; that is, subjects were equally likely to admit bias regardless of whether they read irrelevant or prejudicial pretrial materials.

Excluding pursuant to the Skilling protocol, we are left with 59 and 95 subjects in the control and treatment conditions, respectively. As shown in Table 2 (self-screened row), this self-diagnosing protocol completely failed. The verdict rates remained unchanged

131 Given the high variability of damages awards, particularly the presence of a handful of extreme outliers, the mean can be misleading. Here, for example, the mean for the control condition ($292,540$) is over $220,000$ larger than the trimmed mean of $71,942$, and this difference is driven by a single outlying score of $10,000,000$.

132 The biasing effect of pretrial publicity was not as clearly reflected in certainty scores (not shown), a measure that is not asked of real jurors. The mean rating, from 1 (“certainly not medical negligence”) to 6 (“certainly medical negligence”) was 3.23 in the control condition, and only increased to 3.61 in the treatment condition — a trending but non-significant result, $t(172) = -1.581$, $p = .116$. Although one might assume that binary verdicts can be directly inferred by splitting the scale in half (i.e., those finding for the defendant rate the case from 1-3, while those finding against the defendant provide a 4-6 rating), that is apparently not the case for all subjects. Just under 5% (4 of 94) of those ruling in favor of the defendant rate the case from 1-3, while those finding against the defendant provide a 4-6 rating), that is apparently not the case for all subjects. Just under 5% (4 of 94) of those ruling in favor of the defendant nonetheless gave a rating of 4 or higher; a similar number (3 of 80) of those against the defendant nonetheless gave a rating of 3 or lower. Screening of jurors made little difference on this measure.

133 For the Skilling criteria used, see section II.A.3 supra.

134 Again chi-square statistics from the contingency table are reported for ease of explication, but regression analyses reveal equivalent results. See Appendix A, Table A2.
almost to the digit: before screening, 35% and 52% of control and treatment subjects found against the defendant; after screening, the rates were 34% and 53%. The 34% versus 53% difference continues to be both significant and meaningful, \( \chi^2 (1) = 5.152, p = .023, \phi = .18 \); this amounts to an odds ratio of 2.17, or in other words, a more than doubling of the odds of a verdict against the defendant, given exposure to prejudicial pretrial publicity.

The failure of the *Skilling* protocol is again on dramatic display in the pain and suffering awards. (See Table 2, self-screened row). The median and trimmed mean awards in the control condition remained about the same as they were prior to applying the *Skilling* screen: $0 and $70,536, respectively (SD = $1,366,095). But application of the *Skilling* protocol did not cure the taint of prejudicial pretrial publicity as hoped in the treatment condition; rather, the median and trimmed mean awards remained higher than in the control condition, at $90,000 and $345,404, respectively (SD = $729,406). This difference across conditions continues to be highly significant with a medium effect size, \( U = 3,553, Z = 2.930, p = .003, r = .24 \). Even when defense verdicts are excluded, median damages awards quintupled, from $100,000 to $500,000, \( p = .03 \).

4. **Does Juror Self-Diagnosis Help At All?**

So, far we have compared the exposed group to the unexposed group, after imposing a screen on the basis of juror self-diagnosis. For the reasons explained in Part II.A.1 above, we have defined success as the screened exposed group performing similarly to the screened unexposed group, as this would show that self-diagnoses screens create an "impartial" jury, as the Constitution guarantees. However, even if self-diagnosis does not achieve the gold standard of impartiality, it might still be better than nothing. For that purpose, one might simply examine the exposed condition before and after screening.

Our data fails to support the hypothesis that juror self-diagnoses help at all, but we cannot rule out competing hypotheses that they help somewhat. Our best estimate is that self-diagnoses make matters slightly worse. As shown in Table 2, examining the verdict rates within the exposed conditions for all versus self-screened jury pools, the observed verdict rates against the defendant increase from 52% to 53%. This difference is far from statistical significance.\(^\text{135}\) Indeed, examining damages, nonzero median damages remain unchanged after screening.

The confidence interval around our calculated odds ratios provides further elucidation on this point. For all jurors, pretrial publicity doubles the odds of judgment against the defendant (odds ratio = 2.00). For self-screened jurors, the self-diagnoses actually made matters worse (odds ratio = 2.17). The point estimate (2.17) “is the one value most likely to represent the true” odds ratio.\(^\text{136}\) Nonetheless, the confidence interval around our point

---

\(^\text{135}\) \( \chi^2 (1) = 0.034, p = .854 \).

\(^\text{136}\) Guyatt, Gordon H., et al. 'Users’ guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients.” 271 JAMA 59 (1994) (The article continues: “As one considers values farther and farther from the point estimate, they become less
Juror Self-Diagnoses

estimate ranges from 1.1 to 4.3, meaning that while (at traditional levels of statistical significance) we can rule out the hypothesis that self-screening completely eliminates bias (a 1.0 odds ratio), we cannot rule out the hypothesis that self-diagnosis reduces the bias, say by half (a 1.5 odds ratio). Nor can we rule out the hypothesis that self-diagnosis makes the bias problem twice as worse (a 4.0 odds ratio). It is, of course, difficult to prove a negative. Rather than scholars trying to endlessly falsify smaller and smaller hypotheses about the efficacy of this self-diagnosis procedure, it may be more sensible for those who propose to use this sort of method to instead show that it is at all reliable. To be sure, our data do not support the hypothesis that asking jurors whether they can be impartial is at all diagnostic of partiality versus impartiality, and our best estimate is that it makes matters slightly worse.

5. Opinion-Screening as an Alternative to the Skilling Protocol

Recall that in Skilling, the Supreme Court again endorsed the protocol of judges first asking jurors whether they have any opinion about the case or the parties, then a pair of questions about whether jurors can overcome any opinions and thus base a verdict on only the evidence at trial. The latter two questions then become the basis for screening under the Skilling protocol, and in our experiments described herein.

The juror's task of introspecting to identify an existing opinion (the first question) is a different cognitive task from assessing one's ability to overcome that opinion (the second and third questions). Thus a subject might fail the latter but nonetheless be proficient at the former.

One reason that the answers to the latter questions turn out to be unreliable may be "social desirability" bias. A juror, despite accurately self-diagnosing that he or she cannot overcome a lingering bias, might nonetheless publicly insist that he or she will act impartially. The latter questions essentially ask whether a juror will adhere to the social norms of being a good and fair person – whether he or she will uphold a civic responsibility as any decent person would. Thus there is pressure to respond consistently with the

137 Indeed, to rule out the hypothesis that self-diagnosis screening reduces bias in half (a 1.5 odds ratio), it would be necessary to quadruple our sample size in the self-screened condition (i.e., from N = 154 to N = 616) in order to reach a confidence interval of (1.54-3.03). As the competing hypothesis becomes smaller and smaller, the needed sample size to falsify the hypothesis approaches infinity.

138 See Steven I. Friedland, Paul Bergman, & Andrew E. Taslitz, Evidence Law & Practice (3rd ed., 2007), 375 (“[T]he proponent of the evidence has the burden of proving that a scientific principle or technique is reliable[.]”)


140 See Broeder supra note 91 at 526-27 ("On its face, probably no question commonly asked on voir dire seems as innocuous as," Can you be fair and impartial?" And it was - when addressed to the panel generally.
norm, lest one appear to be a social deviant, one with the undesirable trait of being uncontrollably swayed by bias.\textsuperscript{141}

These considerations suggest a modification to the Supreme Court’s questioning protocol to achieve greater diagnosticity. If courts were to exclude anyone who admits or is unsure of even having an opinion about the case, the response might be more useful than the \textit{Skilling} protocol for constructing the fair and impartial jury required by the Constitution.

To assess the possible efficacy of such a revised protocol, we compared the verdict rates of the pre-\textit{Skilling} protocol jury pool to a jury pool from which we excluded anyone who even admitted (or was unsure about) having an opinion about the defendant. Out of 95 jurors that insisted that they could be fair and impartial in the treatment condition, 61 had admitted (or were unsure about) having an opinion about the defendant, but believed that they could suppress that opinion and focus solely on the evidence at trial. That left 34 who stated they had no opinion whatsoever about the defendant, and thus were eligible under our proposed opinion-eligibility standard.

Despite such an aggressive exclusion protocol, juror bias remains rampant. Prior to the \textit{Skilling} screen, 35\% and 52\% of control and treatment subjects, respectively, found against the defendant, a bias gap of 17\%. After the more stringent screen based on mere opinion, the rates were 33\% and 56\%, a bias gap of 23\%. This difference across control and treatment conditions continues to be both significant and meaningful, $\chi^2 (1) = 4.452$, $p = .035$, $\phi = .21$; odds ratio = 2.53. The bias gap got worse, not better. The failure of the aggressive opinion-screening protocol is again reflected in the damages awards. The trimmed mean awards in the control condition remained about the same as the bias-screened jurors at $75,898 (SD = 1,389,038)$ versus $442,320 (SD = 781,066)$ in the exposed condition. The six-times disparity in damages caused by exposure to publicity remained, regardless of the more aggressive screening protocol.

6. \textbf{The Unexplained Behavior of Fair Jurors Disqualifying Themselves}

If the juror’s actual bias does is not what causes her to say she is biased, then what does cause some jurors to answer the \textit{Skilling} questions differently than other jurors? We hypothesized that perhaps the most earnest and thoughtful jurors may be more likely to admit concerns about their own impartiality. Need for Cognition (NFC) and Cognitive

\textsuperscript{141} Id. (providing example of social desirability bias in a different civic responsibility setting, that of voting in presidential elections).
Reflection Test (CRT) scales were administered to explore individual differences in the ability to self-diagnosis bias.\textsuperscript{142}

NFC is a personality variable reflecting the extent to which people engage in and enjoy exerting cognitive effort. We hypothesized that those higher in Need for Cognition would be more motivated to fully consider the Supreme Court’s prescreening questions, and thus be more likely to self-diagnose bias caused by exposure to the pretrial article about the defendant.

The CRT is designed to assess the degree to which individuals suppress an intuitive and spontaneous wrong answer in favor of a reflective and deliberative right answer. We hypothesized that those performing more highly on the Cognitive Reflection Test would be more likely to stop and think carefully about the Skilling prescreening questions, rather than relying on a quick intuitive assessment of being unbiased, and thus be more likely to self-diagnose bias caused by exposure to the pretrial article about the defendant.

To test the above possibilities, we fitted a binary logistic regression model predicting Skilling eligibility from NFC, CRT, and demographic variables on data from the treatment condition.\textsuperscript{143} As shown in the Appendix, the model itself was far from significant, and contrary to our hypotheses, neither NFC nor CRT nor any of the demographic variables predicted which jurors would disqualify themselves as Skilling ineligible ($\chi^2 (9) = 3.904, p = .918$).\textsuperscript{144} It remains a mystery as to what causes some jurors to disqualify themselves for potential bias, even when it turns out that they were just as likely to rule against the defendant as the other jurors.

\section*{D. Overall Analysis}

It is notable that our two experiments involved very different populations, who had very different base rates for imposing liability, and very different levels of confidence in their own ability to be fair and impartial. (Compare Table 1 and Table 2.) In the control condition 13\% of the law students imposed liability, while 35\% of the respondents in the national sample imposed liability. We also saw very different rates of self-diagnoses, with 45\% of law students saying that they were unable to be fair and impartial (or unsure), versus only 9\% of the respondents in the online national sample. Notwithstanding these differences, we saw similar effects of pretrial publicity and a similar failure of our screening protocol to remove that bias. This finding should enhance readers’ confidence in the external validity of our studies; they do not seem to be driven by peculiarities about a particular subject pool.

\textsuperscript{142} For reference articles on the CRT and NFC, see note 126 supra.
\textsuperscript{143} The control condition was not included because, without knowledge of what biased opinions a juror might or might not harbor, it is impossible to predict the effect of NFC and CRT. In the treatment condition, in contrast, we assume jurors harbor the pretrial publicity bias that we experimentally induced, and thus can articulate hypotheses regarding the ability to self-diagnose that bias.
Juror Self-Diagnoses

Figure 1 displays the combined data from the 248 subjects in Experiments 1 and 2, including verdicts, damages awarded, and whether the juror was screened based on their self-diagnoses as to whether they could be fair and impartial (the Supreme Court’s “Skilling protocol”, as we have called it). This graphic depicts the upwards skew of awards in the exposed condition, compared to the unexposed condition. For the Skilling protocol to successfully remove bias, it would need to edit the distribution on the right to make it appear like the distribution on the left. However, the paucity and improper distribution of juror self-diagnoses shows the failure of this protocol to correct for the induced bias.

Examining verdict rates across both the convenience and national samples, 27 of 95 (28%) unexposed persons found negligence, while 71 of 152 (47%) exposed persons found negligence. This is significantly different as predicted, $\chi^2 (1) = 7.931, p = .005$, odds ratio = 2.181, and indicates successful induction of a pretrial publicity bias. The Supreme Court’s method of excluding jurors – those who thought themselves to be unable to be fair or impartial or unsure – removed 46 of our jurors (about 19% of our sample). Nonetheless the method altogether failed to correct this doubling of the odds of a liability verdict, leaving the percents finding negligence almost unchanged after screening (28% in the exposed condition versus 48% in the exposed condition). The difference remained highly significant, $\chi^2 (1) = 8.332, p = .004$, odds ratio = 2.398.

Examining monetary awards (including $0 verdicts), the 95 unexposed persons had a median and mean (SD) award of $0 and $95,374 ($304,077). The 152 exposed persons had a median and mean (SD) award of $0 and $275,739 ($460,700). This tripling of mean damages awards across conditions is highly significant, $U = 8,872, Z = 3.331, p = .001$. Screening out the jurors based on their self-diagnoses failed to cure this bias, $U = 6,117, Z = 3.290, p = .001$.

Examining monetary awards conditional on a finding of negligence, the 27 unexposed persons who found negligence had a median and mean (SD) award of $60,000 and $339,092 ($498,817). The 71 exposed persons who found negligence had a median and mean (SD) award of $500,000 and $594,197 ($518,597). With median damages awards eight times higher in the exposed condition, this difference across conditions is highly significant, $U = 1,324, Z = 2.920, p = .004$, and indicates that pretrial publicity has an effect above and beyond the impact on verdicts. The application of a self-diagnosis screen again failed to cure, with median damages awards remaining unchanged at $60,000 in the unexposed condition versus $500,000 in the exposed condition, $U = 897, Z = 2.581, p = .010$. 

31
Figure 1 – Plot of Combined data from Experiments 1 & 2 by Condition with Verdict, Damages (Capped at $1.5M), and Self-Screen – Each dot represents a single juror (n=248), with defense verdicts shown in a block on the bottom and plaintiffs’ verdicts shown by amount of damages awarded, as a dot-histogram with $50,000 bins. Jurors who said they were unable to be fair and impartial or were unsure (the “self-screened”) are shown as yellow-striped. Removal of those jurors would not cure the upwards skew (bias) in the condition where subjects were “exposed” to pretrial publicity adverse to the defendant.
E. Limitations

Our study had several limitations. First, we tested a particular type of juror bias—the bias due to pretrial publicity, and indeed used a particular form of pretrial publicity based on a real newspaper article.\textsuperscript{145} Publicity is amenable to manipulation and thus randomized experimentation, unlike other sources of biases, such as personal acquaintance with the parties or racial animus. Whether and to what extent jurors are biased from other sources, and whether they may be better able to self-diagnose those biases, are open questions. We are unaware, however, of any empirical evidence that would suggest that jurors are better able to self-diagnose in these other domains.

Second we should emphasize that we have not calibrated the degree of bias observed in our mock jurors with the amount of bias that may or may not have infected any particular juror in any particular case. (Nor could we.) It is possible that none of the jurors in the Skilling case were biased at all, or that they were collectively more biased than the mock jurors in our study. Relatedly, our data cannot say whether the amount of bias shown in our study (a more than doubling of odds of imposing liability on the defendant) is “too much” bias to be Constitutionally tolerable.\textsuperscript{146} Maybe a court would say that a doubling of odds is “good enough impartiality,” noting that nearly half of the jurors still exonerated the defendant even with the publicity.\textsuperscript{147} Regardless of such line drawing problems, our experiment instead focuses on the epistemic device that the courts use to assess bias. In this sense, our study does suggest that if courts use this device to reassure themselves that jurors were in fact unbiased—below whatever threshold for too much bias they select—then the courts’ conclusions are unwarranted on that basis.

Third, we used a 32-minute abridged civil trial for our experimental stimulus. The condensed stimulus allowed us to utilize a randomized controlled trial experimental design, which is the gold-standard for scientific research. Still, there are reasonable concerns about external validity. Specifically, ours was a civil trial, but pretrial publicity problems often arise in a criminal context instead. We are unaware of any evidence that jurors called for criminal trials are somehow better able to diagnose their own biases than jurors called for civil trials, though one could speculate that the difference in standards (preponderance

---

\textsuperscript{145} Not unlike real instances of pretrial publicity, our stimulus was complex, consisting of multiple components in a 1300-word article. Thus, we are unable to say which aspects of our stimulus were most biasing. Was it the prior instances of misconduct attributed to the physician-defendant? Was it the discussion of the suffering of another plaintiff? Perhaps the large numbers mentioned in the stimulus, even had an anchoring effect? Future research could try to disaggregate these effects. One might speculate that jurors would be better able to self-diagnose for some causes of biases over others.

\textsuperscript{146} Presumably, a small amount of partiality is permissible. See e.g., Skilling, 130 S.Ct. at 2914 (describing prior cases as ones where the “trial atmosphere was utterly corrupted by press coverage”); id, at 2913 (asking whether there was “extraordinary local prejudice”), id, at 2922 (describing the “deep and bitter prejudice” in a prior case). Notably, however, some of these quotations refer to the distinct legal question of whether prejudice should be presumed, as distinct from the question of whether a particular juror suffered from actual prejudice. See id, at 2917 (making this transition in two different analyses).

\textsuperscript{147} On the other hand, courts have held that a single biased juror on the panel of twelve is too much. See sources cited at note 7, supra.
versus beyond a reasonable doubt) could matter. The length of our stimulus also raises the possibility that in a real trial, which may last for days or even weeks and where biasing factors are less proximate, jurors would more heavily weight the evidence presented therein, and thus be less subject to pretrial publicity biases at all. \textsuperscript{148} These limitations apply more to the question of whether jurors are biased, as opposed to whether they are able to self-diagnose that bias, the question studied here.

Fourth, the experiments were conducted with a convenience sample of law students and a national sample of human subjects online, who reached individual judgments rather than collective jury verdicts after deliberation. It is possible that real jurors in real courthouses are somehow better able to diagnose their own biases. \textsuperscript{149} Prior research has shown that “the population of Mechanical Turk is at least as representative of the U.S. population as traditional subject pools.” \textsuperscript{150} Known experimental results have been replicated using the Mturk population. \textsuperscript{151} Still, it is likely that Mturkers may be more easily distracted from the trial compared to real jurors, and may even provide junk responses. Although we paid respondents rather generous bonuses contingent on their measurable performance on attention tasks, such problems could increase noise in the data. It may be that real jurors are more earnest in their efforts to diagnose their own biases. On the other hand, real jurors may have other motivations for saying that they are unbiased (if they have an axe to grind against the defendant, or the social pressure of answering in public) or biased (if they would prefer not to serve on the jury), which would further reduce the diagnosticity of the questioning.

Fifth, we merely tested whether jurors could diagnose their own biases, and thereby provide reliable information to the judge tasked with deciding whether to exclude the juror, or change the venue. One could speculate that "by looking the juror in the eye," as the Supreme Court suggests, judges are able to ascertain whether he or she can be impartial in a

\textsuperscript{148} In fact, we observed some heterogeneity even within our sample as to who people react to pretrial publicity. At least in terms of binary verdicts, many subjects voted the same way as they would have in the unexposed condition, since the exposure only created a 19% bias gap on the margin in the national sample. Indeed, we used this same stimulus with another convenience sample of law students and found an insignificant biasing effect, which thus prevented us from testing whether subjects could self-diagnose those biases in that sample. The experiments reported herein presume the existence of a biasing factor.

\textsuperscript{149} But see Suggs and Sales supra, note 15 at _ (arguing that courtrooms are particularly bad contexts for elucidating candid responses, given modes of questioning, interaction distance, and formality, which potentially makes real world voir dire practices even worse than the experimental procedures employed here) and Rose and Diamond supra, note 17 at 516 (reviewing evidence that trial judges browbeat jurors into saying that they can be fair and will do their duty, thus likely reducing the sensitivity of their self-diagnoses). See also Maureen O’Connor, Terry Connolly, Bruce D. Sales, and John Davis, Pre-Voir Dire Instruction of the Jury Pool: A Natural Experiment, 30 CRIMINAL LAW BULLETIN 518 (1994) (showing that instructing jurors on the law prior to conducting voir dire had no effect on the jurors’ responses).


more holistic way. Attorneys may also use such a holistic approach and their limited number of peremptory challenges to exclude jurors. Our study merely suggests that, in making that holistic assessment, courts and attorneys should give no weight to the content of the juror’s own professions of impartiality or partiality.

III. DISCUSSION

The courts of appeal say that they will defer to trial court determinations as to whether a juror can be impartial, as long as those determinations are based on “substantial evidence.” This study has shown that the juror’s responses to the magic question are not substantial evidence as to their actual impartiality. These experiments show that jurors’ self-diagnoses simply do not provide the courts with a reliable basis for sorting biased jurors from unbiased jurors. Trial courts should not rely upon such unreliable answers.

The data further suggest that courts’ reliance on such unreliable self-diagnoses to seat biased jurors may be a cause of wrongful convictions and wrongful impositions of civil liability. The concern is not just with trial outcomes but also with plea bargains and civil settlements. “Indeed,” as Justice Frankfurter said, “such extraneous influences, in violation of the decencies guaranteed by our Constitution, are sometimes so powerful that an accused is forced, as a practical matter, to forego trial by jury.”

Defendants may be settling or pleading guilty out of a fear that they will be unable to get a fair trial under current doctrine. That fear appears to be warranted.

A. The Cognitive Limitations of Self-Diagnosis

Why does the Skilling method of diagnosing juror bias fail? The self-diagnosis task can be usefully analyzed as part of a more general problem that the psychology literature refers to as “mental contamination.” Wilson & Brekke, for example, define mental contamination as “the process whereby a person has an unwanted judgment, emotion, or behavior because of mental processing that is unconscious or uncontrollable.”

---

152 Skilling, 130 S. Ct. at 2924. But see Kerr et al., supra note 105 at 683 (testing judges who reviewed videotapes of potential jurors questioned about exposure to pretrial publicity, and finding no correlation between the judge’s assessments and the juror’s verdict votes).

153 See e.g., United States v. Martinez–Salazar, 528 U.S. 304, 316, 120 S.Ct. 774, 145 L.Ed.2d 792 (2000). Although this procedure may satisfy the Sixth Amendment, it is not a panacea for fairness, because the other side in litigation will be using their preemptory challenges for discretionary purposes to shape the jury affirmatively, rather than trying to remediate problems of pretrial publicity.

154 See e.g., California v. Boyette, 58 P.3d 391, 414 (Cal. 2002). See also Hannaford-Agor and Waters supra note 113 at 3 (“As a practical matter ... all U.S. jurisdictions give substantial discretion to trial judges with respect to these decisions.”)

155 Irwin, 366 U.S. at 730 (Frankfurter, J. concurring).


157 Id. at 117.
The Wilson and Brekke model outlines at least four separate actions a person must consciously undertake to debias mental contamination: a person must (1) be aware that mental contamination exists; (2) be motivated to correct the bias; (3) be aware of the direction and magnitude of the bias; and (4) be able to adjust his or her response. In the context of the present study, a potential juror must first realize that exposure to pretrial publicity has affected judgment. Second, this realization must trigger a desire to counteract the influence of pretrial publicity. Third, the juror must know the direction of the effect (either more or less likely to convict) and its magnitude (e.g., 5% or 85% shift). Finally, the juror must be able to make the necessary correction (e.g., deliberately lower the estimate of guilt by 5% to compensate for a biased 5% increase, or disqualify herself if she assesses that she will be unable to make that adjustment).

A juror responding to Skilling screening questions might fail at step (1), and not be aware that the condemning article about the defendant has increased the odds of adjudicating the defendant as guilty. A large body of psychological research demonstrates that awareness of mental contamination can be surprisingly difficult to achieve, since people are generally ignorant of the processes by which they form their judgments. Dan Simon has reviewed evidence showing that people can clearly access declarative types of knowledge with clarity, but have difficulty knowing how the arrived at that information. Simon notes that “people habitually insist on their introspective abilities, and when asked about the reasons for their decisions and behaviors, they readily provide spurious explanations.” Simon argues that people suffer from an “illusion of objectivity.”

A step (2) motivating failure would occur if the juror were indifferent about her role as a juror, or simply disagreed with the judicial instructions that pretrial publicity constitutes a bias worth eradication. After all, a juror is not a member of the legal culture, which is


\[159\text{ Alternatively, it may be possible to use other tools of modern psychology to identify bias, such as the “implicit attitudes test,” which measures reaction times of subjects presented with various stimuli, to assess whether subjects are implicitly biased for or against certain races, people, or ideas, even when the subjects are not consciously aware of such biases. See Dale K. Larson, A Fair And Implicitly Impartial Jury: An Argument For Administering The Implicit Association Test During Voir Dire, 3 DePaul Journal for Social Justice 139 (2010)(proposing such use).}\]


\[161\text{ Dan Simon, More Problems with Criminal Trials: The Limited Effectiveness of Legal Mechanisms, Law and Contemporary Problems, Vol. 75, No. 2 at 167 (2012). See also, Mark W. Bennett, Unraveling the Gordian Knot of Implicit Bias in Jury Selection: The Problems of Judge Dominated Voir Dire, the Failed Promise of Batson, and Proposed Solutions, 4 Harvard Law & Policy Review 149 (2010) (“I have come to the conclusion that present methods of addressing bias in the legal system—particularly in jury selection—which are directed primarily at explicit bias, may only worsen implicit bias. Specifically, judge dominated voir dire and the Batson challenge process are well-intentioned methods of attempting to eradicate bias from the judicial process, but they actually perpetuate legal fictions that allow implicit bias to flourish.”)}\]

\[162\text{ Id., at 186.}\]
indoctrinated through the process of law school to believe that the epistemic sterility of a trial under the Rules of Evidence is the optimal way to decide an important question. Instead, jurors may feel that they should rely on all available information, and even their feelings about the parties, just as the jurors do in the rest of their lives.

Even if aware and motivated, it is unlikely that a juror would have any insight into how much bias the pretrial exposure caused (a failure of step (3)). In experimental research, participants usually under- or over-correct, rather than reaching judgments similar to those not exposed to such information. Consider the anchoring effect, wherein a numerical judgment is biased by previous consideration of an uninformative number. In the classic study by Kahneman and Tversky, for example, subjects were asked to estimate the number of African countries in the United Nations; their answers were highly correlated with whatever number was first randomly generated from a wheel of fortune.163 Wilson et al. explicitly warned subjects that anchors would bias their judgments, provided an example, and admonished them to “please be careful not to have this contamination effect happen to you.”164 Clearly aware of the possibility of bias, and even when given incentives for accuracy, subjects rendered slightly different judgments in the direction required for correction, but not nearly enough. The differences were not even statistically significant.165

Given what we know about the limits of human cognition, it should be unsurprising that jurors fail the task of self-diagnosing their own biases. This self-diagnosis task poses to the juror a question, i.e., will the exposure cause you to change your decision from what it would have been had you not been exposed to the biasing factor? Analytically, that is a lot to ask. Indeed, it becomes doubtful whether jurors are even doing that predictive-counterfactual-comparison task, or if they are instead reporting an aspiration to fulfill their civic duty to serve and serve fairly.166

B. A Way Forward: Borrowing the Method Used for Judge Bias

The law takes a different approach for potentially biased judges than it does for potentially biased jurors. This rule for judge disqualification may be a better model for juror disqualification as well.

165 Id. at 397-398.
166 "The influence that lurks in an opinion once formed is so persistent that it unconsciously fights detachment from the mental processes of the average man... No doubt each juror was sincere when he said that he would be fair and impartial to petitioner, but the psychological impact requiring such a declaration before one’s fellows is often its father...” Irvin v. Dowd, 366 U.S. 717, 728, 81 S.Ct. 1639, 1645, 6 L.Ed.2d 751, 759 (1961).
Federal laws require that a judge "shall disqualify himself in any proceeding in which his impartiality might reasonably be questioned." If judges and jurors have similar psychologies, our experiments suggest that it may not be wise for judges to make the decision about whether to disqualify themselves. It may be more sensible to allow a different judge to resolve the threshold question, based on an objective review of the circumstances (i.e., potentially biasing factors).

Substantively however, putting aside who makes the decision, the judge-disqualification question is different than the one posed for jurors. It is not whether the judge can be impartial, but rather whether her "impartiality might reasonably be questioned." In principle, this would seem to be an objective and more stringent standard that would require disqualification more often. As the Supreme Court said in Caperton,

> The difficulties of inquiring into actual bias, and the fact that the inquiry is often a private one, simply underscore the need for objective rules. Otherwise there may be no adequate protection against a judge who simply misreads or misapprehends the real motives at work in deciding the case. The judge’s own inquiry into actual bias, then, is not one that the law can easily superintend or review.

Chief Justice Marshall recognized also this difficulty in the jury context, writing that the juror "may declare that he feels no prejudice in the case; and yet the law cautiously incapacitates him from serving on the jury because it suspects prejudice, because in general persons in a similar situation would feel prejudice."

Our experiments show that Chief Justice Marshall was correct in his skepticism about juries, just as the Caperton court was skeptical about judges hundreds of years later. Since we in fact lack a reliable method for the law to answer the question of whether a particular juror is actually biased, we should return to the more fundamental question of whether the juror’s impartiality can be reasonably questioned at all. On that basis, courts should simply remove all jurors that have been exposed to substantial pretrial publicity or other biasing factors, delaying trial or changing venue if necessary.

---

168 See Merritt v Hunter, 575 P2d 623, 624 (Okla, 1978) ("Even though a judge personally believes himself to be unprejudiced, unbiased and impartial, he should nevertheless certify his disqualification where there are circumstances of such a nature to cause doubt as to his partiality, bias or prejudice.").
169 556 U.S. at 883.
170 Burr, 25 F.Cass at 50.
171 See AMERICAN BAR ASSOCIATION, PREAMBLE, PRINCIPLES FOR JURIES AND JURY TRIALS 14 (Aug. 2005), available at www.abanet.org/juryprojectstandards/principles.pdf. ("[i]n ruling on a challenge for cause, the court should evaluate the juror's demeanor and substantive responses to questions. If the court determines that there is a reasonable doubt that the juror can be fair and impartial, then the court should excuse him or her from the trial.") (emphasis added).
Such a reform would presumably lead to some additional false positives (the elimination of jurors who could have been fair after all), as a necessary cost of reducing the false negatives that our study has shown arise from the attempt to get at actual bias. Given the constitutional demand for an impartial jury – and the need for the courts to be viewed as legitimately reaching accurate outcomes – however, this cost may be acceptable.

~ ~ ~

In sum, these findings challenge a longstanding and ubiquitous practice of the state and federal courts, used in both civil and criminal trials. Nearly 30 years ago, the Supreme Court said that, “[i]t is fair to assume that the method we have relied on since the beginning . . . usually identifies bias.”\(^{172}\) Our study undermines that assumption. Although further research is warranted, it is now fair to put the burden on those who rely upon this particular method of diagnosing bias to show that such reliance is reasonable. One can no longer simply “assume” that it is.

\(^{172}\) Patton, 467 U.S. at 1038 (citing Burr).
Appendix A. Regression Analyses

Table A1. Logistic Regression Predicting the Odds of a Plaintiff Verdict, with Nested Comparisons for National Sample Data. The overall model regressing verdict on condition was statistically significant for the jury both before (Model A1, $\chi^2 (1) = 4.738, p = .029$) and after (Model A2, $\chi^2 (1) = 5.218, p = .022$) applying the self-diagnosis screen, revealing that self-diagnosis failed to cure the bias induced via pretrial publicity. Nested comparisons revealed that adding self-diagnosis, CRT, and NFC (Model B) failed to improve the model, $\chi^2 (3) = 1.06, p = 0.787$; addition of demographic variables (Model C) likewise failed to improve the model, $\chi^2 (4) = 0.34, p = .987$. Thus the observed effects can be confidently attributed solely to the experimental manipulations.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>B</th>
<th>S.E.</th>
<th>Wald's $\chi^2$</th>
<th>$p$</th>
<th>95% CI for Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
</tr>
<tr>
<td>Model A1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>condition</td>
<td>0.69</td>
<td>0.32</td>
<td>4.63</td>
<td>.031</td>
<td>1.06</td>
</tr>
<tr>
<td>constant</td>
<td>-0.60</td>
<td>0.26</td>
<td>5.39</td>
<td>.020</td>
<td>-</td>
</tr>
<tr>
<td>Model A2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>condition</td>
<td>0.77</td>
<td>0.34</td>
<td>5.07</td>
<td>.024</td>
<td>1.11</td>
</tr>
<tr>
<td>constant</td>
<td>-0.67</td>
<td>0.28</td>
<td>5.90</td>
<td>.015</td>
<td>-</td>
</tr>
<tr>
<td>Model B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>condition</td>
<td>0.71</td>
<td>0.33</td>
<td>4.80</td>
<td>.028</td>
<td>1.08</td>
</tr>
<tr>
<td>self-diagnosis</td>
<td>-0.16</td>
<td>0.49</td>
<td>0.11</td>
<td>.743</td>
<td>0.33</td>
</tr>
<tr>
<td>CRT</td>
<td>-0.44</td>
<td>0.13</td>
<td>0.12</td>
<td>.727</td>
<td>0.75</td>
</tr>
<tr>
<td>NFC</td>
<td>0.30</td>
<td>0.32</td>
<td>0.93</td>
<td>.335</td>
<td>0.73</td>
</tr>
<tr>
<td>constant</td>
<td>-0.56</td>
<td>0.54</td>
<td>1.09</td>
<td>.300</td>
<td>-</td>
</tr>
<tr>
<td>Model C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>condition</td>
<td>0.70</td>
<td>0.33</td>
<td>4.60</td>
<td>.032</td>
<td>1.06</td>
</tr>
<tr>
<td>gender</td>
<td>-0.12</td>
<td>0.32</td>
<td>0.00</td>
<td>.970</td>
<td>0.53</td>
</tr>
<tr>
<td>age</td>
<td>-0.94</td>
<td>0.15</td>
<td>0.06</td>
<td>.801</td>
<td>0.71</td>
</tr>
<tr>
<td>race</td>
<td>0.16</td>
<td>0.38</td>
<td>0.19</td>
<td>.664</td>
<td>0.56</td>
</tr>
<tr>
<td>education</td>
<td>-0.02</td>
<td>0.15</td>
<td>0.01</td>
<td>.918</td>
<td>0.74</td>
</tr>
<tr>
<td>constant</td>
<td>-0.49</td>
<td>0.61</td>
<td>0.66</td>
<td>.417</td>
<td>-</td>
</tr>
</tbody>
</table>
Table A2. Logistic Regression Predicting the Odds of a Self-Diagnosis for National Sample Data. The overall model regressing self-diagnosis on condition, CRT, and NFC was not statistically significant, $\chi^2 (3) = 4.806, p = .187$, revealing that jurors equally likely (or rather unlikely) to self-diagnosis bias regardless of exposure to pretrial publicity or psychological constructs related to self-reflection and tendency to think.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>B</th>
<th>S.E.</th>
<th>Wald’s $\chi^2$</th>
<th>$p$</th>
<th>95% CI for Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>condition</td>
<td>-0.40</td>
<td>0.52</td>
<td>0.58</td>
<td>.45</td>
<td>0.24</td>
</tr>
<tr>
<td>CRT</td>
<td>0.29</td>
<td>0.20</td>
<td>2.07</td>
<td>.15</td>
<td>0.90</td>
</tr>
<tr>
<td>NFC</td>
<td>0.63</td>
<td>0.50</td>
<td>1.58</td>
<td>.21</td>
<td>0.70</td>
</tr>
<tr>
<td>constant</td>
<td>-0.49</td>
<td>0.61</td>
<td>0.66</td>
<td>.417</td>
<td>-</td>
</tr>
</tbody>
</table>