TRUTH OR DOUBT? AN EMPIRICAL TEST OF CRIMINAL JURY INSTRUCTIONS

Michael D. Cicchini *
Lawrence T. White **

INTRODUCTION

The Constitution protects a criminal defendant from conviction unless the government can prove guilt beyond a reasonable doubt. However, the Constitution does not require that trial courts use any particular set of words when defining reasonable doubt for the jury. Instead, a broad range of jury instructions have been deemed constitutionally acceptable, provided they do not diminish or dilute the government's high burden of proof.

Many defendants have challenged the constitutionality of their burden of proof instructions based on a seemingly innocuous word: truth. That is, after properly instructing juries on reasonable doubt, many trial courts conclude by instructing the jury not to search for doubt, but instead to evolve, seek, or search for the truth of what they think really happened.

This and similar truth-related language poses several constitutional problems. Most significantly, an instruction to evolve, seek, or search for the truth may lower the government's burden of proof from beyond a reasonable doubt to a mere preponderance of the evidence. That is, if a jury feels the government's version of events is slightly more likely than the defendant's version to be true, it would follow that, in a search for the truth, the jury would be obligated to convict the defendant.

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

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

When defendants challenge the constitutionality of these truth-related jury instructions, appellate courts nearly always affirm their convictions.[1] Many of the courts concede that the truth-related language adds nothing of value, is inaccurate, and could, in theory, dilute the government's burden of proof. Nonetheless, these same courts go on to conclude, without any empirical support, that when faced with these two competing instructions—one to evaluate evidence for reasonable doubt and another to search for the truth—juries will not be influenced by the truth-related language and will instead follow the reasonable-doubt language.

We put this common judicial conclusion to an empirical test. We recruited 300 participants to serve as mock jurors in a hypothetical criminal case. After reading a case summary—including the elements of the charged crime, the trial testimony of three witnesses, and closing arguments—mock jurors were randomly assigned to one of three groups that received different jury instructions on the government's burden of proof.

Group A received an instruction merely to search for the truth and convicted at a rate of 29.6%. Group B received a legally proper reasonable doubt instruction and convicted at a rate of 16%. Group C received a combination instruction, i.e., an instruction on reasonable doubt followed by an instruction not to search for doubt but to search for the truth.

Contrary to the judicial assumption that truth-related language does not lower the burden of proof, the third group convicted at a rate of 29%—nearly *double* the rate of the group that received a legally proper reasonable doubt instruction and statistically identical to the group that received no reasonable doubt instruction at all (but was instructed merely to search for the truth). These findings are strong evidence that adding truth-related language to the end of an otherwise proper burden of proof instruction not only diminishes the government's burden, but actually eviscerates it.

Part I of this article briefly discusses the concept of proof beyond a reasonable doubt, along with its importance to individuals

1. For the purpose of illustrating the various truth-related language used by the courts, we cited cases from thirteen different jurisdictions—ten states and three federal circuits—where defendants appealed their convictions based on truth-related jury instructions. In every single case, the appellate court rejected the defendant's argument and affirmed the conviction. *See infra* notes 13–18.

and society generally. Part II surveys some of the truth-related language used in multiple state and federal jurisdictions. It also examines the constitutional problems created by this language and discusses courts' inadequate responses to these problems.

Part III explains our controlled experiment, including our hypotheses, study design, and empirical findings. Part IV discusses these findings and their significance and argues that courts should immediately terminate their use of truth-based jury instructions so that our constitutional guarantees are fulfilled. Finally, Part V discusses the possible limitations of our study and considers ways that researchers may choose to address these issues in future studies.

I. BEYOND A REASONABLE DOUBT

The concept of a heightened burden of proof for criminal cases was born centuries ago, and the expression "beyond a reasonable doubt" was formulated in 1798.[2] The United States Supreme Court has implicitly recognized this burden of proof since the late 1800s.[3] Then, in 1970, our nation's highest court explicitly held that "the Due Process Clause protects the accused against conviction except upon *proof beyond a reasonable doubt*."[4]

This proof beyond a reasonable doubt standard not only protects individual defendants, but also serves the public interest more broadly. With regard to individuals, this high burden is designed to protect us "from dubious and unjust convictions, with resulting forfeitures of life, liberty and property."[5] With regard to the public interest, the burden "is indispensable to command the respect and confidence of the community in applications of the criminal law."[6] In light of these concerns, therefore, "[i]t is critical

2. *See In re* Winship, 397 U.S. 358, 361 (1970).

3. *See, e.g.*, Davis v. United States, 160 U.S. 469, 488 (1895); Miles v. United States, 103 U.S. 304, 312 (1881).

4. *In re Winship*, 397 U.S. at 364 (emphasis added). For the history and origins of the concept of reasonable doubt, see generally Miller W. Shealy, Jr., *A Reasonable Doubt About "Reasonable Doubt*," 65 OKLA. L. REV. 225 (2013), and Steve Sheppard, *The Metamorphoses of Reasonable Doubt: How Changes in the Burden of Proof Have Weakened the Presumption of Innocence*, 78 NOTRE DAME L. REV. 1165 (2003).

5. *In re Winship*, 397 U.S. at 362 (quoting Bringar v. United States, 338 U.S. 160, 174 (1949)).

6. *Id.* at 364.

that the moral force of the criminal law not be diluted" by a lower burden of proof.[7]

Trial courts vary dramatically in how they instruct jurors on proof beyond a reasonable doubt.[8] Some courts define the phrase on a relative basis by contrasting it with the preponderance of the evidence standard—the burden used in civil cases.[9] Other courts explain the phrase by drawing analogies to important decisions jurors make in their everyday lives.[10] Yet other courts try to define beyond a reasonable doubt simply by rephrasing it—hoping, perhaps, that other words will be more intuitive or instructive.[11] And some courts simply believe that the phrase beyond a reasonable doubt "is self-defining, that there is no equivalent phrase more easily understood . . . that the better practice is not to attempt the definition, and that any effort at further elucidation tends to misleading refinements."[12]

7. *Id*. For a discussion of the various burdens of proof within a model of "probabilistic decision making" for "criminal guilt and punishment," see generally Talia Fisher, *Conviction Without Conviction*, 96 MINN. L. REV. 833 (2012).

8. *See* Victor v. Nebraska, 511 U.S. 1, 5 (1994) (giving courts tremendous leeway in their attempts to instruct jurors on the concept of reasonable doubt). For an extensive survey and discussion of the various definitions of reasonable doubt, see generally Shealy, *supra* note 4, and Hon. Richard E. Welch III, *"Give Me That Old Time Religion": The Persistence of the* Webster *Reasonable Doubt Instruction and the Need to Abandon It*, 48 NEW ENGLAND L. REV. 31 (2013).

9. *See, e.g.*, Ariz. Prelim. Crim. Jury Instructions § 7.08-C (2015) ("In civil cases, it is only necessary to prove that a fact is more likely true than not or that its truth is highly probable. In criminal cases such as this, the State's proof must be more powerful than that. It must be beyond a reasonable doubt."); N.Y. Crim. Jury Instructions 2d, *Reasonable Doubt* (2016) ("On the other hand, it is not sufficient to prove that the defendant is probably guilty. In a criminal case, the proof of guilt must be stronger than that. It must be beyond a reasonable doubt.").

10. *See, e.g.*, CRIM. JURY INSTRUCTIONS (3d Cir.) (2009) ("A reasonable doubt means a doubt that would cause an ordinary reasonable person to hesitate to act in matters of importance in his or her own life."). This analogy to jurors' decision making in their own lives, though commonly used, has been harshly criticized. *See* FED. JUD. CTR., PATTERN CRIM. JURY INSTRUCTIONS 21 comment. (1987) ("[D]ecisions we make in the most important affairs of our lives—choosing a spouse, a job, a place to live, and the like—generally involve a very heavy element of uncertainty and risk-taking. They are wholly unlike the decisions jurors ought to make in criminal cases."); WIS. CRIM. JURY INSTRUCTIONS No. 140 (2000) ("[Reasonable doubt] means such a doubt as would cause a person of ordinary prudence to pause or hesitate when called upon to act in the most important affairs of life."); United States v. Jaramillo-Suarez, 950 F.2d 1378, 1386 (9th Cir. 1991) (discussing the same commentary).

11. *See, e.g.*, CRIM. JURY INSTRUCTIONS (10th Cir.) (2011) (equating beyond a reasonable doubt with being "firmly convinced of the defendant's guilt"); FLA. STAND. JURY INSTRUCTIONS (criminal) § 3.7 (2015) (Fla.) (equating beyond a reasonable doubt with "an abiding conviction of guilt").

12. United States v. Lawson, 507 F.2d 433, 443 (7th Cir. 1974); *see also* United States

Although a wide range of language has been deemed constitutionally acceptable, one thing is constitutionally guaranteed: a court's instruction on the government's burden of proof must not dilute the burden to anything lower than beyond a reasonable doubt. Unfortunately, as the next section demonstrates, this guarantee is more theoretical than real.

II. SEARCHING FOR THE TRUTH

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."[13] Similarly, other courts instruct jurors that, when reaching their verdict, they should "evolve the truth,"[14] "seek the truth,"[15] "search for truth,"[16] or "find the truth."[17] Some courts—again, after 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."[18]

v. Vavlitis, 9 F.3d 206, 212 (1st Cir. 1993) (attempting to define the term "reasonable doubt" "is unnecessary, could confuse the jury, and provides fertile grounds for objections"); United States v. Headspeth, 852 F.2d 753, 755 (4th Cir. 1988) ("We have frequently admonished district courts not to attempt to define reasonable doubt in their instructions to the jury absent a specific request from the jury itself.").

13. State v. Dunkel, 466 N.W.2d 425, 430 (Minn. Ct. App. 1991); *see also* People v. Heron, 85 Cal. Rptr. 2d 424, 424 n.1 (Cal. Ct. App. 1999) (permitting jurors to convict if they "feel an abiding conviction of the truth of the charge"); State v. Marshall, 586 A.2d 85, 154 (N.J. 1991) (instructing jury to "let your verdict declare the truth"); State v. Mabry, 751 P.2d 882, 883 (Wash. Ct. App. 1988) (permitting jurors to convict if the "have an abiding belief in the truth of the charge").

14. United States v. Pine, 609 F.2d 106, 108 (3d Cir. 1979).

15. United States v. Gonzalez-Balderas, 11 F.3d 1218, 1223 (5th Cir. 1994); *see also* State v. Weisbrode, 653 A.2d 411, 417 (Me. 1995) ("The court instructed the jury to seek truth from the evidence."); State v. Aleksey, 538 S.E.2d 248, 251 (S.C. 2000) ("[I]nstructing the jury its 'one single objective' was 'to seek the truth.'"); State v. Benoit, 609 A.2d 230, 231 (Vt. 1992) ("During jury instructions, the trial judge twice referred to a jury's duty to 'seek the truth.'").

16. Commonwealth v. Allard, 711 N.E.2d 156, 159 (Mass. 1999); *see also* People v. Walos, 229 A.D.2d 953, 954 (N.Y. App. Div. 1996) (instructing the jurors that the trial was a "search for the truth"); State v. Needs, 508 S.E.2d 857, 866 (S.C. 1998) (instructing the jurors that they should be "in search of the truth").

17. United States v. Gray, 958 F.2d 9, 13 (1st Cir. 1992).

18. State v. Avila, 532 N.W.2d 423, 429 (Wis. 1995) (*rev'd*, in part, on other grounds).

A. *The Trouble with Truth*

By explaining to jurors the government's burden of proof beyond a reasonable doubt, but then instructing them to determine, evolve, seek, search for, or otherwise divine the truth of what really happened, courts are creating several constitutional problems.

First, tacking truth-related language on to the end of a reasonable doubt instruction is confusing because "the language implies that truth and doubt are two separate concepts . . . ."[19] In fact, truth and doubt are two separate concepts: truth refers to a judgment about whether something happened; doubt refers to the level of certainty in that judgment. Certainly a reasonably attentive juror would wonder why the judge spent so much time explaining reasonable doubt, only to instruct the juror not to search for it, but instead to search for something else entirely.

Second, courts' use of truth-related language probably does much more damage than merely creating confusion. Such language "impermissibly portray[s] the reasonable doubt standard as a defense tool for hiding the truth, and suggest[s] that a jury's scrutiny of the evidence for reasonable doubt is inconsistent with a search for the truth."[20]

Third, and most significantly, truth-related language seems to lower the government's burden of proof from the constitutionally guaranteed beyond a reasonable doubt standard. 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."[21] In other

19. *Id.*

20. State v. Berube, 286 P.3d 402, 411 (Wash. Ct. App. 2012); *see also Avila*, 532 N.W.2d at 429 (defendant argued that the instruction would lead the jury to believe "that finding doubt would mean not finding the truth").

21. United States v. Gonzalez-Balderas, 11 F.3d 1218, 1223 (5th Cir. 1994); *see also* United States v. Harper, 662 F.3d 958, 961 (7th Cir. 2011) (defendant arguing "that the 'truth' language might have been misunderstood by the jurors as an invitation to convict by a mere preponderance of the evidence"); United States v. Pine, 609 F.2d 106, 108 (3d Cir. 1979) (defendant arguing that the truth language "tend[ed] to dilute and thereby impair the constitutional requirement of proof beyond a reasonable doubt"); State v. Dunkel, 466 N.W.2d 425, 430 (Minn. Ct. App. 1991) (defendant arguing that the truth language "unconditionally dilutes the 'beyond a reasonable doubt' burden of proof standard"); *Avila*, 532 N.W.2d at 429 (defendant arguing that the truth language "would be reasonably likely to impose a lesser burden than reasonable doubt upon the State").

words, if a jury feels the government's version of events is slightly more likely than the defendant's version to be true, it would follow that, in a search for the truth, the jury would be obligated to convict the defendant.[22]

These three problems are obvious from the plain language of truth-based jury instructions. Nonetheless, courts have developed two claims in response to defendants' post-conviction arguments. As discussed below, the courts' first claim is a nonstarter on a theoretical level; the courts' second claim is the subject of our empirical test.

B. *Truth with a Capital "T"*

Some courts find these truth-related instructions to be uncontroversial in a big-picture sense: "trials *are* searches for the truth; the burden of proof is just a device to allocate the risk of error between the parties."[23] Therefore, these courts believe that it is entirely proper to instruct jurors not to search the evidence for reasonable doubt, but instead to search for the truth of what really happened. But this clever, economic-based justification does not withstand scrutiny.

The reality, as practicing defense lawyers have contended,[24] is that trials are not about searching for the truth. Instead, courts routinely prevent defendants from introducing relevant evidence of innocence. The most egregious and pervasive example is the routine exclusion of defendants' third-party guilt evidence.[25] An-

22. *See* State v. Aleksey, 538 S.E.2d 248, 251 (S.C. 2000) ("Jury instructions on reasonable doubt which charge the jury to 'seek the truth' are disfavored because they 'run the risk of unconstitutionally shifting the burden of proof to a defendant.'"); *Berube*, 286 P.3d at 411 (instructing the jury to "search for truth and not for reasonable doubt both misstates the jury's duty and sweeps aside the State's burden").

23. *Harper*, 662 F.3d at 961.

24. Criminal defense lawyers have frequently complained about the truth-related language tacked on to the end of burden of proof jury instructions. *See, e.g.*, Erik R. Guenther, *What's Truth Got to Do with It? The Burden of Proof Instruction Violates the Presumption of Innocence*, 13 WIS. DEFENDER, Fall 2005, at 1, 2; *see also* Michael D. Cicchini, *Criminal Court: Guilty by the Preponderance of the Evidence?*, MARQ. U. L. SCH. FAC. BLOG (Nov. 16, 2010), http://law.marquette.edu/facultyblog/2010/11/16/criminal-court-guilty-by-the-preponderance-of-the-evidence/.

25. In many states, a defendant is prohibited from presenting evidence—including eyewitness testimony and even DNA evidence—that a different person committed the crime, simply because the defendant is unable to articulate the third party's motive for committing the crime. This is especially troubling given that prosecutors are not required to prove the defendant's motive in order to convict. *See* Michael D. Cicchini, *An Alternative*

other example is the exclusion of defendants' expert-witness testimony.[26] In fact, truth is so low on the priority scale that even the personal interests of witnesses are elevated above the truth-seeking function.[27]

In short, trials are a set of overly complex (and occasionally irrational) rules that sometimes suppress, rather than reveal, the truth. Even the very nature of a jury's verdict—where the jury decides between guilty and *not guilty*, rather than between guilty and *innocent*—recognizes that trials were not designed to divine the truth of what actually happened, but rather to evaluate whether the government has met its burden of proof.[28]

Even if trials were ultimately concerned with finding the truth—with the burden of proof being a mere "device to allocate the risk of error between the parties"[29]—it is counterproductive for a court to instruct the jury on this supposed ultimate purpose. Here, an analogy to the social sciences can be helpful. In a psychological study, participants are not informed of the study's ultimate purpose, such as whether it is designed to test the participants' willpower, their eagerness to obey authority figures, or their willingness to share with others. Instead, psychologists go to

to the Wrong-Person Defense, 24 GEO. MASON U. CIV. RTS. L.J. 1, 16–17 (2013).

26. Courts offer a variety of creative explanations for excluding defense experts that would educate the jury on phenomena well outside their common knowledge. Sometimes, courts will even exclude the defense expert while permitting the state to call its expert on the same topic—a double standard that is not likely to aid jurors in finding the truth. *See* Danielle E. Chojnacki, Michael D. Cicchini & Lawrence T. White, *An Empirical Basis for the Admission of Expert Testimony on False Confessions*, 40 ARIZ. ST. L.J. 1, 21–26 (2008).

27. For example, defendants accused of child sexual assault often seek to introduce evidence of the complaining witness's prior sexual knowledge. This, of course, is "necessary to rebut the logical and weighty inference that [the child] could not have gained the sexual knowledge he possessed unless the sexual assault[] . . . occurred." State v. Pulizzano, 456 N.W.2d 325, 333 (Wis. 1990). Nonetheless, this prior sexual knowledge evidence is typically *excluded* by rape-shield statutes that aim to protect the complaining witness from embarrassment—even in cases where embarrassment would be highly unlikely—at the cost of keeping relevant evidence of innocence from the jury. *See, e.g.*, State v. Carter, 782 N.W.2d 695, 713 (Wis. 2010) (holding that evidence of the child's prior sexual knowledge would not have been admissible under the rape-shield statute or *Pulizzano*).

28. *See* Yali Corea-Levy, *Making Sense of Reasonable Doubt: Understanding Certainty, Doubt, and Rule-Based Bias Filtering*, 8 AM. U. CRIM. L. BRIEF 48, 48 (2012) ("Jurors are essentially asked to answer two questions . . . (1) Do you think the defendant is guilty?; (2) If so . . . are you certain beyond a reasonable doubt?"); Guenther, *supra* note 24, at 2 ("An acquittal is not a finding of . . . innocence, but merely a finding that the State has not met its burden"; similarly, in light of numerous DNA exonerations, "[a] conviction is not a finding that an accused is actually guilty, but a finding that the State has met its burden of proof beyond a reasonable doubt.").

29. United States v. Harper, 662 F.3d 958, 961 (7th Cir. 2011).

great lengths to eliminate any "demand characteristics," i.e., hints, cues, or other signals that might alert the participant to what the researcher is ultimately testing. The participants are instead urged to focus on the task at hand.[30]

Similarly, jurors should not be instructed on what some courts contend is the trial's ultimate purpose: searching for the truth. Instead, "truth is not the jury's job."[31] "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."[32] So even to the extent our jury-trial system is ultimately concerned with truth, that big-picture objective would be best achieved by instructing jurors on their specific role within that system: to evaluate the evidence through the lens of doubt.

C. *Harmless Error*

Most courts, however, do not attempt to offer this big-picture, trials-are-about-truth defense of truth-based jury instructions. In fact, most courts admit—usually implicitly but sometimes explicitly—that such truth-related mandates are not consistent with the examination of evidence for reasonable doubt.[33] Yet these courts refuse to reverse defendants' convictions and instead allow trial judges to use truth-based jury instructions.[34] Their reasoning

30. *See generally* ELIOT R. SMITH ET AL., SOCIAL PSYCHOLOGY 40, 41, 49 (Psychology Press, 4th ed. 2015) (giving a brief explanation of psychological theories).

31. State v. Berube, 286 P.3d 402, 411 (Wash. Ct. App. 2012).

32. *Id.*; *see also* People v. Katzenberger, 101 Cal. Rptr. 3d 122, 127 (Cal. Ct. App. 2009) (discussing "the jury's serious task of assessing whether the prosecution has submitted proof beyond a reasonable doubt").

33. *See, e.g.*, United States v. Gonzalez-Balderas, 11 F.3d 1218, 1223 (5th Cir. 1994) ("[S]eeking the truth . . . intimates a preponderance of evidence standard."); Commonwealth v. Allard, 711 N.E.2d 156, 160 (Mass. 1999) (truth-related language "was not entirely correct"); State v. Aleksey, 538 S.E.2d 248, 251 (S.C. 2000) (truth-related language is "disfavored" and could "shift[] the burden of proof to a defendant"); *Berube*, 286 P.3d at 411 (truth-related language "misstates the jury's duty"). *But see* State v. Dunkel, 466 N.W.2d 425, 430 (Minn. Ct. App. 1991) (proper to instruct the jury both "to search for the truth and apply the 'beyond a reasonable doubt' standard").

34. *See supra* notes 13–18. Interestingly, two of the cases we uncovered addressed truth-based language in the prosecutor's closing argument. In Washington, a court held that it was error (though harmless) for the prosecutor to argue that the jury should "search for the truth." *Berube*, 286 P.3d at 411. The court believed that this language "misstates the jury's duty and sweeps aside the State's burden." *Id.* Curiously, however, Washington permits the judge to instruct the jury to convict if it has "an abiding belief in the truth of the charge." State v. Pirtle, 904 P.2d 245, 261 (Wash. 1995). This issue of

is nearly uniform across jurisdictions: the courts conclude, without any empirical support, that when faced with two competing instructions—one to evaluate evidence for reasonable doubt and another to search for the truth—jurors will simply disregard the truth-related instruction and instead follow the reasonable doubt instruction.[35]

In reaching this apparently predetermined conclusion, many courts claim that, despite the erroneous truth-related language, the jury instruction was adequate when taken "as a whole" or viewed "in its entirety."[36] Other courts are a bit more specific and uphold convictions by pointing to the force or frequency of an otherwise proper definition of reasonable doubt elsewhere in the instruction.[37] And in perhaps the most creative way of upholding a defendant's conviction, one court acknowledged that truth-related language should be avoided; however, it decided that such erroneous language would be tolerated as long as it is "not combined with other offending terms."[38]

Other courts will, strangely, rely on the placement of truth-related language at the *end* of the instruction to conclude that it did not diminish the burden of proof. For example, one court held that truth-related language in "the last sentence does not diminish the definition of reasonable doubt given in the first two sentences."[39] Similarly, another held that truth-related language alone *does* lower the burden of proof, but is nonetheless accepta-

prosecutorial closing argument was also raised in the Seventh Circuit, where the court *rejected* the defendant's appeal based on the prosecutor's argument that the jury should "search for the truth." *Harper*, 662 F.3d at 961. The court's reasoning was that the defendant was "wrong to equate arguments of counsel with instructions from the court." *Id.* Instead, "instructions from the court carry more weight with jurors than do arguments made by attorneys." *Id.*

35. *See infra* notes 60–63.

36. *See, e.g.*, United States v. Gray, 958 F.2d 9, 13 (1st Cir. 1992); State v. Weisbrode, 653 A.2d 411, 417 (Me. 1995); State v. Dunkel, 466 N.W.2d 425, 430 (Minn. Ct. App. 1991); State v. Marshall, 586 A.2d 85, 155 (N.J. 1991); State v. Avila, 532 N.W.2d 423, 429–30 (Wis. 1995).

37. *See, e.g.*, *Harper*, 662 F.3d at 961 (noting that the search-for-the-truth language, when evaluated against the number of references to the "reasonable doubt standard" did not diminish the burden of proof); United States v. Pine, 609 F.2d 106, 108–09 (3d Cir. 1979) (indicating that he truth-related language, when viewed in the context of the "forceful charge on the requirement of proof beyond a reasonable doubt," did not dilute the burden); State v. Benoit, 609 A.2d 230, 232 (Vt. 1992) (noting that two "seek the truth" references deemed harmless when compared with six "reasonable doubt" references).

38. State v. Needs, 508 S.E.2d 857, 867–68 (S.C. 1998).

39. *Pirtle*, 904 P.2d at 262.

ble when used as a "single reference at the end of the charge" that follows the reasonable doubt instruction.[40]

This judicial conclusion—that jurors will ignore the truth-related language because it appears at the end of the instruction—is contradicted by common sense and scientific evidence. The "recency effect" is a robust and well-documented phenomenon in cognitive psychology.[41] In a strict sense, the recency effect refers to items at the end of a list as being more likely to be remembered than items located elsewhere in the list. In a broader sense, the recency effect refers to decision-makers giving more weight to recently acquired information because such information is highly salient (i.e., distinctive and noticeable) and more likely to come to mind.[42]

In short, the judicial conclusion is in opposition to the science. In reality, the truth-related language is almost impossible to ignore because of its placement at the end of the instruction. In fact, when truth-related language is tacked on at the very end, jurors are likely to view it as modifying, qualifying, or even overriding the reasonable doubt instruction that preceded it. After all, the judge's concluding instruction is clear and explicit: "[Y]ou *should not* search for doubt. You *should* search for the truth."[43]

Despite these obvious problems, courts routinely reject defendants' post-conviction challenges that focus on the plain language of, and inherent contradictions in, the instructions. Instead, courts blindly cling to the assumption that, for whatever reason, the jury was not influenced by the truth-related language. Therefore, the courts go on to conclude such language does not diminish

40. United States v. Gonzalez-Balderas, 11 F.3d 1218, 1223 (5th Cir. 1994).

41. *See* Matt Jones & Winston R. Sieck, *Learning Myopia: An Adaptive Recency Effect in Category Learning*, 29 J. EXPERIMENTAL PSYCHOL. 626, 626 (2003) ("Recency effects (REs) are a robust phenomenon in cognitive psychology. REs are said to occur whenever more recent experiences are better remembered or are more influential in judgments about present situations.") Recency effects have also been demonstrated in three studies that tested order effects in the presentation of legal evidence. *See* Adrian Furnham, *The Robustness of the Recency Effect: Studies Using Legal Evidence*, 113 J. GEN. PSYCHOL. 351, 351–52 (1986).

42. *See* Jon A. Krosnick, Fan Li & Darrin R. Lehman, *Conversational Conventions, Order of Information Acquisition, and the Effect of Base Rates and Individuating Information on Social Judgments*, 59 J. PERSONALITY & SOC. PSYCHOL. 1140, 1140–41 (1990) ("More salient information and information in short-term memory enjoy enhanced impact on social judgments.").

43. State v. Avila, 532 N.W.2d 423, 429 (Wis. 1995) (emphasis added).

the government's burden of proof. As discussed in the next section, we put this judicial assumption and conclusion to an empirical test.

III. THE STUDY

A. *Hypotheses*

Our first hypothesis is that truth and doubt are distinct concepts, which jurors understand to have different meanings, and jurors who search for the truth will convict at a higher rate than jurors who evaluate evidence for reasonable doubt. This hypothesis is not controversial. As discussed above, many courts have conceded that searching for truth and evaluating evidence for reasonable doubt are two entirely different exercises. However, after making this concession, courts simply dismiss the problem by concluding, without any empirical support, that combining the two conflicting instructions on truth and doubt does nothing to lower the government's burden of proof. Rather, the courts contend, the beyond a reasonable doubt language is somehow more powerful, and jurors will therefore disregard the truth portion of the instruction. This leads to our second hypothesis.

Our second hypothesis is that when 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.

B. *Study Design*

To test these hypotheses, we recruited 300 study participants through Amazon's Mechanical Turk, an online platform for conducting social science research. 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."[44] Further, several studies have found a high degree of similarity between the judgments and behaviors of Mechanical Turk "work-

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

ers" and of participants recruited in more conventional ways, such as through university subject pools.[45]

These 300 participants served as mock jurors and rendered a verdict in a hypothetical criminal case.[46] Each participant was required to be an adult and a United States citizen. Because two participants were not U.S. citizens, their data were discarded, leaving us with a sample of 298 mock jurors.

Our sample was large and diverse. Participants hailed from forty-two different states and the District of Columbia. Of the participants, 52% were male. Participants' ages ranged from nineteen years to seventy-six years; the mean (average) age was thirty-seven years and the median age (50th percentile) was thirty-three years.

The ethnic composition of the sample was also diverse and comprised of: 81% non-Hispanic whites, 6% African Americans, 5% Hispanics, 4% Asian Americans, and smaller percentages identifying with other groups. Forty percent of the participants have a four-year college degree, while an additional 36% have completed some college. Fourteen percent reported having prior jury experience.

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.

45. *Id.* at 3–4.

46. Before conducting our study on Mechanical Turk, we conducted a small pilot study with thirty-eight college students to test the clarity of the case materials and to assess the time needed to complete the task. The results of the pilot study were qualitatively similar to the results of our main study on Mechanical Turk. Both studies are on file with author Lawrence White.

Before being asked to render a verdict of guilty or not guilty, the 298 mock jurors were randomly assigned to one of three groups, each of which received a different jury instruction on the government's burden of proof. Group A was instructed only "to search for the truth." This instruction presumably would never be given by an actual criminal court; it was used in our study to establish a baseline and allow for a test of our first hypothesis. The "truth-only instruction," in its entirety, was as follows:

> In reaching your verdict, examine the evidence with care and caution. Act with judgment, reason, and prudence.
>
> Defendants are not required to prove their innocence. The law presumes every person charged with the commission of an offense to be innocent.
>
> The burden of establishing every fact necessary to constitute guilt is upon the State.
>
> *In reaching your verdict, you are not to search for doubt. You are to search for the truth.*[47]

Group B was given an instruction that discussed reasonable doubt and concluded with this mandate: "It is your duty to give the defendant the benefit of every reasonable doubt." This instruction probably would be deemed legally proper in nearly every jurisdiction in the country, but it does draw an analogy to decisions that jurors make "in the most important affairs of life," which has drawn harsh criticism in some jurisdictions.[48] For purposes of our study, however, it does not contain any objectionable truth-related language. The "doubt-only instruction," in its entirety, was as follows:

> In reaching your verdict, examine the evidence with care and caution. Act with judgment, reason, and prudence.
>
> Defendants are not required to prove their innocence. The law presumes every person charged with the commission of an offense to be innocent. This presumption requires a finding of not guilty unless in your deliberations, you find it is overcome by evidence which satisfies you beyond a reasonable doubt that the defendant is guilty.
>
> The burden of establishing every fact necessary to constitute guilt is upon the State. Before you can return a verdict of guilty, the evi-

47. The language in this instruction was extracted from WIS. CRIM. JURY INSTRUCTIONS No. 140 (2015).

48. FED. JUD. CTR., PATTERN CRIM. JURY INSTRUCTIONS (1987) ("[D]ecisions we make in the most important affairs of our lives—choosing a spouse, a job, a place to live, and the like—generally involve a heavy element of uncertainty and risk-taking. They are wholly unlike the decisions jurors ought to make in criminal cases."); United States v. Jaramillo-Suarez, 950 F.2d 1378, 1386 (9th Cir. 1991) (attacking analogy of making important life decisions to reasonable doubt instruction).

> dence must satisfy you beyond a reasonable doubt that the defendant is guilty.
>
> If you can reconcile the evidence upon any reasonable hypothesis consistent with the defendant's innocence, you should do so and return a verdict of not guilty.
>
> The term "reasonable doubt" means a doubt based upon reason and common sense. It is a doubt for which a reason can be given arising from a fair and rational consideration of the evidence or lack of evidence. It means such a doubt as would cause a person of ordinary prudence to pause or hesitate when called upon to act in the most important affairs of life. A reasonable doubt is not a doubt which is based on mere guesswork or speculation. A doubt which arises merely from sympathy or from fear to return a verdict of guilt is not a reasonable doubt. A reasonable doubt is not a doubt such as may be used to escape the responsibility of a decision.
>
> *It is your duty to give the defendant the benefit of every reasonable doubt.*[49]

Finally, Group C was given an instruction that combined concepts of doubt and truth. The instruction is identical to the instruction in Group B except that it adds fourteen words at the very end, concluding with this 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*." The "combination instruction," in its entirety, was as follows:

> In reaching your verdict, examine the evidence with care and caution. Act with judgment, reason, and prudence.
>
> Defendants are not required to prove their innocence. The law presumes every person charged with the commission of an offense to be innocent. This presumption requires a finding of not guilty unless in your deliberations, you find it is overcome by evidence which satisfies you beyond a reasonable doubt that the defendant is guilty.
>
> The burden of establishing every fact necessary to constitute guilt is upon the State. Before you can return a verdict of guilty, the evidence must satisfy you beyond a reasonable doubt that the defendant is guilty.
>
> If you can reconcile the evidence upon any reasonable hypothesis consistent with the defendant's innocence, you should do so and return a verdict of not guilty.
>
> The term "reasonable doubt" means a doubt based upon reason and common sense. It is a doubt for which a reason can be given arising from a fair and rational consideration of the evidence or lack of evidence. It means such a doubt as would cause a person of ordinary prudence to pause or hesitate when called upon to act in the most important affairs of life. A reasonable doubt is not a doubt which is

49. The language in this instruction was also taken from WIS. CRIM. JURY INSTRUCTIONS No. 140 (2015), though it was modified for purposes of this study.

> based on mere guesswork or speculation. A doubt which arises merely from sympathy or from fear to return a verdict of guilt is not a reasonable doubt. A reasonable doubt is not a doubt such as may be used to escape the responsibility of a decision.
>
> *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.*[50]

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

C. *Findings*

To test our first hypothesis—that truth and doubt are different concepts and that searching for the truth will produce a higher conviction rate than examining evidence for reasonable doubt—we must compare the conviction rates of Groups A and B.

Group A received the truth-only instruction, and 29 of 98 mock jurors returned verdicts of guilt for a group conviction rate of 29.6%. Group B received the doubt-only instruction, and only 16 of 100 mock jurors returned verdicts of guilt for a group conviction rate of 16%. The conviction rate among jurors who were instructed to "search for the truth" was nearly *double* the conviction rate for jurors who were instructed to evaluate the government's evidence for proof beyond a reasonable doubt.

This result is significant at the $p < .05$ level, with an exact p-value of 0.023. 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.[51] In plain language, we are more than 97% certain ($1–p$) that the difference in conviction rates between Groups A and B is a real difference and did not occur by chance.

50. The language in this instruction constitutes the entirety of WIS. CRIM. JURY INSTRUCTIONS No. 140 (2015), which combines the concepts of reasonable doubt and the search for the truth. The relevant language in this Wisconsin instruction is very similar (and often identical) to the language that was upheld in jury instructions from California, Maine, Massachusetts, Minnesota, New Jersey, New York, South Carolina, Vermont, and Washington, as well as the First, Third, and Fifth Circuits, among other jurisdictions. *See supra* notes 13–18.

51. *See* Raymond Hubbard, *Alphabet Soup: Blurring the Distinctions Between p's and α's in Psychological Research*, 14 THEORY & PSYCHOL. 295, 298 (2004).

This finding provides strong empirical support for our first hypothesis that truth and doubt are, as many courts concede, two different concepts—and jurors who are instructed to determine the truth of the matter will convict at higher rates than jurors who are instructed to evaluate the evidence for reasonable doubt.

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

As stated above, Group B received a standard reasonable doubt (doubt-only) instruction and convicted at a rate of 16%. In Group C, which received the combination instruction that included both doubt and truth language, 29 of 100 mock jurors returned verdicts of guilt for a group conviction rate of 29%. The conviction rate among jurors who were first properly instructed on reasonable doubt but then told to "search for the truth" was nearly *double* the conviction rate for jurors who were only instructed to evaluate the government's evidence for proof beyond a reasonable doubt.

This result is significant at the $p < .05$ level, with an exact p-value of 0.028. Again, the p-value measures the probability of a Type I error, i.e., obtaining a false positive.[52] Therefore, we are more than 97% certain ($1–p$) that the difference in conviction rates between Groups B and C is a real difference and did not occur by chance.

This finding provides strong empirical support for our second hypothesis that the truth-related language at the end of an otherwise proper reasonable doubt instruction actually diminishes the government's burden of proof.

After participants 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 among the three groups. In fact, all three group means were essentially 7 (fairly certain) on the 10-point scale.[53]

52. *Id.*

53. We also uncovered several subsidiary findings not directly related to the main purpose of our study: (a) women (32%) were much more likely than men (18%) to vote guilty ($p < .01$); (b) younger participants were somewhat more likely than older participants to vote guilty ($r = –.21, p < .001$), but they were slightly less certain ($r = .23, p <$

Participants also answered two attention-check questions—one testing the participant's recollection of the elements of the charged crime and one testing the participant's recollection of the instruction they received.

Overall, the attention-check results were encouraging. Nearly 84% of participants correctly identified the elements of the charged crime.[54] Those who voted not guilty were correct 86.6% of the time, while those who voted guilty were correct 75.7% of the time. This difference is statistically significant ($p < .05$), which suggests that those mock jurors who paid closer attention to the legal elements of the charged crime (sexual assault of a child) were less likely to convict.

More than 65% of participants correctly identified the burden of proof instruction they received. Group C, which received the combined doubt-and-truth instruction, performed the least well on this question. This is not surprising, as juror confusion is the most basic criticism of this instruction. However, the lower level of accuracy on this attention-check question could also be attributable to the question's placement: participants answered it after they rendered a verdict, after they reported their level of certainty in the verdict, and after they were asked to recall the elements of the charged crime.

IV. DISCUSSION: THE CASE AGAINST TRUTH

Despite conceding that reasonable doubt and truth are two different concepts, appellate courts still affirm convictions and permit trial judges to tack on truth-related language to the end of their burden of proof instructions.[55] The appellate courts' reasoning is highly suspect but consistent across jurisdictions: the truth-related language is inaccurate and adds nothing of value; however, in the larger context of an otherwise proper instruction on

.001) in their verdict; (c) more educated participants were somewhat more likely than less educated participants to vote guilty (Spearman's rho = .117, $p < .05$); (d) there was no statistically significant relationship between the ethnicity of participants and their verdicts; and (e) there was no statistically significant relationship between the prior jury experience of participants and their verdicts.

54. 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."

55. *See supra* notes 13–18.

reasonable doubt, it probably does not diminish the state's burden of proof.[56]

Our study, however, demonstrates that this judicial reasoning is flawed. In fact, not only did truth-related language diminish the burden of proof, it actually eviscerated it. Group C in our study—the group that received a reasonable doubt instruction followed by a contradictory instruction "to search for the truth"—convicted at a rate of 29%. This rate is, statistically speaking, *identical* to the 29.6% conviction rate of group A—the group that received no reasonable doubt instruction whatsoever—and is nearly *double* the 16% conviction rate of Group B, which received a legally proper reasonable doubt instruction.

Because "the Due Process Clause protects the accused against conviction except upon proof beyond a reasonable doubt,"[57] our findings provide strong evidence of a serious constitutional problem. In fact, due to our conservative study design, we may have underestimated the seriousness of the problem.

In our study, jurors were instructed only *once* on the burden of proof. Further, in order to hold the case summary constant across groups, the lawyers' closing arguments did *not* include any reference to the three different burden of proof instructions tested. This, however, is dramatically different than real-life trials where juries may be told as many as *five times* "not to search for doubt," but instead "to search for the truth." The burden of proof instruction is often given verbally before opening statements, again before closing arguments, and then in writing for the jury's reference during deliberations. Even more harmful, during closing arguments a prosecutor may parrot the court's instruction and argue to the jurors that they must search for truth, not doubt. And many prosecutors will do this both in their main argument and again in their rebuttal argument—thus leaving their "truth trumpet" ringing in the jury's ears as they begin deliberations.

Given this reiteration, any defense lawyer who argues to the jury that the evidence presented does not constitute proof beyond a reasonable doubt does so at the client's peril. The reason is that the court's truth-related jury instruction allows the prosecutor, in rebuttal argument, to dismiss the constitutionally guaranteed

56. *See supra* notes 36–40; *see infra* notes 60–63.

57. *In re* Winship, 397 U.S. 358, 364 (1970).

burden of proof as a mere "defense tool for hiding the truth," and to argue that the jury's "scrutiny of the evidence for reasonable doubt is *inconsistent* with a search for the truth."[58]

In sum, our study's findings included a near doubling of the conviction rate among jurors who were told, just a single time, to search for the truth. Being told four or five times—as is common in real-life trials—could magnify this effect. To call this a serious constitutional problem is indeed an understatement.

Fortunately, an immediate case-by-case solution to the problem is available: individual 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[59] to eliminate the "search for the truth" and similar language from its burden of proof instruction.

These recommendations should not be controversial on any level. 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,[60] and merely tolerate the language after the fact of conviction because, they conclude, it probably did no harm. That is, "it is not reasonably likely" that it will diminish the burden of proof;[61] it probably does not diminish the burden of proof "when it is not

58. State v. Berube, 286 P.3d 402, 411 (Wash. Ct. App. 2012) (emphasis added).

59. *See, e.g.*, Commonwealth v. Russell, 23 N.E.3d 867, 869 (Mass. 2015) (finding it necessary for the court to use its superintendence power to modernize and mandate the jury instruction on the government's burden of proof).

60. *See, e.g.*, United States v. Gonzalez-Balderas, 11 F.3d 1218, 1223 (5th Cir. 1994) (finding truth-related language "intimates a preponderance of evidence standard"); Commonwealth v. Allard, 711 N.E.2d 156, 159–60 (Mass. 1999) (finding truth-related language "was not entirely correct"); State v. Aleksey, 538 S.E.2d 248, 251 (S.C. 2000) (finding truth-related language is "disfavored" and could "shift[] the burden of proof to a defendant") (quoting State v. Needs, 508 S.E.2d 857, 867–68 (S.C. 1998)); State v. Berube, 286 P.3d 402, 411 (Wash. Ct. App. 2012) (finding truth-related language "misstates the jury's duty").

61. State v. Avila, 532 N.W.2d 423, 429 (Wis. 1995).

combined with other offending terms";[62] and it probably "does not diminish the definition of reasonable doubt . . . but neither does it add anything of substance."[63]

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."[64] 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.[65]

V. LIMITATIONS OF THE STUDY AND FURTHER TESTING

Potential criticisms of our study include our use of the case summary method, the use of a single criminal charge and fact pattern, the lack of deliberation by our mock jurors, our inability to observe participants' level of attention, and the failure to screen for juror bias. We discuss these limitations below.

All empirical studies are flawed in the sense that methodological decisions designed to solve one problem often exacerbate another. For example, controlled experiments that use random assignment solve the problem of causal ambiguity, i.e., determining what produced the effect. But the desire to control extraneous variables may constrain the researcher's ability to generalize the study's results beyond the specific conditions tested.

The present study is not an exception to the rule, so we encourage future researchers to replicate our study and to use different

62. *Needs*, 508 S.E.2d at 867.

63. State v. Pirtle, 904 P.2d 245, 262 (Wash. 1995).

64. United States v. Reynolds, 64 F.3d 292, 298 (7th Cir. 1995) (discussing the risks of any attempt to expound upon reasonable doubt) (quoting United States v. Hanson, 994 F.2d 403, 408 (7th Cir. 1993)).

65. In other parts of the article, we refer to removing truth-related language from the "burden of proof" instruction more specifically, as that is where the offending language is nearly always located. *See supra* Part II. However, merely moving the language to another instruction—such as the "duty to deliberate" instruction in *United States v. Gonzalez-Balderas*, 11 F.3d 1218, 1223 (5th Cir. 1994)—does not solve the problem. In fact, due to the recency effect and the way jurors process information, moving the language to a different instruction could, depending on the instruction's placement and timing, magnify the problem.

test materials and (possibly) different methodologies when they do. The goal of social science is to arrive at conclusions that are supported by multiple converging lines of evidence, with each contributing study being necessarily flawed, but flawed in a different way.

A. *Case Summary Method*

The case summary method employed in our study—where the jury was provided with the elements of the charged crime and a transcript of the witnesses' testimony and the lawyers' closing arguments—is similar to the method that has been used in many peer-reviewed studies, including studies that examined the impact of a defendant's personal characteristics on jury decision making.[66] However, some social scientists have expressed concerns about the ecological validity of these studies and have called for more realistic trial simulations.[67]

Conversely, other social scientists have noted the prohibitive costs of more realistic trial simulations[68] or have failed to observe differences in the reactions of mock jurors to abbreviated and more elaborate case summaries.[69] Therefore, "even highly artificial simulations are not *inherently* distorting and may actually inform us on relationships of real significance for law and human behavior."[70] Further, the more realistic trial simulation methods actually "provide a myriad of additional legally relevant and *ir-*

66. Wayne Weiten & Shari S. Diamond, *A Critical Review of the Jury Simulation Paradigm*, 3 L. & HUM. BEHAV. 71, 77 (1979); *see also* Lawrence T. White, *Juror Decision Making in the Capital Penalty Trial: An Analysis of Crimes and Defense Strategies*, 11 L. & HUM. BEHAV. 113, 116 (1987).

67. *See* Joel D. Lieberman & Bruce D. Sales, *What Social Science Teaches Us About the Jury Instruction Process*, 3 PSYCHOL., PUB. POL. & L. 589, 592 (1997) ("A better methodology is to provide a videotaped trial to participants. The videotaped format provides a highly engaging simulation, and is much more representative of an actual trial. Consequently, greater faith can be placed in studies using this methodology than experiments using transcripts or case summaries.").

68. Shari S. Diamond, *Illuminations and Shadows from Jury Simulations*, 21 L. & HUM. BEHAV. 561, 563–64 (1997).

69. Geoffrey P. Kramer & Norbert L. Kerr, *Laboratory Simulation and Bias in the Study of Juror Behavior: A Methodological Note*, 13 L. & HUM. BEHAV. 89, 89 (1989) ("Results provided no support for the contention that treatment effects act differently as a function of the length of the stimulus trial in which they are embedded. Rather, it is suggested that treatments used in simplified jury simulations may often show similar effects when examined in more realistic, complex settings *if* the treatments are comparable.").

70. *Id.* at 99.

relevant bases on which to make a decision,"[71] including, for example, a defendant's or witness's race and ethnicity.

Seen in this light, the simplicity of the case summary method may actually be its strength. First, researchers who use the case summary method can eliminate extrajudicial factors, including race and ethnicity, which may have an impact on jurors' decision-making processes. Second, the more abbreviated case summary method compresses events in time, thereby reducing the pernicious effect of forgetting, which can also affect jurors' decision-making processes. Third, the case summary method allows researchers to test the impact of a specific component of a trial—in our study, a particular jury instruction—that may get lost in the clutter of a more complex trial simulation. For these reasons, "more abbreviated experimental stimulus materials *can* play an important role in addressing some questions about jury behavior."[72]

B. *Single Fact Pattern*

Like actual trials, simulation studies usually include a particular criminal charge and a unique set of facts. This single fact pattern is then held constant across test conditions. In this way, different effects cannot be attributed to different fact patterns; rather, they can only be attributed to the variable (e.g., the type of jury instruction) that was manipulated by the researcher. However, the drawback of using this procedure is obvious: limited generalizability.

That is, while our findings allow us to conclude that truth-related language diminishes the burden of proof in criminal cases, we cannot know the *extent* to which this effect will also be observed in other cases with different fact patterns. For example, in our study, the alleged crime was touching over clothing, and there was no eyewitness or physical evidence to corroborate the accusation. As a result, mock jurors who received the standard reasonable doubt (doubt-only) instruction convicted at a rate of

71. Diamond, *supra* note 68, at 564 (emphasis added).

72. *Id.* (While certain studies, such as those that test "the credibility of various types of expert testimony," demand "a fairly elaborate simulation," other studies can be accomplished using "a less extensive trial stimulus.").

only 16%. This conviction rate nearly doubled for the group that received the combination (doubt-and-truth) instruction.

Our fact pattern was ideal for testing our hypotheses, but we cannot say that the impact of the doubt-and-truth instruction would be identical when applied to different cases. For example, in cases where there is *more* evidence of guilt—such as an eyewitness or some physical evidence—the impact of the truth-related language could be even greater.[73] Similarly, as discussed previously, cases that include repeated judicial instructions to search for the truth, or prosecutorial closing arguments parroting the court's instructions, may also magnify the impact of the language.[74]

Therefore, in order to assess the generalizability of our findings, future researchers may wish to construct hypothetical cases that use different fact patterns. Researchers may wish to test the impact of the various jury instructions under circumstances in which the evidence presented produces a larger percentage—perhaps even a majority or near majority—of guilty verdicts when using a standard reasonable doubt (doubt-only) instruction. Further, researchers may wish to increase their mock jurors' exposure to the truth-related language, whether through additional instructions from the judge or through prosecutorial closing argument.

C. *Lack of Deliberations*

Our study tested the impact of three different jury instructions on *individual* mock jurors' verdicts. These mock jurors did not deliberate as a group before reaching their decisions. In this sense, our study differed from an actual jury trial and some other jury simulation studies.

73. The impact could be greater in terms of the percentage increase in guilty votes and in terms of turning a pre-deliberation minority of guilty votes into a pre-deliberation majority. *See supra* Part III.C.

74. In our study, the jurors who received an instruction "to search for the truth" were instructed as such only once. In a real-life trial, however, jurors may receive truth-related instructions up to five times: in the court's verbal instructions before and after the evidence, in the prosecutor's closing argument and rebuttal argument, and in the court's written instructions provided to the jury for their deliberations. In fact, given that prosecutors commonly raise this in rebuttal argument, the mandate "to search for the truth" is often the last thing the jury hears before deliberating. *See supra* Part IV.

Some studies show that "deliberations sometimes do influence outcomes," including, for example, a study in which juror deliberations reduced individual juror biases and made them more likely to follow the judge's instructions.[75] However, the evidence on the impact of deliberations is, at best, mixed.[76]

For example, several studies have tested the impact of deliberations on the physical attractiveness bias, examining the tendency for jurors to perceive and treat attractive defendants more favorably than plain-looking defendants.[77] A study in 1974 found that deliberation mitigated the physical attractiveness bias.[78] A study in 1990, however, found that deliberation exacerbated the bias.[79] Most surprising of all, a study in 2008 found a *reversal* of the expected effect: Mock jurors who deliberated were biased *against* the attractive defendant.[80]

In the end, however, having mock jurors deliberate before rendering a verdict is not likely to change the observed pattern of verdicts across conditions. Rather, "[t]he prevailing view . . . is that deliberations play a minor role in determining jury verdicts because the predeliberation majority generally prevails in the end."[81]

D. *Participant Attention Level*

When using Mechanical Turk, as opposed to a more realistic trial simulation method, it is not possible to directly observe participants' level of attention. However, we were able to test partic-

75. Diamond, *supra* note 68, at 565 (describing a 1994 study conducted by J. Kerwin and D.R. Shaffer).

76. *Id.*; *see also* Lieberman & Sales, *supra* note 67, at 635 ("On the basis of these contradictory findings, we cannot assume that deliberation will eliminate the problem of incomprehensible instructions.").

77. *See* Marc W. Patry, *Attractive But Guilty: Deliberation and the Physical Attractiveness Bias*, 102 PSYCHOL. REP. 727, 727 (2008).

78. Richard R. Izzett & Walter Leginski, *Group Discussion and the Influence of Defendant Characteristics in a Simulated Jury Setting*, 93 J. SOC. PSYCHOL. 271, 271 (1974).

79. *See* Robert J. MacCoun, *The Emergence of Extralegal Bias*, *During Jury Deliberation*, 17 CRIM. JUST. & BEHAV. 303, 311 (1990).

80. *See* Patry, *supra* note 77, at 731.

81. Diamond, *supra* note 68, at 564. For purposes of truth-related jury instructions, the real danger lies in cases that have more evidence of guilt than was presented in our study's case summary. That is, given that our findings included a near doubling of guilty verdicts when jurors were told "to search for the truth," a predeliberation minority of guilty votes could, in many cases, easily become a predeliberation majority.

ipant attention level through the use of two post-verdict attention-check questions. As indicated above, these results were largely encouraging and demonstrated that, on the whole, our mock jurors devoted adequate attention to the case study materials before rendering their verdicts.

Nevertheless, we know that some participants failed to pay adequate attention to the task at hand. Six of the 298 participants completed the task in less than two minutes, whereas the average time required was nearly fourteen minutes. While unfortunate, this problem exists with real-life jurors as well. For example, far worse than merely being inattentive, *sleeping* jurors are routinely tolerated as long as the trial judge concludes that the jurors were not sleeping too long or the evidence they missed was not important enough to justify a new trial.[82] In this regard, then, the fact that a small percentage of mock jurors in our study did not give sufficient attention to the task actually mirrors the problems encountered with real-life jurors.

E. *Participant Bias*

In theory, biased jurors are excused from jury duty and do not participate in actual trials. In our Mechanical Turk study, we were not able to screen participants in advance for potential bias. However, four things mitigate this potential problem.

First, many biases—for example, racial biases or a predisposition to view police officers more (or less) favorably than citizen witnesses—would not have been factors in our case. Our hypothetical defendant's (and even the witnesses') race was not provided, and the case summary did not include any testimony by a police officer.

Second, in actual courtrooms across the country, biased jurors find their way onto juries. In fact, to exclude a *subjectively* biased juror, all of the following must happen: the would-be juror must be aware of his or her bias; the judge or the attorneys must devise questions to expose that particular bias; and the would-be juror must actually admit his or her bias to a roomful of fellow citizens. For all of these things to happen is a rare occurrence indeed.

82. *See* State v. Chestnut, 643 S.W.2d 343, 346 (Tenn. Ct. App. 1982) (upholding the denial of defendant's motion for new trial despite undisputed evidence of two jurors sleeping through evidentiary portions of trial).

Third, even when a juror is, by all accounts, *objectively* biased, he or she may still be permitted to serve on the jury. Perhaps the most egregious example occurred when a court permitted the prosecutor's own employee to serve on the defendant's jury, finding that the employee-employer relationship between the juror and the prosecutor was not sufficient to justify the juror's removal.[83]

So once again, this problem—the potential for participant bias—mirrors the problems encountered with real-life jurors. Therefore, our inability to screen for potential bias does not compromise the generalizability of our finding because actual trials also include biased jurors.

Finally, there is a fourth mitigating factor: the random assignment of the study participants to one of three test conditions. The virtue of random assignment is that, when used with large numbers of study participants, it produces groups that are statistically equivalent to each other in all respects. Each group has roughly the same number of mock jurors, the same number of men and women, the same number of well-educated and poorly educated persons, and the same number of biased and unbiased individuals.

When test groups are statistically equivalent at the outset, receive different jury instructions, and then convict at different rates, we can be quite certain that the different conviction rates were produced by the different jury instructions and not by personal characteristics of the mock jurors in a particular group. In plain language, random assignment creates a level playing field where the effects of bias are distributed equally across the test conditions. Therefore, the end result—a difference in conviction rates—can only be attributed to the type of jury instruction received.

83. *See* State v. Smith, 716 N.W.2d 482, 483 (Wis. 2006). The dissent, however, offered a far more rational view, stating, "an objectively reasonable person in the place of the challenged prospective juror would not ordinarily be able to separate his or her economic and loyalty interests from the determinations he or she would be required to make as juror. An employee of a district attorney's office should therefore be struck as a juror for cause when that office is prosecuting a case." *Id*. at 495 (Abrahamson, C.J., dissenting).

CONCLUSION

The Constitution protects a criminal defendant from conviction unless the government can prove guilt beyond a reasonable doubt. Many trial courts properly instruct juries on reasonable doubt, but then conclude by instructing them not to search for doubt, but instead to evolve, seek, or search for the truth.

Defendants have argued that this truth-related language at the end of the burden of proof instruction reduces the government's burden to a mere preponderance of the evidence, and is therefore constitutionally defective. That is, if a jury feels the government's version of events is slightly more credible than the defendant's version, it would follow that, in a search for the truth, the jury is obligated to convict the defendant.

Many courts concede that this truth-related language is improper. However, they go on to deny defendants' instruction-based appeals by merely assuming, without any empirical support, that when faced with these two competing instructions—one to evaluate evidence for reasonable doubt and another to search for the truth—juries will not be influenced by the truth-related language and will instead follow the reasonable-doubt language.

Our study empirically tested this judicial assumption by asking 298 mock jurors to render a verdict in a hypothetical criminal case. After reading a case summary that included the elements of the charged crime, witness testimony, and closing arguments, the mock jurors were randomly assigned to one of three groups, with each group receiving a different burden of proof instruction.

Group A was instructed only to search for the truth and convicted at a rate of 29.6%. Group B received a legally proper reasonable doubt instruction and convicted at a rate of 16%. Group C received a combination instruction that properly defined reasonable doubt and then instructed jurors not to search for doubt, but instead to search for the truth. The conviction rate for this group jumped to 29%—nearly *double* the conviction rate of the group that was properly instructed on reasonable doubt and statistically identical to the conviction rate of the group that received no reasonable doubt instruction whatsoever.

Our findings demonstrate that truth-related language not only reduces the government's burden of proof but actually eviscerates it. The courts' previous justification for tolerating such truth-

related language—it adds nothing of value, poses a theoretical risk, but probably does no actual harm—was not persuasive to begin with. Now that empirical evidence contradicts this justification, all truth-related language should be immediately removed from jury instructions regarding the government's burden of proof beyond a reasonable doubt.