117 COLUMBIA L. REV. ONLINE __ (forthcoming, 2017)

TESTING THE IMPACT OF CRIMINAL JURY INSTRUCTIONS ON VERDICTS: A CONCEPTUAL REPLICATION

Michael D. Cicchini*
Lawrence T. White**

The Constitution protects us from criminal conviction unless the state can prove guilt beyond a reasonable doubt. However, many of our nation's trial courts will conclude their burden of proof instructions by telling jurors not to evaluate the evidence for doubt, but instead "to search for the truth" of what they think really happened.

In our previously published study, we empirically demonstrated that such truth-related language lowers the state's burden of proof below the constitutionally-guaranteed reasonable doubt standard. In this article, we discuss the results of our new empirical study—a conceptual replication and extension of our previous work.

In our new study, we again found a statistically significant difference in conviction rates between mock jurors who were properly instructed on reasonable doubt, and mock jurors who were instead instructed "to search for the truth." Additionally, we identified a cognitive link between the truth-related jury instruction and the increased conviction rate.

More specifically, mock jurors who were instructed "to search for the truth" were nearly twice as likely to mistakenly believe they could convict the defendant even if they had a reasonable doubt about his guilt. Further, mock jurors who held this mistaken belief actually voted to convict the defendant at a rate 2.5 times higher than those who correctly understood the burden of proof.

Our original study, our successful replication, and our newly discovered cognitive explanation for juror behavior combine to provide powerful evidence that truth-related language diminishes the constitutionally-mandated burden of proof. Therefore, in order to protect our Due Process rights, courts should immediately remove such truth-related language from their burden of proof jury instructions.

* Criminal Defense Lawyer, Cicchini Law Office, LLC, Kenosha, Wisconsin. J.D., *summa cum laude*, Marquette University Law School (1999); C.P.A., University of Illinois Board of Examiners (1997); M.B.A., Marquette University Graduate School (1994); B.S., University of Wisconsin—Parkside (1990).

** Professor and Chair of Psychology, Beloit College; Director, Beloit College's Law & Justice Program. Ph.D., University of California, Santa Cruz (1984); M.A., *with distinction*, California State University at Fresno (1979); B.A., *with honors*, Whittier College (1975).

INTRODUCTION

The Constitution protects us from criminal conviction unless the state can prove guilt beyond a reasonable doubt. However, after first defining reasonable doubt, many trial courts will then instruct jurors "not to search for doubt," but instead "to search for the truth" of what they think really happened. Defendants have argued that such truth-related language reduces the state's burden of proof to a mere preponderance of the evidence. That is, if the jury were to find the state's case only slightly more convincing than the defendant's, it would follow that, in a search for the truth, the jury would be obligated to convict.

Appellate courts, however, universally reject this argument. Most appellate courts acknowledge that such truth-related language is inaccurate, is highly-disfavored, and could *in theory* lower the state's burden of proof. However, these courts then go on to conclude, without any empirical support, that such language probably does not cause any *actual* harm.

In our previous study and article, "Truth or Doubt? An Empirical Test of Criminal Jury Instructions," we put this judicial reasoning to the test. In a hypothetical criminal case, we found that mock jurors who were properly instructed on reasonable doubt (N = 100) convicted the defendant at the rate of 16%. However, mock jurors who received the identical case information and instruction, but were also told "not to search for doubt" but instead "to search for the truth" (N = 100), convicted at the much higher rate of 29%.

In this article, we discuss the results of our new study wherein we first attempted a conceptual replication of our previous work, and then attempted to identify a cognitive explanation for why truth-related language produces a higher conviction rate. We recruited 248 participants to serve as mock jurors in a hypothetical, fourth-degree sexual assault case. After reading a case summary, jurors were randomly assigned to one of two groups, each of which received a different instruction on the state's burden of proof. Group 1 was properly instructed on reasonable doubt (N = 124), and convicted at the rate of 22.6%. Group 2 received the same instruction, but was then told "not to search for doubt" but instead "to search for the truth" (N = 124), and convicted at the rate of 33.1%. This *nearly 50% increase* in conviction rates is statistically significant.

In our new study, we also asked jurors a post-verdict question about their subjective understanding of the burden of proof. We found that jurors who were first instructed on reasonable doubt, but then told "not to search for doubt" but instead "to search for the truth," were *nearly twice as likely* to believe they could convict the defendant even if they had a reasonable doubt about his guilt. Even more significant, jurors who held this mistaken belief (regardless of the group to which they were

randomly assigned) actually convicted at a rate *2.5 times* that of jurors who correctly understood the burden of proof.

Part I of this article explains the burden of proof in criminal cases, and examines the truth-related language that trial courts commonly tack-on to the end of their reasonable doubt jury instructions. Part II explains our previous study, including our study design and our statistical findings. Part III, the heart of this article, explains our new study—a conceptual replication and extension of our previous work. In this part we outline our study objectives, formally state our hypotheses, discuss our study design, and explain our statistical findings.

Part IV then explains the significance of our findings for trial judges, jury instruction committees, and appellate courts. We also discuss the cognitive link between jury instructions and conviction rates, i.e., truth-related language causes jurors to misunderstand the state's burden of proof, which in turn causes them to convict even when they have a reasonable doubt about guilt. Based on our successful replication and this new finding, we reiterate our argument from our previous article: In order to protect our Due Process rights, courts should immediately terminate their use of truth-based jury instructions.

Finally, Part V discusses the study limitations that we have corrected by virtue of this conceptual replication, as well as the study limitations that still exist, but could be addressed by other researchers in future studies.

I. THE BURDEN OF PROOF IN CRIMINAL CASES

In 1970, the Supreme Court of the United States explicitly held that "the Due Process Clause protects the accused against conviction except upon proof beyond a reasonable doubt."[1] However, trial courts are given tremendous leeway in how they instruct jurors on this burden of proof.[2] And in so doing, many trial courts will conclude their instructions *not* by telling jurors to examine the state's evidence for reasonable doubt, but instead by telling them to decide the *truth* of what they think really happened. In our previous article we provided examples, from thirteen jurisdictions, including these:[3]

> After giving an otherwise legally proper instruction on proof beyond a reasonable doubt, many courts will then instruct jurors that, when reaching their verdict, they should "[d]etermine what you think the truth of the matter is and act accordingly."[4] Similarly, other courts instruct jurors that, when reaching their verdict, they should "evolve the truth,"[5] "seek the truth,"[6] "search for truth,"[7] or "find the truth."[8] Some courts—again, after

[1] In re Winship, 397 U.S. 358, 364.
[2] *See* Victor v. Nebraska, 511 U.S. 1, 5 (1994).
[3] Michael D. Cicchini & Lawrence T. White, *Truth or Doubt? An Empirical Test of Criminal Jury Instructions*, 50 U. RICH. L. REV. 1139, 1143 (2016).
[4] State v. Dunkel, 466 N.W.2d 425, 430 (Minn. Ct. App. 1991).
[5] United States v. Pine, 609 F.2d 106, 108 (3d Cir. 1979).
[6] United States v. Gonzalez-Balderas, 11 F.3d 1218, 1223 (5th Cir. 1994).
[7] Commonwealth v. Allard, 711 N.E.2d 156, 159 (Mass. 1999).
[8] United States v. Gray, 958 F.2d 9, 13 (1st Cir. 1992).

> properly instructing jurors on the concept of reasonable doubt—will explicitly contradict themselves by further instructing jurors that "you should *not* search for doubt. You should search for the *truth*."[9]

Defendants have frequently challenged such truth-related language on appeal. One defense argument is that instructing the jury to determine, evolve, seek, find, or search for the truth of what they think happened diminishes the state's burden of proof. That is, "'seeking the truth' suggests determining whose version of events is more likely true, the government's or the defendant's, and thereby intimates a *preponderance of evidence* standard."[10] More to the point, "truth is not the jury's job."[11] Rather, "The question for any jury is whether the burden of proof has been carried by the party who bears it. In a criminal case . . . [t]he jury cannot discern whether that has occurred without examining the evidence for reasonable doubt."[12]

Our nation's appellate courts, however, have consistently rejected this argument. With only slight variations in their reasoning, courts conclude that while such truth-related language is disfavored—and could, in theory, diminish the state's burden of proof—it probably does no actual harm.[13] This, in turn, allows appellate courts to affirm defendants' convictions, and allows trial courts to continue to instruct juries to search for the truth of what they think really happened, rather than to examine the state's evidence for reasonable doubt.

Based on the plain language of these truth-related jury instructions, we believed the courts' thinking was quite obviously flawed; therefore, we decided to put their judicial reasoning to an empirical test.

II. THE ORIGINAL STUDY: TRUTH OR DOUBT?

In our previous study and article we measured the conviction rates in a hypothetical criminal case.[14] We recruited 200 study participants[15] through Amazon's Mechanical Turk for the purpose of testing the following hypothesis: "[W]hen truth-related language is added to an otherwise proper beyond a reasonable doubt instruction, the truth-related language not only contradicts but also *diminishes* the government's burden of proof."[16]

To test this hypothesis, each study participant served as a mock juror and received the same case summary materials. More specifically:

---

[9] State v. Avila, 532 N.W.2d 423, 429 (Wis. 1995) (emphasis added) (rev'd, in part, on other grounds).
[10] *Gonzalez-Balderas*, 11 F.3d at 1223 (emphasis added).
[11] State v. Berube, 286 P.3d 402, 411 (Wash. Ct. App. 2012).
[12] *Id.*
[13] Cicchini & White, *supra* note 3, at 1158-59, n. 60-63.
[14] *Id.* at 1155-56.
[15] Our previous study actually consisted of 300 participants; this number was then reduced to 298 mock jurors after excluding participants that were not Untied States citizens. However, 98 of the mock jurors were placed into a third test condition designed to test a separate, preliminary hypothesis, which we do not rehash in this article. Therefore, we simply describe our previous study as having 200 study participants.
[16] Cicchini & White, *supra* note 3, at 1150.

> Every mock juror read the same fact pattern in a hypothetical case of sexual assault of a child. The defendant in the case was alleged to have touched a fifteen-year-old child's buttocks, over the clothing, for purposes of sexual arousal or gratification. The case summary began with an instruction on the charged crime, including its elements, followed by a 625-word synopsis of court testimony from three individuals: the alleged child victim, the child's mother, and the defendant. The child's accusation was not corroborated by an eyewitness or physical evidence. In essence, the case consisted, as most real-life sexual touching cases do, of an allegation and a denial. The case summary concluded with an 850-word transcript of the prosecutor's and defense lawyer's closing arguments, each arguing the points most favorable to their case.[17]

Before being asked to render a verdict, these 200 mock jurors were randomly assigned to one of two groups, each of which received a *different* instruction on the state's burden of proof. Jurors in the doubt-only group (N = 100) received a legally proper, 269-word burden of proof instruction that concluded as follows: "It is your duty to give the defendant the benefit of every reasonable doubt."[18]

Jurors in the doubt-and-truth group (N = 100) received the same instruction except that the conclusion was changed to read as follows: "While it is your duty to give the defendant the benefit of every reasonable doubt, you are not to search for doubt. You are to search for the truth."[19]

The doubt-only group (that received the legally proper instruction) convicted at a rate of only 16%. However, the doubt-and-truth group (that was told "not to search for doubt" but instead "to search for the truth") convicted at the much higher rate of 29%.[20] More specifically:

> This result is significant at the $p < .05$ level, with an exact *p*-value of 0.028. . . . [T]he *p*-value measures the probability of a Type I error, i.e., obtaining a false positive. Therefore, we are more than 97% certain ($1–p$) that the difference in conviction rates between [the groups] is a real difference and did not occur by chance.
>
> This finding provides strong empirical support for our . . . hypothesis that the truth-related language at the end of an otherwise proper reasonable doubt instruction actually diminishes the government's burden of proof.[21]

We concluded that "[b]ecause 'the Due Process Clause protects the accused against conviction except upon proof beyond a reasonable doubt,' our findings provide strong evidence of a serious constitutional problem."[22]

---

[17] *Id*. at 1151.
[18] *Id*. at 1152-53.
[19] *Id*. at 1153-54.
[20] *Id*. at 1155.
[21] *Id*.
[22] *Id*. at 1157 (quoting In Re Winship, 397 U.S. 358, 364 (1970)).

III. THE NEW STUDY: A CONCEPTUAL REPLICATION

A. *Objectives*

First, our main objective is to test the reliability of our previous finding by replicating the study. In order to do this, we designed and conducted a conceptual replication, rather than a direct replication. A conceptual replication retests the original hypothesis, but intentionally varies specific features of the original methodology.[23] A benefit of conceptual replication is that it addresses one of the common weaknesses of psychological research, including our original study: limited generalizability.

More specifically, in our original study, each of the 200 mock jurors read the same case summary about a child sexual assault allegation. All of the material was held constant between the two test conditions. This consistency allowed us to isolate the effect of the variable being tested: the closing mandate "not to search for doubt" but instead "to search for the truth." However, with such consistency comes limited generalizability. As we cautioned, "we cannot say that the impact of the doubt-and-truth instruction would be identical when applied to different cases."[24]

A conceptual replication allows us to address this limitation by testing our hypothesis under a different set of circumstances. As discussed below in part III.C., and again in part V., our new study has a larger sample size, a different fact pattern, and includes stronger evidence of the defendant's guilt. We also provided mock jurors with a different, and much shorter, underlying instruction on reasonable doubt. However, the variable being tested—the mandate "not to search for doubt" but instead "to search for the truth"—is the identical language that we tested in our previous study.

Second, in addition to replicating our study, we also extend our study so as to identify a cognitive link between the change in the test conditions (i.e., adding truth-related language to one group's burden of proof instruction) and the change in juror behavior (i.e., a higher conviction rate). In order to accomplish this secondary objective, we added an additional, post-verdict question to our test materials. We discuss this, and our overall study design, in part III.C. First, we will formally state our hypotheses.

B. *Hypotheses*

Our first hypothesis is that when truth-related language is added to an otherwise proper reasonable doubt instruction, the truth-related language will diminish the state's burden of proof, i.e., mock jurors will convict at a higher rate. This is the identical hypothesis we tested in our original study.

Our second hypothesis is that mock jurors who receive the truth-related language at the end of their reasonable doubt instruction will *subjectively* interpret their instruction to permit conviction even if they have a reasonable doubt about the defendant's guilt.

C. *Study Design*

[23] *See, e.g.*, Stefan Schmidt, *Shall We Really Do It Again? The Powerful Concept of Replication is Neglected in the Social Sciences*, 13 REV. OF GEN. PSYCH. 90 (2009).
[24] Cicchini & White, *supra* note 3, at 1162.

To test these hypotheses we recruited 250 study participants—a 25 percent increase in the sample size of our original study—through Amazon's Mechanical Turk. As we explained in our earlier research:[25]

> Mechanical Turk has many advantages, including "easy access to a large, stable, and diverse subject pool, the low cost of doing experiments, and faster iteration between developing theory and executing experiments."[26] Further, several studies have found a high degree of similarity between the judgments and behaviors of Mechanical Turk "workers" and of participants recruited in more conventional ways, such as through university subject pools.[27]

These 250 participants served as mock jurors and rendered a verdict in a hypothetical criminal case. To ensure data quality, we monitored the participants and immediately rejected those who completed the task in less than three minutes; we replaced them with new participants in order to maintain our desired sample size. Each participant was required to be an adult and a United States citizen. After data collection was completed, we discovered that one participant was not a U.S. citizen and one failed to render a verdict; their data were discarded, leaving us with a sample of 248 mock jurors.

Our sample was large and diverse. Participants hailed from 42 different states. Fifty-two percent of participants were female. Participants' ages ranged from 19 years to 73 years; the mean (average) age was 35.8 years and the median age (50th percentile) was 32 years. The ethnic composition of the sample was also diverse: 74% non-Hispanic whites, 10% African-Americans, 5% Hispanics, 5% Asian-Americans, 5% mixed race, and 1% other. Fifty-six percent of the participants reported at least a four-year college degree, while an additional 35% have completed some college. Thirteen percent reported having prior jury experience.

Every mock juror read the same fact pattern that involved two adults interacting at a party, and concluded with an accusation of a misdemeanor fourth-degree sexual assault, i.e., the defendant's sexual touching of the alleged victim without her consent. The case summary began with an instruction on the charged crime, including its elements, followed by an 887-word summary of the trial evidence. The evidence consisted of testimony from two witnesses—the accuser and the defendant—and a factual stipulation entered into between the prosecutor and defense lawyer.[28]

There were no eyewitnesses to the alleged sexual assault. The accuser immediately reported the incident to law enforcement. The defendant denied the allegation. Both the accuser and the defendant testified and admitted to consuming

---

[25] *Id*. at 1150-51.

[26] Winter Mason & Siddharth Suri, *Conducting Behavioral Research on Amazon's Mechanical Turk*, 44 BEHAV. RES. 1, 1–2 (2012).

[27] *Id*. at 3-4.

[28] We included a factual stipulation for two reasons. First, it allowed us to shorten the summaries of the witnesses' testimony by removing "identity" as an issue in the case. And second, the data used in this study was obtained as part of a larger data collection effort that included a third group. The inclusion of the factual stipulation allowed us test an additional (third) hypothesis that is *not* related to this study, but may form the basis for a future article.

alcohol during the party at which the sexual assault allegedly occurred. The defendant, however, also admitted to consuming other drugs earlier in the day, and admitted to a prior, unrelated instance of untruthful conduct. In order to shorten the overall length of the case summary materials, we did not include closing arguments from the lawyers. We did, however, instruct the jury that the definition of "evidence" includes the testimony of witnesses as well as the factual stipulation.

Before being asked to render a verdict of guilty or not guilty, the 248 mock jurors were randomly assigned to one of two test conditions, each of which received a different jury instruction on the state's burden of proof. Group 1 (N = 124) received a legally proper, 94-word jury instruction that explained the presumption of innocence, placed the burden of proof on the state, and identified the burden of proof as beyond a reasonable doubt. This doubt-only instruction, in its entirety, reads as follows:

> The defendant is presumed to be innocent of the charge. This presumption continues during every stage of the trial and your deliberations on the verdict. It is not overcome unless from all the evidence in the case you are convinced beyond a reasonable doubt that the defendant is guilty as charged. *The government has the burden of proving the guilt of the defendant beyond a reasonable doubt.*
>
> This burden of proof stays with the government throughout the case. The defendant is never required to prove his innocence or to produce any evidence at all.[29]

Group 2 (N = 124) received an identical jury instruction, with one exception. The instruction given to Group 2 concluded with this additional mandate: "While it is your duty to give the defendant the benefit of every reasonable doubt, you are not to search for doubt. You are to search for the truth." This doubt-and-truth instruction, in its entirety, reads as follows:

> The defendant is presumed to be innocent of the charge. This presumption continues during every stage of the trial and your deliberations on the verdict. It is not overcome unless from all the evidence in the case you are convinced beyond a reasonable doubt that the defendant is guilty as charged. The government has the burden of proving the guilt of the defendant beyond a reasonable doubt.
>
> This burden of proof stays with the government throughout the case. The defendant is never required to prove his innocence or to produce any evidence at all.
>
> *While it is your duty to give the defendant the benefit of every reasonable doubt, you are not to search for doubt. You are to search for the truth.*[30]

---

[29] CRIM. JURY INSTRUCTIONS No. 2.03 (7th Cir.) (1998).

[30] This jury instruction includes the Seventh Circuit instruction in its entirety. *Id*. The additional, truth-related language added to the end of this instruction is the identical language tested in our original study, and is taken from Wisconsin's pattern jury instruction on the burden of proof. WIS. CRIM. JURY INSTRUCTIONS No. 140 (2015). This language is similar, and often identical, to the truth-related language

After rendering their verdict, mock jurors were asked to answer a question about how they interpreted their burden of proof instruction. The question posed to all jurors was as follows: "You received an instruction from the judge explaining the prosecutor's burden of proof. Which of the following do you believe is the *most accurate summary* of the judge's instruction?" Jurors were instructed to "choose only *one* answer." Their options were as follows:

> A. If I have a reasonable doubt about the defendant's guilt, I must not convict the defendant.
>
> B. Even if I have a reasonable doubt about the defendant's guilt, I may still convict the defendant if, in my search for the truth, the evidence shows the defendant is guilty.

The study was approved by Beloit College's Institutional Review Board. We discuss results below.

D. *Findings*

To test our first hypothesis—that adding truth-related language to the end of an otherwise proper reasonable doubt instruction will diminish the government's burden of proof—we must compare the conviction rates of Groups 1 and 2.

In Group 1, which received the doubt-only instruction, only 28 of 124 mock jurors returned verdicts of guilt for a group conviction rate of 22.6%. In Group 2, which received the doubt-and-truth instruction, 41 of 124 mock jurors returned verdicts of guilt for a group conviction rate of 33.1%. That is, the conviction rate among jurors who were told "not to search for doubt" but instead "to search for the truth" was almost 50% higher than the conviction rate for jurors who were simply instructed to evaluate the state's case for reasonable doubt.

This result is significant at the $p < .05$ level, with an exact p-value of 0.033. The p-value measures the probability of a Type I error, i.e., the risk of obtaining a false positive when testing a hypothesis, given the two sample sizes and the difference in conviction rates between the two groups. In plain language, we are more than 96% certain $(1 - p)$ that the observed difference in conviction rates between Groups 1 and 2 is a real difference and did not occur by chance.[31]

After mock jurors rendered their verdict, they were asked to report how certain they were (on a 10-point scale) that they had made a correct decision. There were no statistically significant differences in levels of certainty between the doubt-only group

---

used in the thirteen different jurisdictions we identified in our original study and article. Cicchini & White, *supra* note 3, at 1143, n. 13-18.

[31] A statistical test for the difference between two proportions produced a *Z*-score of -1.84. In a one-tailed test, the *p*-value is 0.033. Researchers use a two-tailed test (also called a two-sided test) when they cannot predict if a test variable will increase or decrease scores. We used a one-tailed test because we had empirical evidence (from our first study) that truth-instructed jurors would convict at a higher rate, not a lower rate.

and the doubt-and-truth group. In fact, both group means were essentially 6.6 (fairly certain) on the 10-point scale.[32]

Participants also answered an attention-check question that tested their recollection of the elements of the charged crime. The question included five potential elements, only three of which were correct. The attention-check results were encouraging. Nearly 92% of participants correctly identified the elements of the charged crime.[33]

To test our second hypothesis—that mock jurors receiving the doubt-and-truth instruction would *subjectively* interpret it to permit conviction even if they had a reasonable doubt about the defendant's guilt—participants were reminded that they had received an instruction from the judge about the state's burden of proof. Participants were then asked to indicate the most accurate summary of the judge's instruction by choosing either answer A or B. As indicated above, A is the correct interpretation of the constitutionally-mandated burden of proof; B is the incorrect interpretation, as it permits conviction even when there is a reasonable doubt about the defendant's guilt.

In Group 1, which received the doubt-only instruction, only 15% of participants selected answer B; that is, only 15% believed they could convict the defendant if they had a reasonable doubt about his guilt. However, in Group 2, which received the doubt-and-truth instruction, 28% selected answer B; that is, 28% believed they could convict the defendant even if they had a reasonable doubt about his guilt. This difference is highly significant ($p = 0.01$).

Perhaps more importantly, when analyzing the responses across *both* groups 1 and 2, a juror's understanding of the burden of proof instruction was an incredibly strong predictor of his or her verdict. Of those participants who selected the legally-correct answer A—that they could *not* convict if they had a reasonable doubt about the defendant's guilt—only 21% voted guilty. Of those who selected the legally-incorrect answer B—that they *could* convict despite their reasonable doubt about the defendant's guilt—54% voted guilty. This difference is highly significant ($p < .001$).

IV. DISCUSSION: AN EVEN STRONGER CASE AGAINST TRUTH

Appellate courts consistently reject defendants' challenges to truth-related jury instructions. The courts simply decide, without any empirical evidence, that even though truth-related language is improper, disfavored, and could *in theory* lower the burden of

[32] We also uncovered several subsidiary findings not directly related to the main purpose of our study: (a) women (34%) were more likely than men (22%) to vote guilty ($p < .04$); (b) there were no statistically significant relationships between a participant's verdict and his or her age, education, ethnicity, or prior jury experience; and (c) mock jurors who voted guilty were significantly more certain than other jurors were that they had made the correct decision (a mean score of 7.5 versus a mean score of 6.3 on a 10-point scale, $p < .001$).

[33] Our standard for a correct answer was high; a mock juror who identified the correct elements of the charged crime, but also an incorrect element, was classified as "incorrect." Those mock jurors who voted not guilty were correct 93% of the time, while those who voted guilty were correct 87% of the time. This difference is not large enough to be statistically significant, but it suggests that those mock jurors who paid closer attention to the legal elements of the charge (fourth degree sexual assault) were less likely to convict.

proof, it probably causes no *actual* harm.[34] Our findings in this study, however, provide strong empirical evidence that this judicial reasoning is badly flawed.

In this study, our first finding confirms our hypothesis that adding truth-related language to the end of an otherwise proper reasonable doubt instruction diminishes the state's burden of proof. That is, the jurors in Group 1, who were instructed simply to evaluate the state's evidence for reasonable doubt, convicted at a rate of 22.6%. However, the jurors in Group 2, who were instructed "not to search for doubt" but instead "to search for the truth," convicted at a rate of 33.1%—a conviction rate *nearly 50% higher* than Group 1's rate. This replicates the finding in our original study, which also revealed a statistically significant gap in conviction rates when testing the identical hypothesis.[35]

Our second finding in this study is, in some ways, even more compelling. We hypothesized that jurors who received the doubt-and-truth instruction would be more likely to subjectively interpret the burden of proof to permit conviction even if they had a reasonable doubt about the defendant's guilt. What we found was that in Group 1 (doubt-only), only 15% of jurors believed they could convict the defendant if they had a reasonable doubt about his guilt. However, in Group 2 (doubt-and-truth), 28%—*nearly double*—believed they could convict the defendant even if they had a reasonable doubt about his guilt.

Even more striking, when analyzing the responses of all participants across groups, jurors who mistakenly believed they could convict, even if they had a reasonable doubt about guilt, found the defendant guilty 54% of the time. This conviction rate is *more than 2.5 times* the conviction rate (21%) of jurors who correctly understood the burden of proof—a highly significant difference.

These findings suggest that we have identified the cognitive mechanism explaining why the "search for the truth" language produces a much higher conviction rate. Specifically, the truth instruction (TI) produces in jurors a mistaken belief (B) about the legally-mandated burden of proof, and jurors base their verdicts (V) on that mistaken belief.

That is, in our original study we demonstrated the impact of the truth-related jury instruction on jurors' conviction rates, but we did not attempt to explain why, in a cognitive sense, the truth-related language led so many jurors to find the defendant guilty. In this study, however, we have demonstrated empirically that TI → B and B → V. The mistaken belief B is a cognitive variable that mediates (or explains) the impact of TI on V. In plain language, telling jurors "to search for the truth" leads them to form an incorrect understanding of the state's burden of proof. This misunderstanding, in turn, leads many jurors to vote guilty, even when the state has not met its burden.

Because "the Due Process Clause protects the accused against conviction except upon proof beyond a reasonable doubt,"[36] courts that tack-on truth-related language to their burden of proof instructions are creating a serious constitutional problem. Therefore, as we proposed in our original article:

---

[34] Cicchini & White, *supra* note 3, at 1158-59.

[35] *Id.* at 1155 (finding conviction rates of 16% in the doubt-only group and 29% in the doubt-and-truth group, with a *p*-value of 0.028).

[36] In re Winship, 397 U.S. 358, 364 (1970).

> [I]ndividual trial judges should exercise their broad discretion and simply eliminate truth-related language from their burden of proof instruction. A larger, more comprehensive solution to the problem is equally simple: in states where truth-related language is part of a state-wide model jury instruction, the state's jury-instruction committee should modify its instruction accordingly. Alternatively, if the committee fails to do so, the state's supreme court can use its supervisory powers to eliminate the "search for the truth" and similar language from its burden of proof instruction.[37]

Further, such a simple change should not be controversial in any imaginable way. As we explained in our original article:

> Of the numerous court decisions cited in this article, none contend that truth-related language is necessary, that the government is entitled to it, or that it offers any benefit whatsoever. Rather, appellate courts concede that the language is erroneous, and merely tolerate the language after the fact of conviction because, they conclude, it probably did no harm. . . . These lukewarm justifications did not warrant keeping the truth-related language in the first place. Instead, such language should have been eliminated long ago, as the best thing that could ever be said about it is that it "presents a risk without any real benefit." Since our study's empirical evidence demonstrates that the "risk" is actually a reality—i.e., truth-related language diminishes and even eviscerates the government's burden of proof—such language has no place whatsoever in criminal jury instructions.[38]

And now, our second study has replicated the result of our first study; further, we have also identified a cognitive mechanism that serves as a bridge or link between the legally-defective, truth-based instruction and the jurors' higher conviction rate. This makes an even more compelling case for the removal of truth-related language from burden of proof jury instructions.

V. STUDY LIMITATIONS AND FURTHER TESTING

In our original study and article, we identified five potential limitations that researchers may wish to address in future studies.[39] In this new study, we have addressed two of those five ourselves.

First, by conducting a conceptual replication, rather than a direct replication, we have expanded the generalizability of our findings. More specifically, in our original article we cautioned that "we cannot know the *extent* to which this effect will also be observed in other cases with different fact patterns."[40] Therefore, in this new study we

[37] Cicchini & White, *supra* note 3, at 1158 (internal citation omitted).
[38] *Id.* at 1158-59 (internal citations omitted).
[39] *Id.* at 1159-1165.
[40] *Id.* at 1161.

changed the fact pattern. For example, instead of a delayed report by a child accuser, we used an immediate report by an adult accuser. We also incorporated more evidence of guilt than in our first study, including the defendant's drug use on the day of the incident, and the defendant's prior, unrelated instance of untruthful conduct. Both of these pieces of evidence tend to diminish the credibility of the defendant's testimony.

In addition to changing the fact pattern, we also changed other parts of the case summary to further expand the generalizability of our findings. For example, we eliminated closing arguments of the lawyers on both sides, and even changed the underlying jury instructions. We added an instruction telling jurors that "evidence" includes the testimony of witnesses, which was designed to correct any misconception that physical evidence is required in order to convict. We also changed the underlying jury instruction on reasonable doubt. Instead of the lengthy, 269-word doubt-only instruction from our original study, we used a much shorter, 94-word doubt-only instruction.[41] What remained unchanged from our original study, however, was the closing mandate (for one of the two groups) that "[w]hile it is your duty to give the defendant the benefit of every reasonable doubt, you are not to search for doubt. You are to search for the truth."

Second, we also corrected the problem of some participants' inadequate attention level. In our previous study, we did not reject any study participants, and included their data regardless of the amount of time they spent on the task. In this new study, however, we monitored the incoming data and rejected the work of study participants who spent fewer than three minutes on the task. These participants were replaced before we concluded the data collection process. Our attempt to ensure quality responses was apparently successful, as reflected in the results of our attention-check question. In this new study, nearly 92 percent of mock jurors were able to correctly identify the legal elements of the charged crime. In our previous study, less than 84 percent of mock jurors correctly identified the legal elements.

This leaves three potential study limitations for other researches to address in the future: our use of the case summary method; the lack of juror deliberations; and participant bias. As we explained in our original study, however, the case summary method may actually be the best method for testing the impact of a jury instruction as it eliminates extraneous variables, such as witnesses' race and ethnicity, from the equation.[42] And with regard to participant bias, this problem mirrors the problem with real-life juries and, for purposes of controlled studies like ours, is mitigated by the random assignment of participants to test conditions.[43]

CONCLUSION

In our previous study and article we demonstrated that mock jurors who were first instructed on reasonable doubt, but then told "not to search for doubt" but instead "to

[41] As explained earlier in this article, states are given tremendous leeway when instructing juries on reasonable doubt. And while the two doubt-only instructions used in our two studies are dramatically different in length and content, both have been held to be legally proper.
[42] Cicchini & White, *supra* note 3, at 1160-61.
[43] *Id*. at 1164-65.

search for the truth," convicted at a much higher rate than mock jurors who received a legally-proper reasonable doubt instruction.

In this new study—a conceptual replication and extension of our previous work—we substantially changed the case summary materials before randomly assigning mock jurors to the test groups. Jurors assigned to Group 1 received a legally proper reasonable doubt instruction (doubt-only) and convicted at the rate of 22.6%. Jurors assigned to Group 2 received the same reasonable doubt instruction, but were then told "not to search for doubt" but instead "to search for the truth" (doubt-and-truth). Jurors in this group convicted at the rate of 33.1%. This *nearly 50% increase* in the conviction rate is statistically significant, and replicates the major finding in our original study.

In addition to replicating our original study, we also attempted to provide a cognitive explanation for the difference in conviction rates between Group 1 and Group 2 in our new study. In order to do this, we asked mock jurors a post-verdict question about their subjective understanding of the burden of proof instruction. Of the jurors in Group 1 (doubt-only), only 15% believed they could convict the defendant if they had a reasonable doubt about his guilt. However, of the jurors in Group 2 (doubt-and-truth), 28% held this mistaken belief. This *near doubling* was statistically significant.

Even more noteworthy, however, was that a mock juror's subjective interpretation of the burden of proof instruction was a strong predictor of his or her verdict. When analyzing verdicts across groups, of those jurors who correctly believed they could convict only if the state proved guilt beyond a reasonable doubt, only 21% voted guilty. Conversely, of those jurors who mistakenly believed they could convict even if they had a reasonable doubt about guilt, 54% voted guilty. Stated differently, jurors who held this mistaken belief convicted at a rate that was *2.5 times* that of jurors who properly understood the burden of proof.

This finding was highly significant (p-value < 0.001), and identifies the cognitive link between the truth-related jury instruction and the higher conviction rate. That is, the truth-related jury instruction creates a mistaken belief in jurors about the state's burden of proof; this mistaken belief, in turn, leads jurors to convict defendants even when the state has not met its constitutionally-imposed burden.

Our original study, our successful replication of that study, and our new empirical finding regarding the cognitive explanation for juror behavior, all combine to provide powerful evidence that truth-related language in jury instructions diminishes the state's burden of proof. And because the Constitution protects us from conviction unless the state can prove guilt beyond a reasonable doubt, we have also provided powerful evidence of a serious constitutional problem in states that permit such truth-related language in their jury instructions.

The solution to this problem is incredibly simple. In light of our empirical findings, courts can no longer justify the continued use of truth-related language under the theory that it probably does not cause any actual harm. Consequently, all truth-related language should be immediately removed from burden of proof jury instructions. Instead, these jury instructions should simply conclude, without qualification, that the state has the burden of proof beyond a reasonable doubt.