



QUANTIFYING REASONABLE DOUBT

Daniel Pi, Francesco Parisi,† and Barbara Luppi‡*

TABLE OF CONTENTS

I. INTRODUCTION	455
II. BACKGROUND	459
A. <i>The Origins of Reasonable Doubt</i>	459
B. <i>The Search for a Definition</i>	463
III. INTERPRETING REASONABLE DOUBT	469
A. <i>Articulable Doubt</i>	469
B. <i>Reasonable Person’s Doubt</i>	472
C. <i>Probabilistic Doubt</i>	477
1. Juror Responsiveness.....	477
2. Frequency and Certainty	485
3. Ratio of Social Costs.....	489
4. Heterogeneity and Tailoring the Standard.....	492
5. Merging Doctrinal β and Efficient β	495
IV. THE REASONABLE DOUBT STANDARD BY JURISDICTION	499
V. CONCLUSION	507

I. INTRODUCTION

In a criminal trial, the prosecution bears the burden of proving the defendant’s guilt *beyond a reasonable doubt*. The principle is a cornerstone of the common law:¹ every criminal conviction in the United

* Mitchell Hamline School of Law. E-mail: daniel.pi@mitchellhamline.edu.

† University of Minnesota, School of Law and University of Bologna, Department of Economics. E-mail: parisi@umn.edu.

‡ University of Modena and Reggio Emilia, Department of Economics and John Hopkins University. E-mail: barbara.luppi@unimore.it. Many thanks to Vincent Todd, Lokys Gust, Jessica Ford, Ryan Bravo, Odise Carr, Favio Ramirez-Caminatti, and Lisa Kaster for their invaluable research assistance. Special thanks also to Tomasz Mielniczuk for his insightful suggestions and editorial assistance.

1. See *United States v. Pepe*, 501 F.2d 1142, 1143 (10th Cir. 1974) (“[W]e recognize that the reasonable doubt standard is a constitutional cornerstone of the criminal justice system.”); see also Herschel E. Richard, Jr., *Insanity—The Burden of Proof*, 30 LA. L. REV. 117, 124 (1969).

States,² as well as those in the United Kingdom,³ Canada,⁴ Australia,⁵ and New Zealand,⁶ filters through this most stringent standard of proof.⁷

Yet despite its paramount importance in determining the fates of the accused, courts have heretofore demonstrated an exceeding reluctance to articulate what it means for jurors to be persuaded of a defendant's guilt "beyond a reasonable doubt."⁸ Appeals courts have viewed all attempts at meaningful elucidation with disfavor, insisting that "reasonable doubt" should not be defined,⁹ while paradoxically insisting that it is not

2. In the United States, not only must the overall certainty of guilt meet the reasonable doubt threshold, but *every element* of the offense must meet this standard. *In re Winship*, 397 U.S. 358, 364 (1970).

3. See *Woolmington v. Dir. of Pub. Prosecutions* [1935] AC 462 (HL) (appeal taken from Eng.).

4. See *R. v. Lifchus* [1997] 3 S.C.R. 320 (Can.).

5. See *Green v The Queen* (1971) 126 CLR 28 (Austl.).

6. See *R v. Wanhalla* [2007] 2 NZLR 573 at [5] (N.Z.).

7. The criminal standard of proof in civil law systems is essentially analogous to the reasonable doubt standard. The judge must possess an "*intime conviction*": an "inner, deep-seated, personal conviction." Kevin M. Clermont & Emily Sherwin, *A Comparative View of Standards of Proof*, 50 AM. J. COMP. L. 243, 246, 250 (2002). Translation problems render precise comparisons difficult. Nevertheless, comparative law scholars seem to have embraced the equivalency of reasonable doubt and civil law standards of proof. The French Ministry of Justice, in an English language overview of the French legal system, translates its standard of proof as "beyond reasonable doubt." *The French Legal System*, FRENCH MINISTRY OF JUSTICE 10 (2012), http://www.justice.gouv.fr/art_pix/french_legal_system.pdf.

8. See, e.g., EDWARD JAMES DEVITT ET AL., FEDERAL JURY PRACTICE AND INSTRUCTIONS 473 (5th ed. 2000); LEONARD B. SAND ET AL., MODERN FEDERAL JURY INSTRUCTIONS ¶ 4.01 (2002); Robert C. Power, *Reasonable and Other Doubts: The Problem of Jury Instructions*, 67 TENN. L. REV. 45, 61–62 (1999).

9. See, e.g., *Holland v. United States*, 348 U.S. 121, 140 (1954) (citation omitted) ("Attempts to explain the term 'reasonable doubt' do not usually result in making it any clearer to the minds of the jury."); *Adorno v. Melvin*, 876 F.3d 917, 922 (7th Cir. 2017) ("As a matter of Illinois common law, trial judges are instructed not to attempt to define reasonable doubt for the jury."); *United States v. Herman*, 848 F.3d 55, 57 (1st Cir. 2017) (citation omitted) ("As an initial matter, we have repeatedly noted 'that reasonable doubt does not require definition.' Thus, 'an instruction which uses the words reasonable doubt without further definition adequately apprises the jury of the proper burden of proof.'"); *United States v. Fields*, 660 F.3d 95, 96 (1st Cir. 2011) ("Our decisions hold that 'reasonable doubt does not require definition.'"); *United States v. Walton*, 207 F.3d 694, 696 (4th Cir. 2000) ("There is no constitutional requirement to define reasonable doubt to a jury. The Supreme Court has never required trial courts to define the term."); *United States v. Oriakhi*, 57 F.3d 1290, 1300 (4th Cir. 1995) ("It is well settled in this circuit that a district court should not attempt to define the term 'reasonable doubt' in a jury instruction absent a specific request for such a definition from the jury."); *Smith v. Commonwealth*, 410 S.W.3d 160, 169 (Ky. 2013) (citations omitted) ("In 1978, this Court amended RCr 9.56 to eliminate a former articulation of the concept of 'reasonable doubt' and to explicitly provide that the jury should not be instructed upon a definition of 'reasonable doubt.' In *Commonwealth v.*

“undefinable.”¹⁰

The reluctance of courts to settle on a consistent and tractable analysis of reasonable doubt carries profound consequences for the administration of criminal justice. Numerous studies of mock juries and judicial surveys have found that laypeople and lawyers are wildly heterogeneous in how they understand the term “reasonable doubt.”¹¹ Consequently, like cases are assuredly not treated alike—one defendant may be convicted of a crime, while another prosecuted under substantially identical facts may be acquitted, only because the fact-finders in their respective cases happened to understand the term “reasonable doubt” differently. This is not a mere theoretical possibility. It is a real problem, infecting the outcomes of all criminal trials.

Yet curiously, while eschewing all attempts to define the standard of proof, courts evince no diffidence endorsing precise numerical objectives when expounding upon the relative undesirability of erroneous outcomes. William Blackstone famously declared in his *Commentaries on the Laws of England*, “[T]he law holds, that it is better that ten guilty persons escape, than that one innocent suffer.”¹² A multitude of judges, lawyers, and scholars have enthusiastically quoted or paraphrased Blackstone, affirming this putative principle of law, though sometimes varying the numerical specifics. John Fortescue thought the ratio ought to be twenty

Callahan, we extended the well-settled prohibition of defining reasonable doubt to all points in a trial proceeding, stating ‘trial courts shall prohibit counsel from any definition of “reasonable doubt” at any point in the trial . . .’; *Colburn v. State* 990 So. 2d 206, 217 (Miss. Ct. App. 2008) (“It is a long-standing rule that defining ‘reasonable doubt’ for the jury is improper.”); *Postelle v. State*, 267 P.3d 114, 133 (Okla. Crim. App. 2011) (“We have consistently and repeatedly held that reasonable doubt is self-explanatory, and that rather than clarifying the meaning of the phrase, definitions of reasonable doubt tend to confuse the jury.”); *Matthews v. State*, 478 S.W.3d 781, 782–83 (Tex. App. 2015) (“We did not hold, and we do not now hold that giving such an instruction is a wise thing for trial courts to do.”); *State v. Levitt*, 148 A.3d 204, 211 (Vt. 2016) (“[W]e continue our observation that attempting to define reasonable doubt is a ‘hazardous undertaking,’ and continue to discourage trial judges from trying such an explanation.”).

10. *Victor v. Nebraska*, 511 U.S. 1, 26 (1994) (Ginsburg, J., concurring).

11. See, e.g., Dorothy K. Kagehiro, *Defining the Standard of Proof in Jury Instructions*, 1 PSYCHOL. SCI. 194, 196 (1990); Dorothy K. Kagehiro & W. Clark Stanton, *Legal v. Quantified Definitions of Standards of Proof*, 9 LAW HUM. BEHAV. 159, 164 (1985); L.S.E. Jury Project, *Juries and the Rules of Evidence*, 1 CRIM. L. REV. 208 (1973); C.M.A. McCauliff, *Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees*, 35 VAND. L. REV. 1293, 1325 (1982); Rita James Simon & Linda Mahan, *Quantifying Burdens of Proof: A View from the Bench, the Jury, and the Classroom*, 5 L. & SOC’Y. REV. 319, 325, 327–28 (1971).

12. WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND 352 (1769).

to one.¹³ Matthew Hale proposed five to one.¹⁴

Facially, it seems that any Blackstonian ratio necessarily implies a precise measure of the reasonable doubt standard.¹⁵ Thus is it perplexing that courts affirmatively disclaim the existence of a numerical formulation of reasonable doubt, for they have enthusiastically endorsed a principle of law which necessarily implies exactly that. We might hypothesize that the implication escaped the attention of careless writers, yet it seems at least some judges were, in fact, clearly aware of the implication. Consider, for example, Judge Weinstein, writing in *United States v. Fatico*:

If, in the case of proof “beyond a reasonable doubt,” the figure of 95% Or 99% Is [sic] used, it means that we would rather have, respectively, twenty or one hundred guilty persons go free than more than one innocent person be convicted. Blackstone would have put the probability standard for proof “beyond a reasonable doubt” at somewhat more than 90%, for he declared: “It is better that ten guilty persons escape than one innocent suffer.”¹⁶

And yet—more curiously still—on the occasions when inventive defense attorneys proposed to include such language defining “reasonable doubt” in jury instructions, courts have uniformly rejected it.¹⁷

The scholarly record is replete with complaints—from both academics and judges—about the law’s schizophrenic treatment of reasonable doubt. In this Article, we take stock of the current state of scholarship, drawing together several disparate strands of historical, philosophical, economic, and experimental research. We consider several new arguments in favor of numerical quantifications of the reasonable

13. JOHN FORTESCUE, COMMENDATION OF THE LAW OF ENGLAND 45 (Francis Grigor trans., Sweet and Maxwell 1917) (1537).

14. MATTHEW HALE, HISTORIA PLACITORUM CORONAE: THE HISTORY OF THE PLEAS OF THE CROWN 289 (1736). For an amusing compilation of the various values which have been proposed throughout history, see Alexander Volokh, *n Guilty Men*, 146 U. PA. L. REV. 173, 175–76 (1997).

15. See, e.g., Alan D. Cullison, *Probability Analysis of Judicial Fact-Finding: A Preliminary Outline of the Subjective Approach*, 1 U. TOL. L. REV. 538, 565 (1969); John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065, 1077 (1968); Stuart Nagel, *Bringing the Values of Jurors in Line with the Law*, 63 JUDICATURE 189, 191–92 (1979). But see Michael L. DeKay, *The Difference Between Blackstone-Like Error Ratios and Probabilistic Standards of Proof*, 21 L. & SOC. INQUIRY 95, 99 (1996).

16. *United States v. Fatico*, 458 F. Supp. 388, 411 (E.D.N.Y. 1978).

17. See *Coleman v. State*, 13 N.E. 100, 103 (Ind. 1887) (“[I]t was not error to refuse the defendant’s request to instruct that it was better that ‘ten guilty persons escape than that one innocent suffer.’”).

doubt standard, inferred from Blackstonian ratios, and suggest a practicable legal basis for such quantification.

II. BACKGROUND

A. *The Origins of Reasonable Doubt*

The standard of proof in criminal trials has undergone much evolution over the centuries. In the earliest practice of English law, very little adjudicative procedure was formally defined. The strict separation between evidence production and evidence evaluation were blurred. Jurors, drawn from small communities, were expected to have knowledge of the accused, his alleged victims, and the events precipitating the dispute, serving dually as witnesses in the case and finders of fact.¹⁸

Moreover, inasmuch as the division between witness and fact-finder was obscured, so too were the divisions between law and religion or legality and morality. Jurors swore oaths to the verity of their testimonies and judgments. To convict falsely was not only a failure of their civic duty, but a sin against the accused and against “God.”¹⁹ In capital crimes, the sin would have been tantamount to murder, for the

18. See, e.g., FREDERIC WILLIAM MAITLAND & FRANCIS C. MONTAGUE, A SKETCH OF ENGLISH LEGAL HISTORY 56–57 (1915) (“Originally the jurors are called in, not in order that they may hear, but in order that they may give, evidence. They are witnesses. They are the neighbours of the parties; they are presumed to know before they come into court the facts about which they are to testify.”); Thomas A. Green, *The Jury and the English Law of Homicide, 1200–1600*, 74 MICH. L. REV. 413, 424 (1976) (“The jury’s opinion, however, may also have been based on what it had long known about either the homicide or the defendant, or on what it had just learned in the rumor mill of the gaol delivery session. It was the jurors’ duty to render a verdict on the basis of their knowledge of the facts, whether that knowledge was first-hand or gleaned from those whom they chose to believe.”); Theodore Waldman, *Origins of the Legal Doctrine of Reasonable Doubt*, 20 J. HIST. IDEAS 299, 308 (1959) (“The distinction between a witness and a juror was not one clearly made throughout the history of trial by jury, and it is this distinction with its ramifications which led to the modern system. . . . It was not until 1705 that the jury could come from the body of the county rather than the immediate vicinity of the place in which the facts at issue occurred. . . . When juries were summoned an attempt was made to get persons who would be likely to know the facts in the case.”). The proposition that juries relied principally on firsthand knowledge in determining a verdict has been a topic of some scholarly dispute, however. See generally Daniel M. Klerman, *Was the Jury Ever Self-Informing?*, 77 S. CAL. L. REV. 123 (2003).

19. See Anthony A. Morano, *A Reexamination of the Development of the Reasonable Doubt Rule*, 55 B.U. L. REV. 507, 510 (1975); 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, 1172 (2003) (“The decision of each juror, the mechanism by which the jurors would find the matter, was not a question for the law but a matter of private conscience. As a matter of conscience, the juror had to swear on oath to convict.”).

innocence of the defendant would undermine the moral justification for his killing by the state, and the erring juror would bear the responsibility. Such were the stakes in very many cases, as death was the preferred sanction in the early common law.²⁰ Thus were jurors exceedingly eager to acquit but on the surest showing of proof.²¹ To convict wrongly—according to the metaphysics of the age—entailed damnation of their souls.²²

The original standard of proof in common law criminal trials was therefore implicitly guilt beyond *any* doubt. The “any doubt” standard was a byproduct of jurors’ religious incentives.²³ Over several centuries, it calcified to a principle of law. In contradistinction to our modern practice, *unreasonable* or fanciful doubts were not to be set aside under the “any doubt” standard. Any cause for hesitation was sufficient to acquit, no matter how far-fetched. Consequently, early English criminal trials rarely resulted in convictions.²⁴

Gradually, several forces bore against the “any doubt” standard. First, given the scarcity of convictions, we can infer that pre-Enlightenment English criminal law suffered from systematic underdeterrence.²⁵ This inefficiency may well have exerted substantial pressure for change. Yet this alone was not sufficient, for the entrenchment of religious doctrine foreclosed the feasibility of reform.²⁶

20. THOMAS A. GREEN, VERDICT ACCORDING TO CONSCIENCE: PERSPECTIVES ON THE ENGLISH CRIMINAL TRIAL 1200–1800 5 (1985).

21. *Id.* at 27, 98; see also Barbara Shapiro, *The Beyond Reasonable Doubt Doctrine: Moral Comfort or Standard of Proof?*, 2 L. & HUMAN. 149, 151–52 (2008).

22. See generally JAMES Q. WHITMAN, THE ORIGINS OF REASONABLE DOUBT: THEOLOGICAL ROOTS OF THE CRIMINAL TRIAL (2008) (tracking the evolution of the reasonable doubt standard from its religious origins).

23. We should make clear that we are rehearsing what we find to be the prevailing narrative in the scholarly literature; we have undertaken no independent historical research. We should hasten to add that the religious entanglement theory, advanced by Morano, Sheppard, and Whitman is a point of ongoing debate among historians. See Shapiro, *supra* note 21, at 161–62.

24. See Shapiro, *supra* note 21 (first citing J.S. COCKBURN, A HISTORY OF ENGLISH ASSIZES 1558–1714 (1972); then citing GREEN, *supra* note 20; then citing BARBARA A. HANAWALT, CRIME AND CONFLICT IN ENGLISH COMMUNITIES 1300–1348 (1979)).

25. See Shapiro, *supra* note 21. However, we recognize a colorable argument that the harshness of sanctions could, under a Beckerian model of deterrence, have resulted in efficient deterrence despite low conviction rates. See Gary Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169 (1968).

26. See Morano, *supra* note 19, at 509–13. But see Shapiro, *supra* note 21, at 151–54. We have chosen to characterize this factor in motivating legal change in somewhat agnostic terms, as its significance has been a point of some controversy. We would be remiss not to cite John H. Langbein, whose research on the topic enjoys considerable authority. See JOHN H. LANGBEIN, THE ORIGINS OF ADVERSARY CRIMINAL TRIAL (2003).

A second impetus for change grew out of the skeptical turn in Enlightenment era philosophy. The epistemological inquiries of the seventeenth and eighteenth centuries undermined mystical notions of “perfect knowledge.” Science and reason displaced revelation. Naïve expectations of total certainty in all matters—scientific, philosophical, and practical—gave way to more realistic theories of knowledge. Epistemologists undertook ever more careful and rigorous analyses of justification. And the fundamentals of Scholastic absolutism—a priori knowledge, divine revelation, and deductive moral reasoning—withered under improving standards of philosophical argumentation.²⁷ The trends in philosophical research thus militated against a legal standard of proof, which required a juror’s belief to be free from any doubt.

A third force at work was the rise of liberalism and legal positivism. Legal philosophers—most notably Austin and Bentham—counterarguing against Thomistic conceptions of legality, persuasively rebutted the supposed necessary connection between law and morality. The concept of a manmade *positive law* came to be understood as distinct from *divine law* or *positive morality*,²⁸ which encouraged jurors to disentangle their notions of civic duty from their feelings of religious or moral desert.

Finally, the rise of Protestantism in England generated an opportunity for courts to introduce a clear break from the Catholic dogma of the preceding age. The establishment of new religious foundations allowed English lawyers of the Enlightenment to lay new groundwork, liberating them from past practices to reconceive the relationship between religion and legal systems generally. Indeed, the desire to distinguish Protestant practices—legal and otherwise—from those of the Catholic Church may have even generated an affirmative incentive to do things differently for the mere sake of doing things differently.²⁹

27. The scholarly literature on the influence of Enlightenment epistemology on the law of evidence and standards of proof is vast. The following sources are helpful starting points in exploring this nexus between modern philosophy and legal history. See LANGBEIN, *supra* note 26; Waldman, *supra* note 18; Morano, *supra* note 19; Shapiro, *supra* note 21.

28. WILFRID E. RUMBLE, JOHN AUSTIN: THE PROVINCE OF JURISPRUDENCE DETERMINED 10 (1995); see also JEREMY BENTHAM, PRINCIPLES OF MORALS & LEGISLATION 218 (Batoche Books, ed. 2000). It would be impossible to treat this topic in summary, as the development of legal positivism was *the* central issue of general jurisprudence over the past two centuries. Leslie Green provides a helpful survey on this topic. See Leslie Green, *Legal Positivism*, STAN. ENCYCLOPEDIA PHIL. ARCHIVE, <https://plato.stanford.edu/archives/spr2018/entries/legal-positivism/> (last visited Jan. 20, 2020).

29. Shapiro, *supra* note 21, at 162 (“This reported decline in faithfulness to oaths would seem to signal a general decline in the influence of Christian fears about salvation on the behaviour of all trial participants.”); see also Waldman, *supra* note 18, at 300 (“A main facet

Which of these four factors were predominant in effecting the reform of law is a matter of some debate among legal historians.³⁰ Yet the subtleties of their archaeological arguments—however intriguing—need not detain us. There was indubitably a confluence of all four forces contributing to the pressure for legal change.

The “reasonable doubt” standard’s displacement of the “any doubt” standard was gradual. Its first occurrence has been a topic of controversy. Many legal historians locate the first recorded usage of the term “reasonable doubt” in the treason trials of Dublin in 1798.³¹ However, subsequent scholarship places its origin in the Boston Massacre trials of the 1770s—which in addition to antedating the Dublin trials, involved express litigation over the standard of proof.³² Whatever its origins, by the latter half of the eighteenth century, the ingredients for change were in the air, and it is plausible that the reasonable doubt standard emerged in multiple jurisdictions independently. The new standard percolated through the American states and British commonwealth over the next century, and it emerged the de facto standard in criminal trials by the dawn of the twentieth century.³³

of the conflict between Protestants and Catholics during and after the Reformation concerned the criteria of religious faith From attempts to answer these questions a protestant tradition developed which stressed probable rather than absolute certainty concerning articles of religious faith.”); *id.* at 301 (“The unreasonable sceptic or ‘papist’ are ignored implicitly, since they refuse to accept ‘reasonable’ arguments.”).

30. Some of the more important players in these debates have been: BARBARA SHAPIRO, *BEYOND REASONABLE DOUBT AND PROBABLE CAUSE: HISTORICAL PERSPECTIVES ON THE ANGLO-AMERICAN LAW OF EVIDENCE* (1993); BARBARA SHAPIRO, *PROBABILITY AND CERTAINTY IN SEVENTEENTH-CENTURY ENGLAND: A STUDY OF THE RELATIONSHIPS BETWEEN NATURAL SCIENCE, RELIGION, HISTORY, LAW, AND LITERATURE* (1983); Barbara Shapiro, “*To a Moral Certainty*”: *Theories of Knowledge and Anglo-American Juries 1600–1850*, 38 *HASTINGS L.J.* 153, 154 (1986); *see also* Morano, *supra* note 19; WHITMAN, *supra* note 22; HENRY VAN LEEUWEN, *THE PROBLEM OF CERTAINTY IN ENGLISH THOUGHT 1630–1690* (1963).

31. This identification seems to have originated with Leonard May suggesting Finney’s Case and Bond’s Case as the genesis of the phrase in legal doctrine. *See* Leonard Mary, *Some Rules of Evidence: Reasonable Doubt in Civil and Criminal Cases*, 10 *AM. L. REV.* 642, 656–58 (1876); *see also* Finney’s Case, 26 *How. St. Tr.* 1019 (Ire. 1798); Bond’s Case, 27 *How. St. Tr.* 523 (Ire. 1798). May’s contention gained purchase with subsequent publications, wherein the claim was repeated, leading to Justice Brennan’s majority opinion in *In re Winship*. 397 U.S. 358, 361 (1970) (“[I]ts crystallization into the formula ‘beyond a reasonable doubt’ seems to have occurred as late as 1798.”); 9 JOHN WIGMORE, *EVIDENCE* § 2497 (3d ed. 1940); CHARLES MCCORMICK, *LAW OF EVIDENCE* § 341 (2d ed. 1972).

32. Morano, *supra* note 19, at 515–16 (citing several cases using that formulation prior to the Dublin Treason trials). Morano identifies the origins of the reasonable doubt formulation in *Rex v. Preston* and *Rex v. Wemms*. *See* *THE ADAMS PAPERS: LEGAL PAPERS OF JOHN ADAMS* 46–314 (L. Kinvin Wroth & Hiller B. Zobel eds., 1965); *see also* Sheppard, *supra* note 19, at 1190.

33. Morano, *supra* note 19, at 519–27.

B. The Search for a Definition

In 1970, the Supreme Court imbued the reasonable doubt standard with constitutional imprimatur, holding it to be the requisite standard in every criminal trial, guaranteed under the due process clause of the Fifth Amendment.³⁴ The reasonable doubt standard was thus transformed from a practical expedient for achieving higher rates of conviction into a fundamental liberty owed to criminal defendants.

In his majority opinion in *In re Winship*, Justice Brennan extolled the importance of securing everywhere and evermore the utmost burden upon the prosecution.³⁵ Similar paeans to the liberal virtues of the reasonable doubt standard abound,³⁶ yet we should not forget its history. As we discussed in the foregoing section, the reasonable doubt standard evolved not from a more permissive threshold for conviction, but rather from a more stringent one. The reasonable doubt standard was not intended to bolster but rather to weaken protection for criminal defendants, instructing jurors to convict despite harboring some doubts, so long as those doubts were deemed “unreasonable.”

Yet despite Brennan’s rhetorical posture,³⁷ his opinion demonstrates a clear recognition of the motivating principle undergirding reasonable doubt: that certainty should be understood as a *spectrum* and not a set of discrete alternatives. Justice Harlan’s concurrence recognizes the principle even more explicitly.

Lamentably, the Court, even while acknowledging the probabilistic nature of judgment and extolling at length upon the value of consistency, declined to specify *what* they believed the threshold for reasonable doubt ought to be. Harlan acknowledged this difficulty, characterizing “the labels used for alternative standards of proof” as being “vague and not a very sure guide to decision making.”³⁸ Yet notwithstanding this defect, Harlan evidently believed that the holding could be useful for some purpose.³⁹ Possibly he thought it would have been at least minimally

34. *In re Winship*, 397 U.S. at 364.

35. *Id.* at 362 (“This notion—basic in our law and rightly one of the boasts of a free society—is a requirement and a safeguard of due process of law in the historic, procedural content of ‘due process.’”).

36. See, e.g., Scott Sundby, *The Reasonable Doubt Rule and the Meaning of Innocence*, 40 HASTINGS L.J. 457, 457 (1989) (citing several exuberant celebrations of the ideal).

37. In fairness to Brennan, the controversy in *In re Winship* arose in a context where the lower court had applied a preponderance standard, and the Court was indeed ratcheting *up* the standard of proof for juvenile proceedings in that case. *In re Winship*, 397 U.S. at 369.

38. *Id.* at 369–70.

39. It is possible that Harlan regarded precision to be theoretically unattainable, citing with approval the eminent evidence scholar of the first half of the twentieth century, John

useful to bound the different standards: i.e., that *whatever* reasonable doubt meant, it should be understood as something more stringent than clear and convincing evidence and preponderance of the evidence. Unfortunately, as we shall discover, even this modest hope seems disappointed in fact.

Subsequent Supreme Court jurisprudence has proven to be rather less than illuminating on the meaning of reasonable doubt. In *Jackson v. Virginia*, the Court approvingly cites a pre-*Winship* case,⁴⁰ *Holland v. United States*, in which the Court ruled upon the adequacy of a jury instruction defining reasonable doubt as “the kind of doubt . . . which you folks in the more serious and important affairs of your own lives might be willing to act upon.”⁴¹ The *Holland* court held:

We think this section of the charge should have been in terms of the kind of doubt that would make a person hesitate to act, rather than the kind on which he would be willing to act. But we believe that the instruction as given was not of the type that could mislead the jury into finding no reasonable doubt when in fact there was some. A definition of a doubt as something the jury would act upon would seem to create confusion rather than misapprehension.⁴²

However, quoting from an even earlier nineteenth century case, *Miles v. United States*,⁴³ the *Holland* court bafflingly adds, “Attempts to explain the term ‘reasonable doubt’ do not usually result in making it any clearer to the minds of the jury.”⁴⁴

We subsequently observe the “hesitate to act” gloss commonly rehearsed in pattern jury instructions. Its popularity with trial courts seems to stem from the Court’s approving stance in *Holland*. As a historical matter, the precise origin of the hesitate to act analysis is unclear. The Court points to *Bishop v. United States*,⁴⁵ which in turn cites

Wigmore: “The truth is that no one has yet invented or discovered a mode of measurement for the intensity of human belief. Hence there can be yet no successful method of communicating intelligibly . . . a sound method of self-analysis for one’s belief.” *In re Winship*, 397 U.S. at 369 (citing 9 JOHN WIGMORE, EVIDENCE 325 (3d ed. 1940)).

40. 443 U.S. 307, 317 n.9 (1979).

41. 348 U.S. 121, 140 (1954).

42. *Id.* (citation omitted).

43. 103 U.S. 304 (1880) (establishing that an improper reasonable doubt instruction can be a basis for finding error).

44. *Id.* at 312.

45. 107 F.2d 297, 303 (D.C. Cir. 1939).

to an Alabama case, *Posey v. State*.⁴⁶ In remarks critical about the instruction, Judge Newman posits *Posey* as a possible point of origin;⁴⁷ however, we have discovered similar language appearing in many court opinions (including one Supreme Court opinion) significantly antedating *Posey*.⁴⁸ We can locate no instance of the instruction earlier than the 1840s, which may have been the approximate timeframe of its genesis. In any event, the approving citation of *Holland* in *Jackson* renewed endorsement of the hesitate to act formulation, and it has since proliferated across many jurisdictions.⁴⁹

Could the hesitate to act instruction be the legal definition of reasonable doubt? Evidently not. While many jurisdictions continue to employ the language, its use has become controversial. The Federal Judicial Center Committee to Study Criminal Jury Instructions remarks in a comment to its preferred formulation:

[T]he committee has rejected the [hesitate to act] formulation because the analogy it uses seems misplaced. In the decisions people make in the most important of their own affairs, resolution of conflicts about past events does not usually play a major role. Indeed, decisions 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.⁵⁰

The First Circuit Court of Appeals has also remarked unfavorably upon the hesitate to act construction, writing, “The momentous decision to acquit or convict a criminal defendant cannot be compared with ordinary decision making without risking trivialization of the constitutional standard.”⁵¹ On similar reasoning, the Supreme Court, having occasion to reconsider its position some years after *Jackson*,

46. 93 So. 272, 273 (Ala. Ct. App. 1922).

47. Jon O. Newman, *Beyond “Reasonable Doubt”*, 68 N.Y.U. L. REV. 979, 982–83, n.14 (1993).

48. See, e.g., *Holt v. United States*, 218 U.S. 245, 254 (1910); *Pettine v. New Mexico*, 201 F. 489, 496 (8th Cir. 1912); *United States v. American Naval Stores Co.*, 172 F. 455, 456 (S.D. Ga. 1909).

49. See, e.g., COLORADO JURY INSTRUCTIONS § B:01 (MODEL CRIMINAL JURY INSTRUCTIONS COMM. COLO. SUPREME COURT 2017), https://www.courts.state.co.us/userfiles/file/Court_Probation/Supreme_Court/Committees/Criminal_Jury_Instructions/2017/COLJI-Crim%202017%20-%20Final.pdf.

50. FED. JUDICIAL CENTER, PATTERN CRIMINAL JURY INSTRUCTIONS 29 (1987).

51. *United States v. Noone*, 913 F.2d 20, 29 (1st Cir. 1990).

expressed dissatisfaction with the language⁵² and cited Judge Newman, who puzzled over the phrase:

I was always bemused by its ambiguity. If the jurors encounter a doubt that would cause them to 'hesitate to act in a matter of importance,' what are they to do then? Should they decline to convict because they have reached a point of hesitation, or should they simply hesitate, then ask themselves whether, in their own private matters, they would resolve the doubt in favor of action, and, if so, continue on to convict?⁵³

An alternative elaboration, popular until relatively recently, was to equate proof beyond a reasonable doubt with the no-less-cryptic language of "moral certainty." Historically, moral certainty was an independent standard of proof, understood to be distinct from reasonable doubt and any doubt.⁵⁴ However, it gradually merged and became synonymous with reasonable doubt by the early nineteenth century.⁵⁵ The language was common in pattern instructions until *Cage v. Louisiana*, when the Supreme Court ruled that a moral certainty jury instruction was reversible error.⁵⁶ In *Cage*, the Court also expressed disapproval for instructions which framed reasonable doubt in terms of "grave uncertainty" and "actual substantial doubt."⁵⁷

This disapproval might be taken for progress, for it would seem that after *Cage*, we have at least learned what a reasonable doubt analysis *is not*. Alas, there was no such progress, as subsequent cases would drastically weaken the holding in *Cage*. Many courts deemed the use of moral certainty language to be harmless error and, in *Victor v. Nebraska*, the Supreme Court agreed, revealing a new test: that the standard of proof instruction, taken as a whole, should not convey a concept less than due process requires.⁵⁸ And *what due process requires* is, of course, proof beyond a reasonable doubt, which the Court again elected not to analyze.

A thin straw upon which to pull us from the circularity appears in Justice Ginsburg's concurring opinion, which is comparatively responsive to scholarly complaints about the reasonable doubt standard. Ginsburg expressed displeasure with several common formulations, noting that

52. *Victor v. Nebraska*, 511 U.S. 1, 24–25 (1994) (Ginsburg, J., concurring); *id.* at 34 (Blackmun, J., concurring in part and dissenting in part).

53. *Id.* at 24–25 (Ginsburg J., concurring) (quoting Newman, *supra* note 47, at 983).

54. Sheppard, *supra* note 19, at 1195.

55. *See id.* at 1202–03.

56. 498 U.S. 39, 41 (1990) (per curiam).

57. *Id.*

58. 511 U.S. 1, 5, 22–23 (1994).

“[a]t least two of the Federal Courts of Appeals have admonished their District Judges not to attempt a definition,”⁵⁹ but acknowledging empirical research tending to show that the words “reasonable doubt” were not self-defining to laypersons.⁶⁰ Ginsburg concluded, “even if definitions of reasonable doubt are necessarily imperfect, the alternative—refusing to define the concept at all—is not obviously preferable.”⁶¹

The definition which Ginsburg preferred as “clear, straightforward, and accurate,” is Pattern Criminal Jury Instruction 21, devised by the Federal Judicial Center:

As I have said many times, the government has the burden of proving the defendant guilty beyond a reasonable doubt. Some of you may have served as jurors in civil cases, where you were told that it is only necessary to prove that a fact is more likely true than not true. In criminal cases, the government’s proof must be more powerful than that. It must be beyond a reasonable doubt.

Proof beyond a reasonable doubt is proof that leaves you firmly convinced of the defendant’s guilt. There are very few things in this world that we know with absolute certainty, and in criminal cases the law does not require proof that overcomes every possible doubt. If, based on your consideration of the evidence, you are firmly convinced that the defendant is guilty of the crime charged, you must find him guilty. If on the other hand, you think there is a real possibility that he is not guilty, you must give him the benefit of the doubt and find him not guilty.⁶²

However, it is doubtful how “helpful” Pattern Criminal Jury Instruction 21 would be to any jury facing a close case. To characterize the certainty one must possess to convict as “[being] firmly convinced” only kicks the can down the road. How convinced must one be to be “firmly” convinced? If a juror believes the defendant probably committed a crime, but harbors some doubt, it cannot be more helpful to say that his

59. *Id.* at 25 (Ginsburg, J., concurring) (first quoting *United States v. Adkins*, 937 F.2d 947, 950 (4th Cir. 1991) (“This circuit has repeatedly warned against giving the jury definitions of reasonable doubt, because definitions tend to impermissibly lessen the burden of proof.”); and then quoting *United States v. Hall*, 854 F.2d 1036, 1039 (7th Cir. 1988) (“[A]t best, definitions of reasonable doubt are unhelpful to a jury An attempt to define reasonable doubt presents a risk without any real benefit.”)).

60. *Id.* at 26.

61. *Id.*

62. FED. JUDICIAL CENTER, *supra* note 50, at 28.

belief must be “firm” or that his doubts must be “a real possibility” than to say that his doubt must be “reasonable.”

Although Ginsburg sensibly recognizes that an imperfect definition is better than no definition, it is dubious whether “firmly convinced” is, in any meaningful sense, a definition at all.

Another common gloss of reasonable doubt is the charge that the jury should acquit “if [their] minds are wavering or the scales are even.”⁶³ This formulation, which is evidently popular in New York trial courts, is generally discouraged,⁶⁴ suffering difficulties essentially similar to the hesitate to act instruction (and several more additionally). Still another common analysis is that reasonable doubt is doubt “based on reason.”⁶⁵ This surely cannot be false, though we can think of little else to recommend it.

Where does that leave us? To summarize: *In re Winship* establishes that every crime must be proven beyond a reasonable doubt.⁶⁶ What proof beyond a reasonable doubt means, the Court riddles, is something weaker than absolute certainty but stronger than clear and convincing evidence.⁶⁷ *Jackson* seems to suggest that a reasonable doubt is one which would cause a reasonable person to hesitate to act (or not act after a hesitation) in matters of importance in their own lives. *Cage* seems to suggest that defining reasonable doubt in terms of “moral certainty,” “grave uncertainty,” or “actual substantial doubt” are so egregiously incorrect as to be reversible error.⁶⁸ But then *Victor* indicates—*contra Jackson*—that *all* of the foregoing definitions are to be disfavored, but—*contra Cage*—that such disfavored language should not be grounds for reversal unless, taken as a whole, the instruction conveys a concept less than due process requires.⁶⁹ This brings us full circle, for what due process requires is, of course, proof beyond a reasonable doubt.

And just in case we should doubt whether a non-circular definition can be had, Ginsburg taunts, “[W]e have never held that the concept of reasonable doubt is undefinable.”⁷⁰

63. *People v. Bebee*, 481 N.Y.S.2d 404, 404 (N.Y. App. Div. 1984).

64. *Id.*

65. *See Johnson v. Louisiana*, 406 U.S. 356, 360 (1972).

66. 397 U.S. 358, 361 (1970).

67. *Id.* at 364.

68. *Cage v. Louisiana*, 498 U.S. 39, 41 (1990) (per curiam).

69. *Victor v. Nebraska*, 511 U.S. 1, 35 (1994).

70. *Id.* at 26.

III. INTERPRETING REASONABLE DOUBT

At the level of utmost generality, there are at least three categories of proposed interpretations of the term reasonable doubt. First, it has been proposed that a reasonable doubt is a doubt which is *articulable* (call this “articulable doubt”). Second, it has been proposed that reasonable doubt is the doubt that a “reasonable person” would entertain (call this the “reasonable person’s doubt”). Finally, it has been proposed that reasonable doubt is some probabilistic threshold (call this “probabilistic doubt”). We consider each category in turn.

A. *Articulable Doubt*

The species of articulable doubt are several.⁷¹ The simplest version is that an articulable doubt is any doubt that *could be* verbalized. This minimal interpretation is devised to exclude doubts so nebulous as to resist reduction to language.

Alternatively, we might understand an articulable doubt to be the kind of doubt that an *actual juror* would verbalize if called upon to do so.⁷² Or we might understand it as a doubt *actually* articulated—either in the jury room or by defense counsel in court. These subdivide into further variations. For example, we might further require an articulable doubt to tell a *story*, a plausible narrative of innocence.⁷³ Such narratives may include further conditions, such as logical coherence, a degree of “fit” with the evidence presented, or a degree of “relative fit” when compared with the prosecution’s narrative.⁷⁴

We need not explore all the multifarious ways an articulable doubt standard may be constructed. All interpretations of reasonable doubt qua articulable doubt—whatever the particular specifications—share a

71. Sheppard, *supra* note 19, at 1204–16 (offering an especially detailed and thoughtful analysis of the articulable doubt construction; although, in Sheppard’s framing, there is some conflation of articulable doubt and the reasonable person’s doubt, which we identify as distinct).

72. Several courts have upheld—although sometimes with criticism—the formulation that a reasonable doubt is “a doubt for which you can give a reason if called upon to do so.” *Vargas v. Keane*, 86 F.3d 1273, 1275–77 (2d Cir. 1996) (citing 1 N.Y. CRIMINAL JURY INSTRUCTION § 3.07 at 92 (1983)) (holding that a variation of New York’s pattern jury instruction was not reversible error); *see also* *State v. Coward*, 972 A.2d 691, 705 (Conn. 2009); *State v. Johnson*, 410 S.E.2d 547, 553 (S.C. 1991).

73. For example, see Michael S. Pardo & Ronald J. Allen, *Juridical Proof and the Best Explanation*, 27 L. & PHIL. 223, 238–39 (2008).

74. *See* LARRY LAUDAN, TRUTH, ERROR AND CRIMINAL LAW: AN ESSAY IN LEGAL EPISTEMOLOGY 80–86 (2006).

common defect. By requiring a doubt to be effable in order to be sufficient for acquittal, the burden of proof is ineluctably shifted to the defense.⁷⁵

A juror can have very strong doubts which are not articulable and very weak doubts which are.⁷⁶ Conspiracy theorists can certainly *articulate* reasons to doubt commonly accepted propositions (for example, that humans have walked on the moon or that Elvis Presley died in 1977). Indeed, conspiracy theories often satisfy even the more restrictive conditions on articulable doubt: they form coherent narratives, which are *consistent* with existing evidence. Yet most people would nevertheless regard most conspiracy theories as being premised upon distinctly *unreasonable* doubts.

It cannot be, therefore, that an articulable doubt is necessarily a strong doubt. Neither can it be that an inarticulable doubt is necessarily a weak doubt. We can doubt the truth of a proposition if the evidence for it is weak, even when we lack any affirmative *reasons* for believing its negation. The articulable doubt standard seems to invert the burden of proof so that it falls upon the defense to supply some articulable reasons for doubting the defendant committed the crime. Merely undermining the affirmative claims of the prosecution, without more, would not seem to suffice under an articulable doubt standard.

Surely this cannot be the result proponents of an articulable doubt analysis intend. Yet it is an inescapable consequence of the formulation unless we radically redefine “articulable” to the point it no longer resembles its ordinary meaning. This alone is reason to reject the articulable doubt analysis.

It is worth acknowledging that there is an elegance to the proposition that reasonable doubt should be analyzed literally as *a doubt for which reasons can be given*. But it is fundamentally the wrong question for the

75. Not all courts regard this as sufficiently troubling to merit reversal. *See Vargas*, 86 F.3d at 1275–77; *Coward*, 972 A.2d at 705; *Johnson*, 410 S.E.2d at 553. However, some courts have taken notice of the burden-shifting problem. *See, e.g., Dunn v. Perrin*, 570 F.2d 21, 24 (1st Cir. 1978) (citations omitted) (“That definition of reasonable doubt was the exact inverse of what it should have been. Instead of requiring the government to prove guilt, it called upon petitioners to establish doubt in the jurors’ minds.”); *United States v. MacDonald*, 455 F.2d 1259, 1263 (1st Cir. 1972) (“Individual jurors were not charged with either articulating a supportable *ratio decidendi*.”); *see also Adams v. South Carolina*, 464 U.S. 1023, 1025 (1983) (Marshall, J., dissenting from denial of certiorari) (“[W]hen a jury is told that a reasonable doubt is a doubt that can be articulated, the prosecutor’s burden of proof is unconstitutionally eased.”); *Butler v. South Carolina*, 459 U.S. 932, 934–35 (1982) (Marshall, J., dissenting from denial of certiorari).

76. This point was observed by the Supreme Court of New Jersey: “[The jury instruction] erroneously implies that the jury must find an articulable reason to support its doubts about the State’s case. Jurors may harbor a valid reasonable doubt even if they cannot explain the reason for the doubt.” *State v. Medina*, 685 A.2d 1242, 1246 (N.J. 1996) (citations omitted).

purpose to ask merely whether one *can express* reasons for doubting. The central issue in a criminal trial is not whether a juror's doubt possesses certain formal properties (i.e., being expressible), but rather whether the *magnitude* of that doubt is sufficiently serious. We do not believe—and should not want—jurors to vote to acquit simply because they can concoct an articulable reason for disbelief. Neither do we believe—nor should we want—jurors to vote to convict simply because they *lack* the expressive resources to communicate a rationale to back sincerely felt reservations.

Now, the proponent of an articulable doubt standard might concede that it is the magnitude of doubt which matters after all but still contend that articulable doubt can serve as a kind of proxy. Strong doubts *tend* to be articulable, and inarticulable doubts *tend* to be weak. If we had no better measure of dubiousness than the expressibility of doubt, then the case for reasonable doubt qua articulable doubt might still be argued viable.

Yet even under a proxy theory, the articulable doubt analysis still fails. For any assertion that a defendant committed some criminal act, it is always *possible* to articulate reasons to doubt. For example, in any given case, it is *articulable* that all the witnesses were hallucinating, that police investigators framed the accused, or that the criminal act was committed by the defendant's unknown evil twin. Such counter narratives certainly count as *reasons* for doubting, and it seems self-evident that an infinitude of such counternarratives *can be* articulated.

Thus, articulable doubt cannot serve as a proxy because articulable doubts can be constructed for every possible case. And this cannot be what reasonable doubt means if it is to be distinguished from the any doubt standard. For articulable doubt to have any bite at all, it must require not only that a doubt be articulable, but also that it be actually entertained by the fact-finder.⁷⁷

77. Surprisingly, courts which otherwise tolerate barer formulations of articulable doubt instruction have seemed to balk at the inclusion of additional such criteria. For example, in *People v. Antommarchi*, the Court of Appeals of New York, while generally approving of a “doubt for which a juror can give a reason” formulation, found impermissible the following instruction: “[I]f you have a reasonable doubt, I repeat, a reasonable doubt, on any relative point or material element or on the evidence or lack of it, and when one or more of your fellow jurors questioned you about it, you would be willing and able to give him what you believe is a fair, claim explanation for your position based upon the evidence or the lack of evidence in this particular case.” 604 N.E.2d 247, 251–52 (Ohio 1992). Similarly, the Second Circuit in *Chalmers v. Mitchell* found the emendation that reasonable doubt is a “doubt for which some *good* reason can be given” to be error, albeit harmless error, following *Victor v. Nebraska*. 73 F.3d 1262, 1266 (2d Cir. 1996) (emphasis added); see *Victor v. Nebraska*, 511 U.S. 1, 25 (1994). In *Humphrey v. Cain*, the Fifth Circuit reversed the lower court on the basis of an instruction including a “good reason” criterion. 138 F.3d 552, 554 (5th Cir. 1998) (en banc).

But if we define reasonable doubt as an articulable doubt which the fact-finder actually entertains, we cannot avoid the follow-on question: *how strongly* should the doubt be entertained? Howsoever we set the threshold, we will be left in practice to rely upon jurors' self-reflective evaluations on the gravity of their doubts. Yet if we think jurors' introspections—i.e., their self-assessments of their belief levels—are a reliable measure of certainty, then what purpose does the proxy variable serve? If the end result of the articulable doubt analysis ultimately relies upon jurors' self-reflection, then the only effect of adopting the "articulable" criterion is to exclude strong but inarticulable doubts. In the end, the desirable traits of the articulable doubt standard are redundant, and only the negative byproduct remains.

To put it more simply, articulable doubt—once the idea is fully cashed out—seems ultimately to require that we entrust the measurement to jurors' introspective judgments of their certainty. But if the whole point of articulable doubt were to devise a proxy for introspective judgments of certainty, and if the proxy ultimately requires those very same introspective judgments of certainty, then there is no purpose for using a proxy. Adding the condition that doubts be articulable contributes nothing to the process except the undesirable side effect of excluding the subset of doubts which are reasonable but which a juror lacks the articulateness to express.

As a last-ditch effort, proponents of the articulable doubt analysis might contend that the articulable condition helps to *guide* jurors in their introspections. The contention would be that articulation serves as some sort of a psychological aid in the process of introspection. No evidence has ever been proffered in the legal literature that articulation would have this effect in fact. But even if it did, the proper role for it would be as mere guidance. In the best case, articulable doubt could serve as a helpful suggestion to jurors; but it could not rise to the role of a legal definition or a charge to the jury.

B. Reasonable Person's Doubt

It has also been proposed that reasonable doubt might be understood as the doubt that a reasonable person would harbor if presented with evidence for a claim. The principal problem here is that the proposed gloss is grossly indeterminate. Does it mean that *any* doubt that a reasonable person would entertain is perforce a reasonable doubt? Or does it designate a *sufficiently serious* doubt that a reasonable person would have? If the former, then the line separating reasonable doubt from any doubt would seem to be very thin indeed. If the latter, then once

again we are left to ponder *how serious* the reasonable person's doubt should be in order to qualify as a reasonable doubt.

Let us first consider the reasonable person's doubt as being any doubt that a reasonable person would entertain. Imagine a case goes to trial and, after receiving all of the evidence, the reasonable juror is left 99.99% certain that the defendant committed the crime alleged by the prosecution. Does the 0.01% *doubt of the reasonable juror* qualify as a reasonable doubt because the juror is a reasonable person? If it were a 0.001% doubt or a 0.0001% doubt—no difference? Are we to understand that any doubt—however minuscule—that the reasonable juror would entertain is a reasonable doubt *ipso facto*? And should this alone suffice to acquit?

We think no one would endorse such an interpretation.⁷⁸ A more sensible version of the reasonable person's doubt might be to parse *the reasonable person's doubt* as being that which would motivate a reasonable person to acquit. Yet this is plainly question-begging. It says nothing more than that the reasonable doubt standard is the standard (assuming there is only *one*) that a reasonable person would adopt. It leaves wholly untouched the question *why* the reasonable person should prefer any standard over any other. And more seriously, it fails to specify in any meaningful or tractable way which standard the reasonable person would choose.

This variant of the reasonable person's doubt analysis seems to place the standard itself within the ambit of "fact questions," yet it is a question the litigants are not on notice (nor permitted) to address with evidence.⁷⁹ More troublingly still, mock jury experiments seem to indicate that the way laypeople interpret the term reasonable doubt (both with and without pattern instructions) varies wildly, ranging anywhere from *less than 50%* and *up to 100%* certainty with a high degree of variance. The extreme heterogeneity may also be observed in the wildly divergent averages found between studies.⁸⁰ To the extent the reasonable person is

78. *But see* State v. Jones, 582 So. 2d 973, 975 (La. Ct. App. 1991).

79. *See* Erik Lillquist, *Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability*, 36 U.C. DAVIS L. REV. 85, 87–88 (2002). This topic is discussed in greater detail *infra* Section III.C.2.

80. For an extremely convenient compilation of this line of research, see Reid Hastie, *Algebraic Models of Juror Decision Processes*, in INSIDE THE JUROR: THE PSYCHOLOGY OF JUROR DECISION MAKING 84 (Reid Hastie ed., 1993), and especially *id.* at 102 tbl.4.1 (comparing studies using direct ratings), *id.* at 103 tbl.4.2 (comparing studies using parallel ranking), and *id.* at 105 tbl.4.3 (comparing studies using decision theory). *See also* Mandeep K. Dhami et al., *Instructions on Reasonable Doubt: Defining the Standard of Proof and the Juror's Task*, 21 PSYCHOL. PUB. POL'Y & L. 169, 169 (2015); Irwin A. Horowitz & Laird C. Kirkpatrick, *A Concept in Search of a Definition: The Effects of Reasonable Doubt Instructions on Certainty of Guilt Standards and Jury Verdicts*, 20 L. & HUM. BEHAV. 655,

meant to signify an objective measure, he must be some idealized representative—a Hercules—for there is no point of convergence apparent in the opinions of real ordinary persons. There is much disagreement among the passengers of the Clapham omnibus.

Law-trained individuals are only marginally more consistent when offering quantifications of the reasonable doubt standard. This difference may be due to the different methodologies typically used to extract a quantification from subjects. Mock jury estimations of reasonable doubt are inferred from behavior in an experimental setting, whereas judges and lawyers tend to be asked directly via survey.

Nevertheless, significant inconsistency still exists. One survey asking judges to estimate in probabilistic terms the meaning of reasonable doubt found an average of approximately 90% certainty, although the range of responses was large, with one judge placing it as low as 50%, fourteen judges responding 80%, fifty-six judges responding 90%, thirty-one judges responding 95%, and twenty-one judges responding 100%. Interestingly, the judges surveyed did not necessarily favor round numbers, with one judge apiece responding 92%, 93%, 94%, and 97%.⁸¹ Other judicial surveys have found similarly wide variations in judicial quantifications, although a modal clustering around 90% seems to be normal.⁸²

For the foregoing reasons, it is dubious whether the *doubt which would motivate a reasonable person to acquit* can be a useful definition. Let us next consider an analysis of the reasonable person's doubt as being the *reasonable doubt that a reasonable person would have*. Obviously, there is an alarming circularity in this definition. Yet courts and scholars have endorsed this analysis with apparent seriousness. For example, the Appellate Division of New York upheld a jury instruction defining reasonable doubt as "the reasonable doubt of a reasonable man or woman."⁸³

Professor Sheppard offers a perspicuous general analysis of the semantic circle:

The need to define limit₁ (doubt) by limit₂ (reasonable) by limit₃ (grave) or limit₄ (common sense) or limit₅ (articulable: that you can say it aloud) is a self-referential game. Once we move beyond one limit, the justifications for each limit are the other

656–57 (1996); Daniel B. Wright & Melanie Hall, *How a "Reasonable Doubt" Instruction Affects Decisions of Guilt*, 29 BASIC & APPLIED SOC. PSYCHOL. 91, 91–93 (2007).

81. McCauliff, *supra* note 11.

82. See, e.g., *United States v. Fatico*, 458 F. Supp. 388, 410 (E.D.N.Y. 1978); Simon & Mahan, *supra* note 11.

83. *People v. Gonzales*, 77 A.D.2d 654, 657 (N.Y. App. Div. 1980) (Lazer, J., dissenting).

justifications. . . . [T]he law is drawn to use reason to justify and control its delegation of power in the juror. But this justification must itself be justified, hence giving rise to the smoke and mirrors about what “reasonable” must mean here. These further justifications have no essential basis, and eventually the screen fades to black, with the law having made no better claim either for the juror to make decisions of guilt or for limits of reason on that power than the simple assertion that jurors have that power.⁸⁴

The point is perhaps made rather more acutely if we add the alternative “or limit (reasonable person).”

To draw out the intuition more clearly still, notice what happens when using the proposed definition to substitute equivalent terms. If reasonable doubt were defined as the *reasonable doubt that a reasonable person would have*, then it follows by substitution that reasonable doubt means the *reasonable doubt that a reasonable person would have that a reasonable person would have*. And repeating the substitution, by transitivity, it follows that reasonable doubt means the *reasonable doubt that a reasonable person would have that a reasonable person would have that a reasonable person would have*. Such a definition is clearly question begging.

Putting circularity problems aside for the moment, we acknowledge that the reasonable person construction has a clear intuitive appeal. The term *reasonable person* is a ubiquitous (if somewhat ambiguous) legal locution. Its origin is indeterminate. Many sources identify its genesis in the tort classic, *Vaughan v. Menlove*,⁸⁵ citing Dean Prosser’s *Law of Torts*.⁸⁶ However, this misidentification seems to be the result of a typographical error. Prosser cites *Menlove* as having been decided in “1738.” The case was in fact decided in 1837. Several subsequent articles have followed Prosser in identifying *Menlove* as the earliest appearance of the “reasonable person,” repeating the erroneous dating.⁸⁷ However,

84. Sheppard, *supra* note 19, at 1235–36.

85. (1836) 173 Eng. Rep. 232, 232.

86. William L. Prosser, LAW OF TORTS 150 nn.14 & 16 (4th ed. 1971) (“[The ‘reasonable man of ordinary prudence’ was first mentioned in *Vaughan v. Menlove*, 1738. . . .”).

87. See, e.g., Morano, *supra* note 19, at 514 n.51; see also Dolores A. Donovan & Stephanie M. Wildman, *Is the Reasonable Man Obsolete? A Critical Perspective on Self-Defense and Provocation*, 14 LOY. L.A. L. REV. 435, 435 n.2 (1981); Christ Gaetanos, *Essay—Assumption of Risk: Casuistry in the Law of Negligence*, 83 W. VA. L. REV. 471, 485 n.49 (1981); Gary L. Putnam, *The State of Legal Malpractice in Oklahoma*, 9 TULSA L.J. 129, 133 n.16 (1973).

we have found several cases applying the reasonable person standard, using that terminology, prior to 1837 (though none before 1738).⁸⁸

Whatever its radix, the hypothetical beliefs and behavior of the reasonable person have come to serve as the catchall standard, upon which the law routinely heaps its thorniest line-drawing problems. Wheresoever we should look in the law—from the interpretation of contractual language, to the duty of care, to an employee’s perceptions of sexual harassment—the reasonable person lurks not far. Booting the standard of proof problem to the reasonable person will therefore strike lawyers as a familiar and thematically consistent maneuver.

Yet, however tempting it may be to boot the definition of reasonable doubt to the reasonable person, such a move would first have to overcome circularity problems discussed earlier. We do not mean to imply that these problems cannot be resolved. It is conceivable that a better formulation of the reasonable person’s doubt analysis could be devised. Our remarks are intended merely to point out that attempts to accomplish this have, *as yet*, failed to produce a non-circular, meaningful analysis. This does not foreclose future innovations, and we think that some construction of the reasonable person’s doubt analysis could be viable, if the problems raised in this section could be satisfactorily addressed.

Finally, it is worth remarking that the reasonable person’s doubt can—and has been—combined with other analyses in formulating jury instructions. For example, “doubt for which a reasonable man can give a reason,” which seems to merge articulable doubt and the reasonable person’s doubt, was deemed to be an acceptable gloss by the Fifth Circuit in *Bernstein v. United States*.⁸⁹ And as we have seen already, the hesitate to act gloss is also sometimes expressed in reference to the reasonable person—specifically, what might cause the *reasonable person* to hesitate.⁹⁰

88. The earliest usage we have found appears in an admiralty case concerning negligence:

If the conduct of the Respondent was not wilful and with full knowledge, yet it appears to us to have been a crassa negligentia, and that any *reasonable man*, upon inquiry, and the least reflection, upon reading the orders given to the Prize master M’Neil, (and he ought to have read them) or upon the circumstances attending the whole transaction, must have been satisfied, that the Betsy was a prize to the Argo. . . . *Reasonable care, attention, prudence, and fidelity, are expected from the master of a ship, and if any misfortune or mischief ensues from the want of them, either in himself or his mariners, he is responsible in a civil action.*

Purviance v. Angus, 1 U.S. 180, 184–85 (1786) (emphasis added).

89. 234 F.2d 475, 487 (5th Cir. 1956).

90. *Victor v. Nebraska*, 511 U.S. 1, 24–25 (1994) (Ginsburg, J., concurring).

C. Probabilistic Doubt

A third category of approaches to interpreting reasonable doubt is to assign a precise probabilistic threshold as the standard of proof.⁹¹ For example, that a criminal charge is proven beyond a reasonable doubt if and only if the juror believes with greater than $z\%$ certainty that it is true. Probabilistic doubt avoids many of the problems encountered in the investigation of articulable doubt and reasonable person's doubt. It does not invoke a proxy concept, nor boot the problem, nor risk circularity. It directly states the numerical measure of the relevant variable: the magnitude of jurors' requisite certainty level for conviction. However, probabilistic doubt entails fresh complications which demand further investigation.

1. Juror Responsiveness

Let us first address a practical problem: whether specifying a certainty threshold $z\%$ is a meaningful instruction to jurors. Regardless whether we posit that the threshold of belief should be 80% certainty, 95% certainty, or whatever specific number—would such an instruction be operationalizable? There are several compelling reasons to think it might not be. Behavioral economists have discovered, in supernumerary experimental settings, that human decision-making is exceedingly poor when confronting problems involving quantified probabilities. Indeed, the research suggests not only a high rate of error but *systematic* deviation over large samples.⁹² Consequently, instructing a juror to apply an 80% certainty threshold may not result in the application of an 80% certainty threshold in real cases.

The problem is potentially serious, yet not necessarily fatal. First, mock jury experiments show that jurors are at least *responsive* to quantified standard of proof instructions (and relatively *unresponsive* to verbal instructions).⁹³ For example, when given the same set of facts, mock jurors were observed to be more likely to convict when they were told to apply an 80% threshold than if they were told to apply a 90% threshold.

91. It is worth mentioning that the "threshold" approach—of assigning full criminal liability when the evidence suggests greater than $z\%$ probability of guilt—is not the only possible framework for adjudication. See Talia Fischer, *Conviction Without Conviction*, 96 MINN. L. REV. 833, 879–80 (2012) (arguing for a tailoring of sanction severity to epistemic certainty along a continuum rather than discrete outcomes).

92. See Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124, 1130–31 (1974).

93. Nagel, *supra* note 15, at 195; see also Kagehiro & Stanton, *supra* note 11.

Second, to the extent that the *real* threshold jurors apply diverges from the theoretical threshold prescribed by law, there exist many correctives available to the policymaker to remedy the discrepancy. For example, the crudest fix would simply be to account for the expected cognitive distortion when specifying the certainty threshold. Suppose it were known that jurors systematically undershot, such that prescribing a threshold of $z\%$ certainty tended to result in the application of a $0.9z\%$ certainty threshold in practice. Thus, for instance, an instruction that jurors should convict if they are more than 80% certain the defendant is guilty would result in jurors applying a 72% certainty threshold in fact, due to the effect of the cognitive bias ($0.8 \times 0.9 = 0.72$).

Such a regularity can easily be manipulated to achieve the desired effects. Suppose our policy goal were to have a reasonable doubt standard of 80%. Assuming that the real rule applied by jurors is $0.9z\%$ when a certainty threshold of $z\%$ is given, we simply anticipate the effect and instruct jurors to apply a threshold of 88.8% certainty in order to achieve the desired 80% threshold—i.e., assuming that jurors would undershoot the instructed threshold by a fixed deviation, $0.\bar{8} \times 0.9 = 0.8$ gets us to the desired standard of proof.⁹⁴

This strategy is generalizable. Any *systematic* deviation in jurors' application of quantified instructions can be "hacked." If, on the other hand, the effect of cognitive bias is *not systematic* in this way, then it may not be so simply combatted, but it also becomes less clear why we should regard it as seriously problematic in the first place.

Another mechanism for combatting problems encountered due to cognitive bias are "debiasing strategies." Broadly defined, debiasing strategies are policies designed to communicate information or present choices so as to reduce the effects of framing.⁹⁵

94. Of course, our example is deliberately simplistic. Our intention is merely to illustrate the principle of compensating for bias. An earnest attempt to implement such a strategy would involve more sophisticated models of subjective perceptions of probability. See, e.g., R. Duncan Luce & Peter C. Fishburn, *Rank- and Sign-Dependent Linear Utility Models for Finite First-Order Gambles*, 4 J. RISK & UNCERTAINTY 29, 32–33 (1991); JOHN QUIGGIN, GENERALIZED EXPECTED UTILITY THEORY 28–31 (1993); Amos Tversky & Daniel Kahneman, *Advances in Prospect Theory: Cumulative Representation of Uncertainty*, 5 J. RISK & UNCERTAINTY 297, 300–01 (1992).

95. Attempts to debias jurors using jury instructions has not—in experiments conducted with mock juries—yielded encouraging results, although insulating strategies have demonstrated some success. See Allison C. Smith & Edith Greene, *Conduct and Its Consequences: Attempts at Debiasing Jury Judgments*, 29 L. & HUM. BEHAV. 505, 523–24 (2005) (noting the experiment conducted in the context of the reasonableness of defendant relating to tort negligence). See generally Daniel Pi et al., *Biasing, Debiasing, and the Law*, in THE OXFORD HANDBOOK OF BEHAVIORAL ECONOMICS AND THE LAW 143 (Eyal Zamir & Doran Teichman eds., 2014) (describing techniques for debiasing and insulating non-

Yet *even if* offsetting the effect of bias and debiasing strategies were ineffective, there are still several reasons to think probabilistic articulations of reasonable doubt are preferable to the alternatives. First, mock jury experiments reveal that providing test subjects with quantified reasonable doubt instructions *at least* results in the desired relative effects. In an important study, Professors Kagehiro and Stanton found that jurors were practically incapable of distinguishing between “preponderance of the evidence,” “clear and convincing evidence,” and “beyond a reasonable doubt” when given exclusively verbal instructions.⁹⁶ Altering the verbal jury instructions for each of the three standards while holding fixed the facts of a mock case, Kagehiro and Stanton observed no significant effect on juror decision-making.⁹⁷ Indeed, what little effect *was* observed ran counter to the putative objectives of the law—jurors behaved as if reasonable doubt were a weaker standard than clear and convincing evidence, and as if clear and convincing evidence were weaker than preponderance of the evidence. By contrast, quantified instructions elicited the intended separation.

Instruction	Plaintiff/Prosecution Verdicts
Unquantified	
Preponderance	31%
Clear & Convincing	38%
Reasonable Doubt	43%
Quantified	
Preponderance	66%
Clear & Convincing	52%
Reasonable Doubt	31%

Table 1⁹⁸

Interestingly, even when verbal instructions were *combined* with quantified measures, the consistency of juror decision-making was hindered rather than improved.⁹⁹ Only probabilistic instructions *alone*

rational individuals in a legal setting); Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1527 (1998) (providing a broad overview of the applicability of behavioral economics to legal analysis).

96. Kagehiro & Stanton, *supra* note 11, at 172.

97. *Id.* at 164–65. Kagehiro & Stanton used the California and *Addington v. Texas* definitions in the first of their four experiments described in their article. *Id.* at 163 (citing 441 U.S. 418 (1979)); *see also* CAL. JURY INSTRUCTIONS—CRIMINAL [CALJIC] § 2.90.

98. Kagehiro & Stanton, *supra* note 11.

99. See the second experiment in *id.* at 169 (“Combining the legal and quantified definitions of standards of proof in the same jury instructions did not remedy the ineffectiveness of the legal definitions, perhaps because the legal definitions were presented

(i.e., without verbal supplementation) produced behaviors most consistent with what the law prescribes.¹⁰⁰ Other studies obtained similar results.¹⁰¹ In the absence of a probabilistic analysis, jurors seem to be utterly incapable of distinguishing even the coarse-grained distinctions the law has established between preponderance of the evidence, clear and convincing evidence, and reasonable doubt. This does not seem to be a deficiency in the particular formulations devised by courts, but rather a defect of *unquantified instructions* as a class. It is unlikely that tweaking the wording would have any effect because the evidence suggests that verbal formulations of the standard of proof, by their very nature, have *no* measurable effect whatsoever on jurors.¹⁰²

If quantified instructions only effected a sorting at the broadest level of analysis, then this would still represent a substantial improvement over present practice. The principle that the law should be consistent—that it should treat like cases alike—is assuredly among the law’s most fundamental objectives. And ultimately the magnitude of such inconsistency as has been observed must surely outweigh whatever marginal deviations might arise due to cognitive bias. Whatever else one might say about how badly jurors understand numerical probabilities, it is a well-documented fact that they are poorer still at understanding verbal instructions.

Courts have evinced diffidence in acknowledging that present formulations of standards of proof could fail so abysmally to communicate a meaningful distinction to jurors. In a surprising exception to the general silence, the Supreme Court in *Addington v. Texas* attempted to provide some account.¹⁰³ Justice Burger opined:

Candor suggests that, to a degree, efforts to analyze what lay jurors understand concerning the differences among these three tests or the nuances of a judge’s instructions on the law may well be largely an academic exercise; there are no directly relevant empirical studies. Indeed, the ultimate truth as to how the standards of proof affect decisionmaking may well be

first.”). Surprisingly, in the combined treatment, jurors treated reasonable doubt as being a *weaker* standard than clear and convincing. *Id.*

100. *Id.* (showing that verdicts for plaintiffs decreased as the standard increased from preponderance, to clear and convincing, to reasonable doubt).

101. See, e.g., Dorothy K. Kagehiro, *Defining the Standard of Proof in Jury Instructions*, 1 PSYCHOL. SCI. 194, 196 (1990); L.S.E. Jury Project, *supra* note 11, at 220 (showing that jurors who were given a corroborating warning, meant to decrease certainty, convicted more quickly).

102. See Joel D. Lieberman & Bruce D. Sales, *What Social Science Teaches Us About the Jury Instruction Process*, 3 PSYCHOL. PUB. POL’Y & L. 589, 598–99 (1997).

103. 441 U.S. 418, 423–25 (1979).

unknowable, given that factfinding is a process shared by countless thousands of individuals throughout the country. We probably can assume no more than that the difference between a preponderance of the evidence and proof beyond a reasonable doubt probably is better understood than either of them in relation to the intermediate standard of clear and convincing evidence. Nonetheless, even if the particular standard-of-proof catchwords do not always make a great difference in a particular case, adopting a “standard of proof is more than an empty semantic exercise.” In cases involving individual rights, whether criminal or civil, “[t]he standard of proof [at a minimum] reflects the value society places on individual liberty.”¹⁰⁴

Candor indeed! Burger adds in a footnote, “There have been some efforts to evaluate the effect of varying standards of proof on jury factfinding, . . . but we have found no study comparing all three standards of proof to determine how juries, real or mock, apply them.”¹⁰⁵ We should wonder how the Court would respond to Kagehiro and Stanton’s findings. If it were to insist upon Burger’s position in *Addington*, it would be difficult to regard its posture as anything less than monstrous: that we should be indifferent to the *actual effect* of jury instructions—even when effective alternatives are known—merely to pretend at some grotesque theater of justice. Though there is undoubtedly some value in the expressive use of law—over and above an “empty semantic exercise”¹⁰⁶—if this were *all* the law intended for standards of proof, then it should be a deeply troubling admission.

Yet even if jurors were capable of appreciating probabilistic doubt instructions, some scholars have raised additional objections to the use of numerical probabilities to define the reasonable doubt standard. One concern is that the use of a probabilistic standard of proof could invite more widespread use of *statistical evidence* in trials.¹⁰⁷ Professor Tribe forcefully argued against the use of probabilistic methods in legal procedure on this ground, enumerating several plausible ways ordinary people might err in reasoning about statistical information.¹⁰⁸

104. *Id.* at 424–25 (footnote omitted) (quoting *Tippett v. Maryland*, 436 F.2d 1153, 1166 (4th Cir. 1971)).

105. *Id.* at 424 n.3 (citation omitted).

106. *Id.* at 425.

107. Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1359–60 (1971).

108. *Id.* at 1329, 1331–32. Although the issue here is miscalculation rather than cognitive bias, the division is porous. *Id.* at 1340 n.40. In addition to Tversky & Kahneman, we would be remiss not to include some mention of the seminal observations of systematic

Tribe's objection is substantially similar to the problem of cognitive bias. Let us consider a "worst case" scenario to illustrate Tribe's concern: suppose the prosecution in a murder trial introduced evidence tending to show that killers were slightly more likely to be left-handed than individuals in the general population. And suppose the defendant were left-handed. The unsophisticated juror, incapable of appreciating the meaning of the statistical claim, might accord this evidence undue weight and thereby convict a defendant.

This certainly does *feel* problematic. But why? Is the problem that the prosecutor introduced statistical evidence or that the statistical evidence contributed to the verdict? Tribe takes the position that both are problems,¹⁰⁹ but this seems inaccurate. Consider: a juror with sufficient sophistication reasoning about probabilities would not have accorded the handedness of the defendant very much weight. It is only because the juror is, by the premise of the hypothetical, unsophisticated and incapable of appreciating the meaning of the statistical claim, and that he accorded the evidence *undue* weight that we have an issue. The problem is not that the statistical evidence was presented or considered but, rather, that it was treated ineptly. And if this is the problem, then we think it is not after all so different from concerns about cognitive bias. It is superficially distinct but fundamentally the same objection—jurors are bad with probabilities.

Parsing Tribe's criticisms somewhat more finely, another dimension along which we could distinguish his objection is that cognitive bias tends to be regarded as an unconscious effect, whereas the kinds of errors Tribe describes are the result of conscious (albeit erroneous) deliberation.¹¹⁰ In other words, Tribe's complaint is not merely that the typical juror possesses perverse probabilistic intuitions, but rather that the typical juror is also *unskilled* at deliberate mathematical reasoning.¹¹¹ It is the difference between misunderstanding and miscalculation.

The empirical evidence on whether laypeople typically lack the skills required to reason statistically is mixed.¹¹² However, tools exist to

statistical error. See M. Allais, *Le Comportement de l'Homme Rationnel Devant le Risque: Critique des Postulats et Axiomes de l'École Américaine*, 21 *ECONOMETRICA* 503, 508 (1953); Daniel Ellsberg, *Risk, Ambiguity, and the Savage Axioms*, 75 *Q.J. ECON.* 643, 643 (1961).

109. Tribe, *supra* note 107, at 1332.

110. *Id.* at 1371 (discussing how a "factual presumption of guilt" may force jurors to suspend the presumption of innocence until conclusion of the trial).

111. See *id.* at 1388–89.

112. See, e.g., David L. Faigman & A.J. Baglioni, Jr., *Bayes' Theorem in the Trial Process: Instructing Jurors on the Value of Statistical Evidence*, 12 *L. & HUM. BEHAV.* 1, 1, 7 (1988) (finding that jurors incorporated Bayesian updating, albeit with systematic bias); D.H. Kaye & Jonathan J. Koehler, *Can Jurors Understand Probabilistic Evidence?*, 154 *J. ROYAL STAT. SOC'Y* 75, 79–80 (1991) (finding attenuated updating); William C. Thompson, et al.,

constrain the risk of miscalculation—as long as the problem is miscalculation rather than unconscious bias. Indeed, it is *easier* to correct an unskilled juror than a cognitively biased juror because they can simply be educated on how to properly reason about probabilities. Improper statistical inference may be corrected through the use of expert testimony. For example, if a scientific test were known to be 99% accurate and it implicated the defendant in a crime, the expert testifying to the result could be obliged by the court to disabuse jurors of the mistaken inference that the test implied a 99% probability of guilt.¹¹³ Moreover, attorneys for the parties also have incentives to guard jurors from conclusions they might improperly infer from statistical evidence. This is precisely the effect we expect adversarial litigation to have with respect to *non*-statistical inference, and there is no principled reason why the same principles should not apply to statistical inference.¹¹⁴

In any case, it should be noted that Tribe's main concern about a probabilistic *standard of proof* relates not to the standard itself, but rather to its tendency to cause an increase in the introduction of probabilistic evidence in trials.¹¹⁵ This worry is somewhat moot, because the introduction of statistical evidence has become quite common in modern criminal litigation, despite the use of non-probabilistic jury instructions on the standard of proof. In other words, the putatively bad consequences Tribe feared have happened anyway. If we should continue to resist probabilistic doubt, it cannot be because it will have some bad secondary effect but, rather, because it is inherently problematic.

Tribe does discuss some reasons why a probabilistic standard of proof should be disfavored, independent of its supposed propensity to encourage the use of statistical evidence. First, Tribe was concerned that jurors' use of Bayesian priors would tend to undermine the presumption

Do Jurors Give Appropriate Weight to Forensic Identification Evidence?, 10 J. EMPIRICAL LEGAL STUD. 359, 359 (2013) (finding test subjects were approximately consistent with Bayesian expectations).

113. The probability is Bayesian and depends on the prior probability.

114. The questions—whether jurors understand statistical evidence, and how best to present statistical evidence so as to be understood by jurors—motivate a substantial body of research. The present consensus—if there is one—seems to be that statistical evidence *can* be meaningfully understood by jurors, contrary to Tribe's assertions. How best to communicate that information remains a topic of ongoing research. See, e.g., C. Aitkin, et al., *Expressing Evaluative Opinions: A Position Statement*, 51 SCI. & JUST. 1 (2011); Jane Goodman, *Jurors' Comprehension and Assessment of Probabilistic Evidence*, 16 AM. J. TRIAL ADVOC. 361, 375–76 (1992); K.A. Martire, et al., *On the Interpretation of Likelihood Ratios in Forensic Science Evidence: Presentation Formats and the Weak Evidence Effect*, 240 FORENSIC SCI. INT'L 61, 61–62 (2014); see also Faigman & Baglioni, *supra* note 112, at 13–14; Kaye & Koehler, *supra* note 112, at 80; Thompson et al., *supra* note 112, at 375.

115. Tribe, *supra* note 107, at 1330.

of innocence.¹¹⁶ This concern is easily rebutted. It should be observed that subjective rational choice theory posits that jurors will behave *as if* they were relying upon priors regardless whether they are invited to employ an explicitly statistical framework. Tribe is simply mistaken about what subjective rational choice theory claims about behavior when he contends that the absence of an explicitly mathematical framework would eliminate dependence upon prior probabilities in the formation of judgments.¹¹⁷ More specifically, he is mistaken in treating the claims about the representability of juror decision-making as claims about the deliberative processes of juror decision-making.¹¹⁸ And he is further mistaken in assuming that a decision-maker who is not given an explicitly numerical prior probability will employ *no* priors in decision-making.¹¹⁹ All jurors will behave as if their judgments depended on priors, regardless of whether they are given an explicitly probabilistic standard of proof instruction. Assuming priors without examination offers no better protection against injustice than a blindfold shields against a firing squad.

Second, making explicit the treatment of prior probabilities should be regarded as desirable even by Tribe's own normative criteria. It provides occasion for the court to mitigate (or at least condemn) the use of invalid priors—for example, assumptions about race, gender, or religious belief—in the formation of jurors' judgments. Indeed, the very *meaning* of the "presumption of innocence" may be concretized in the articulation of a sufficiently low Bayesian prior. Failing to acknowledge prior probabilities does not eliminate their use. Treating them explicitly allows courts to address the very objections Tribe contemplates.

Third, even if it were granted *arguendo* that jurors were pathologically incapable of reasoning probabilistically about *evidence*, this would not be an argument against a probabilistic definition of the *standard of proof* directly. There is no logical reason why the courts could not forbid the introduction of statistical evidence while explaining the standard of proof in probabilistic terms. As we have previously discussed, the Kagehiro and Stanton experiments demonstrate that probabilistic instructions improve juror consistency, independently of the nature of the evidence is presented in trial.¹²⁰

116. *See id.* at 1358.

117. *See id.* at 1354.

118. *See id.*

119. *See id.*

120. Kagehiro & Stanton, *supra* note 11.

We cannot respond to the very many other specific complaints Tribe raises,¹²¹ which would take us far afield of our present inquiry; however, with respect to the criminal trials in particular, Tribe's concern seems broadly to be that quantifying the reasonable doubt standard would tend to undermine defendants' fundamental protections. All of the data we have found suggests the opposite effect.¹²²

It is possible that it is not the *effect* but rather the *principle* implicit in a probabilistic quantification which is what *really* bothered Tribe.¹²³ The proposition that the law should explicitly acknowledge a tolerable rate of error in a matter which requires *no* doubt would, he claimed, be anathema to the principles of criminal justice.¹²⁴ This is mistaken on two grounds. First, as we have discussed, the reasonable doubt standard arose in contradistinction to the any doubt standard, and it is inherent in the historical usage of it that *some* uncertainty *is* acceptable. Second, the practical effect is that *not* giving a quantified standard of proof tends to favor the prosecution. We imagine few criminal defendants would object to the indignity of the courts expressly articulating an acceptable error distribution, if it improved their prospects of an acquittal in actual practice.

On the question of feasibility therefore, it seems that probabilistic doubt is at least no worse than current jury instructions, and potentially a good deal better (depending on the extent to which jurors are biased, and our capacity to mitigate such biases).

2. Frequency and Certainty

Let us now consider, if a standard of proof instruction were to specify a numerical threshold—*z*% certainty—what the numerical value of *z* should be. Two obvious possibilities avail themselves immediately: (i)

121. Other authors have persuasively taken up this task. *See, e.g.*, Michael O. Finkelstein & William B. Fairley, *A Comment on Trial by Mathematics*, 84 HARV. L. REV. 1801, 1801 (1971).

122. *See, e.g.*, Kagehiro & Stanton, *supra* note 11, at 175; Nagel, *supra* note 15, at 195; Michael J. Saks & Robert F. Kidd, *Human Information Processing and Adjudication: Trial by Heuristics*, 15 L. & SOC. REV. 123 (1980) (arguing that the absence of a quantified standard of doubt renders jurors more susceptible to cognitive biases and increases conviction rates).

123. *See* Tribe, *supra* note 107, at 1373–74.

124. *Id.* at 1374 (“The jury is charged that any ‘reasonable doubt,’ of whatever magnitude, must be resolved in favor of the accused. Such insistence on the greatest certainty that seems reasonably attainable can serve at the trial’s end, like the presumption of innocence at the trial’s start, to affirm the dignity of the accused and to display respect for his rights as a person—in this instance, by declining to put those rights in deliberate jeopardy and by refusing to sacrifice him to the interests of others.”).

that z should be whatever threshold affects *efficient* outcomes (call this “efficient z ”) and (ii) that z should be determined by legal doctrine (call this “doctrinal z ”).

Let us begin with efficient z . A tempting basis for balancing social costs is Blackstone’s maxim that “the law holds, that it is better that ten guilty persons escape, than that one innocent suffer.”¹²⁵ Let us refer to all such schemata, identifying a ratio of type-I errors to type-II errors, as “Blackstonian ratios” (which we denote $\beta \leq 1/x$).

The broadly economic motivation, to achieve an optimal Blackstonian ratio, is expressed in Justice Harlan’s concurrence in *In re Winship*:

Because the standard of proof affects the comparative frequency of these two types of erroneous outcomes, the choice of the standard to be applied in [criminal or civil cases] should, in a rational world, reflect an assessment of the comparative social disutility of each. . . . In a criminal case . . . we do not view the social disutility of convicting [sic] an innocent man as equivalent to the disutility of acquitting someone who is guilty.¹²⁶

Legal scholars have long recognized the logical relationship between Blackstonian ratios and the reasonable doubt standard.¹²⁷ It is often implicit when scholars or judges claim that the reasonable doubt threshold should be set at 90% or 91% certainty of guilt. This number is the result of calculating the specific ten-to-one odds articulated by Blackstone, such that $10:1 = \frac{10}{11} = .\overline{90}$, or equivalently plugging $\beta = 1/10$ into Formula 1 below:¹²⁸

$$z = \frac{1}{\beta + 1}$$

This straightforward algebraic manipulation converts a Blackstonian ratio into a standard of proof. However, there are several reasons to be skeptical of such a conversion. First, if the objective of the criminal law were to effect an error distribution of $x:1$, then the z produced by Formula 1 would produce this result only if there were no intervening effects. For

125. BLACKSTONE, *supra* note 12.

126. 397 U.S. 358, 371–72 (1970).

127. See, e.g., *United States v. Fatico*, 458 F. Supp. 388, 411 (E.D.N.Y. 1978); Jonathan Levy, *A Principled Approach to the Standard of Proof for Affirmative Defenses in Criminal Trials*, 40 AM. J. CRIM. L. 281, 290–92 (2013); Michael S. Pardo, *Evidence Theory and the NAS Report on Forensic Science*, 2010 UTAH L. REV. 367, 370–71 (2010).

128. This is simply $\beta = 1/z - 1$, solving for z . See John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065, 1073 (1968); see also *Fatico*, 458 F. Supp. at 411.

example, given a pool of defendants whose cases all warranted a 50% certainty of guilt, we would have to assume that this certainty level implied that half the defendants were guilty and that half were innocent.

This is a potentially dubious assumption for several reasons. First, evidentiary rules forbid jurors from hearing relevant evidence, i.e., evidence which ought to affect their certainty level if “its probative value is substantially outweighed by a danger of . . . unfair prejudice, confusing the issues, misleading the jury, undue delay, [or] wasting time.”¹²⁹ Although these insulating strategies are valuable, we should be cognizant of their cost: withholding relevant information will create a wedge between juror certainty and the frequency of guilt. And it is easy to imagine the accumulation of many such evidentiary exclusions resulting in substantial deviation between jurors’ subjective certainty of guilt and the objective frequency of guilt. Additionally, the prosecutorial discretion whether to pursue cases and asymmetric detection rates could also affect prior probabilities of guilt. This information is typically not available to jurors and so their certainty levels will not account for these filtering effects.

How seriously we should take the divergence between subjective certainty and objective frequency is debatable. Professor Lillquist suggests a rather extreme wedge, positing the following hypothetical distribution:

Certainty of Guilt	Innocent Defendants	Guilty Defendants
.05	3	4
.15	3	4
.25	3	4
.35	2	8
.45	2	8
.55	2	8
.65	2	8
.75	1	12
.85	1	12
.95	1	12

Table 2¹³⁰

129. FED. R. EVID. 403.

130. Lillquist, *supra* note 79, at 101. In fairness, Lillquist provides the following table as an example, and it is not obvious how seriously we should take the particular values given as an approximation of reality. Nevertheless, he does provide this table in contradistinction to an earlier set of data which he characterizes as unrealistic, implying that he regards these values as being at least a plausible possibility.

The point of Lillquist's example is to demonstrate that a 70% certainty threshold, applied to the pool of defendants represented in Table 2, would result in three false convictions and forty-six false acquittals, nearly twice as favorable to defendants as a 10:1 ratio.¹³¹ This hypothetical is meant to highlight the disconnect between adopting a 90% reasonable doubt threshold and ensuring a 10:1 error distribution. His example demonstrates how, in some circumstances, a 70% threshold could generate a better than 10:1 distribution of error (whereas a 90% threshold would result in an even more defendant friendly 68:1 error distribution).

But it is debatable how plausible a hypothetical this is. The distribution of innocent and guilty defendants that Lillquist gives in Table 2 is very strange. We should be reluctant to accept that a juror would possess a 5% certainty level for three innocent defendants and four guilty defendants (the first row of Table 2).¹³² Absent intervening effects, we would expect the ratio of innocent defendants to guilty defendants, for which the certainty of guilt is 5%, to be approximately 20:1.¹³³ The reasons for believing that such an extreme disconnect between jurors' subjective certainty and objective frequency would arise in reality is largely unargued. Lillquist writes, "it is possible that the plea bargaining process eliminates more guilty defendants than innocent ones (either because of the strength of the evidence or because innocent defendants are less likely to agree to a plea, knowing that they are innocent)."¹³⁴

This is an extravagant claim. It is implausible that plea bargaining would result in such a drastic redistribution of guilty and innocent defendants at the trial stage. However, *even if* this were the case, it is unclear why it should matter. Suppose the prosecutor, anticipating a certainty threshold of 90%, only charged cases wherein he felt the evidence supported a 90% probability of guilt. If there were, for example, 100 prospective defendants who satisfied the criterion, then we might imagine ninety-one defendants are guilty, and nine are innocent. Suppose further that 100% of the guilty defendants accepted a plea deal and 100% of the innocent defendants elected to go to trial. Assuming then that plea bargaining were a perfect filter, 100% of the defendants—all nine—who go to trial would be innocent. But recalling that we started with the assumption that the prosecutor only charged suspects for whom the evidence supported a 90% or greater probability of guilt, all of the

131. *Id.*

132. *Id.*

133. This is the very *meaning* of a probabilistic measure of certainty. In the absence of intervening effects, there is no distinction between subjectivist and frequentist probability.

134. Lillquist, *supra* note 79, at 101.

innocent defendants would be found guilty. What is the net effect? Nine innocent defendants would be punished, having been found guilty at trial; and ninety-one guilty defendants would be punished, having accepted a plea deal. After all this, the end result is an error distribution which satisfies the desired 10:1 ratio.

Nevertheless, the point remains that it is, in principle, possible for jurors' subjective certainty to diverge from objective frequency of guilt, and the deviation may be sufficiently serious to cause concern. If subjective certainty diverges significantly from objective frequency, then the standard of proof cannot be expected to achieve the desired error ratio. Although we are skeptical whether Lillquist has identified the most likely causes of such divergence, the end effect is not implausible. This would tend to undermine the normative basis for the relationship between Blackstonian ratios and standards of proof.

However, there is another way of reading Blackstone's maxim. The formulation—that it is better ten guilty persons escape than that one innocent suffer—is ambiguous. It could be understood as expressing a desired frequency outcome. Or it could be understood as pronouncing a normative claim: that the *social cost* of convicting one innocent person is equal to the *social cost* of acquitting ten guilty.

3. Ratio of Social Costs

In the late 1960s, researchers began exploring the second interpretation of Blackstone's maxim, i.e., as specifying the relative disutility of type-I and type-II errors (rather than the relative frequency of errors).¹³⁵ A majority of scholars advancing probabilistic doubt analyses have since followed this approach.¹³⁶ Justice Harlan's opinion, quoted earlier, also seems to adopt such an interpretation, explicitly employing the language of "social disutility" (although he confusingly mentions "comparative frequency" in the very same sentence).¹³⁷

If we understand Blackstonian ratios as expressing *error disutility* rather than *error frequency*, then it is a natural next step to formulate the standard of proof in terms of efficiency. Let us denote the utility of a

135. Alan D. Cullison, *Probability Analysis and Judicial Fact-Finding: A Preliminary Outline of the Subjective Approach*, 1 U. TOL. L. REV. 538, 557–59 (1969); John Kaplan, *Decision Theory and the Fact-finding Process*, 20 STAN. L. REV. 1065, 1071 (1968).

136. Michael L. DeKay, *The Difference Between Blackstone-Like Error Ratios and Probabilistic Standards of Proof*, 21 L. & SOC. INQUIRY 95, 118–26 (1996); Nagel, *supra* note 15, at 191–95; *see, e.g.*, Lillquist, *supra* note 79, at 147–62.

137. *In re Winship*, 397 U.S. 358, 371 (1970) ("Because the standard of proof affects the comparative frequency of these two types of erroneous outcomes, the choice of the standard to be applied in [criminal or civil cases] should, in a rational world, reflect an assessment of the comparative social disutility of each." (emphasis added)).

false conviction by $U(CI)$, a false acquittal by $U(AG)$, a correct conviction by $U(CG)$, and a correct acquittal by $U(AI)$. One proposed formulation of the socially optimal standard of proof given by scholars advancing subjectivist efficiency analyses is:¹³⁸

$$z = \frac{1}{\frac{U(CG) - U(AG)}{U(AI) - U(CI)} + 1}$$

The meaning of a Blackstonian ratio, under this interpretation, is simply that $\beta = \frac{U(AG)}{U(CI)}$.¹³⁹ Scholars advocating such a formulation have regarded statements of Blackstonian ratios insufficient to determine the value of z because the statement of β by itself leaves indeterminate the requisite values of $U(CG)$ and $U(AI)$.¹⁴⁰

Yet this is an overhasty conclusion. We can *infer* that the benefit of a correct acquittal is equivalent to the cost of a false conviction and that the benefit of a correct conviction is equivalent to the cost of a false acquittal. What, after all, is the cost of a false conviction but the *forgone benefit* of a correct acquittal? And likewise, what is the cost of a false acquittal but the *forgone benefit* of a correct conviction? In other words, $U(AG)$ and $U(CI)$ may be understood as representing opportunity costs. Understood thusly, Formula 2 is double counting and Formula 1 should be regarded the correct formulation of the efficient standard of proof, assuming the further premise that $\beta = \frac{U(AG)}{U(CI)}$.

The foregoing considerations are potentially progress. However, Professor DeKay rejected arguments of a substantially similar nature, pointing out several problems with the social cost minimization formulation of efficient z .¹⁴¹ First, DeKay believes that social cost interpretations fail to track what the utterers of Blackstonian ratios intended.¹⁴² Second, he objects to the simplification (from Formula 2 to Formula 1) on the ground that setting $U(CG) = U(AI) = 0$ is methodologically improper.¹⁴³ Third, he points out that social welfare maximization still requires jurors to estimate probabilities of guilt.¹⁴⁴ Although we may determine that social welfare is maximized when jurors acquit for probabilities less than z and convict for probabilities

138. DeKay, *supra* note 136, at 111.

139. In other words, the Blackstonian ratio is a ratio of disutilities of the errors rather than a ratio of probabilities of the errors. For example, $\beta = \frac{U(AG)}{U(CI)}$, as opposed to, $\beta = \frac{p(AG)}{p(CI)}$.

140. See DeKay, *supra* note 136, at 111; Lillquist, *supra* note 79, at 108.

141. DeKay, *supra* note 136, at 110–18.

142. *Id.* at 112–15.

143. *Id.* at 115–17.

144. *Id.* at 111.

greater than z , the same aforementioned wedge would prevent jurors from maximizing social welfare for the very same reasons it prevented jurors from effecting the desired error frequency.

The first objection is arguable. Statements of Blackstonian ratios do genuinely seem, as we noted earlier, ambiguous as between the error frequency interpretation and social cost interpretation. It is not *obvious* that the utterers of Blackstonian ratios intended to express a relationship between type-I and type-II errors; therefore, it does not seem unreasonable to explore the alternative interpretation. The second objection has less to do with the assumption itself but, rather, the absence of a principled justification for it; however, the opportunity cost rationale we have suggested avoids the problem of arbitrariness.¹⁴⁵ The third objection is a more serious concern. Jurors can maximize *decision utilities* (or something functionally equivalent to that concept) despite having incomplete probabilistic information. But unless those decision utilities approximate objective expected utilities, jurors' efforts to maximize will not result in efficient outcomes in actual fact. That is a problem.

The issue boils down to whether the maximization of decision utilities can stand on its own as a policy objective, independent of its consequences. There are surely reasons to be skeptical whether it can do so. What ordinarily motivates a utility maximization approach is that it tends to produce desirable *outcomes*. When that result cannot be guaranteed, then one could argue that the process is defective.

Several points can be made in defense of what we might call *procedural utility maximization* (i.e., the making of decisions calculated to maximize subjective utility). First, in at least *some* cases, the divergence between subjective probability and objective probability will be inconsequential. Even if it is *possible* for a utility-maximizing procedure to result in inefficient outcomes, the mere *tendency* to produce relatively more efficient outcomes can still justify the use of those procedures. Second, utility maximization embeds the normative objective of reducing the social cost.¹⁴⁶ This analysis *expresses* more than merely setting a target distribution of type-I and type-II errors. Even though the outcomes will be identical as between a distribution-of-errors objective and a utility maximization objective, the latter does more than merely prescribe an outcome. The analysis *says* something about what the standard of proof is meant to achieve.

However, the *main* advantage of the social cost interpretation, as compared with the error frequency interpretation, is that it invites

145. See *id.* at 115–17.

146. See Francesco Parisi, *Positive, Normative and Functional Schools in Law and Economics*, 18 EUR. J.L. & ECON. 259 (2004).

empirical inputs. The error frequency interpretation smacks of arbitrariness. If we understand Blackstone's maxim as merely expressing a desired ratio of type-I and type-II errors, then there is no sensible way of disagreeing. It tells us nothing, other than that William Blackstone wanted that ratio 10:1. By contrast, the social cost interpretation frames the ratio in terms which can be analyzed. The social cost of type-I and type-II errors invites empirical investigation into the real-world social costs of type-I errors and the real-world social costs of type-II errors. Thus, the ratio is not merely a normative decree but, rather, an empirical claim about what processes would effect a maximization of social welfare.

Of course, this does not negate the problem that, *even if* it were known what the efficient standard of proof were, the wedge between subjective probability and objective probability will tend to interrupt the production of an efficient outcome. It would not be an unreasonable position, therefore, to consider the social cost interpretation of Blackstonian ratios to be an insufficient ground upon which to base so important a principle as the criminal standard of proof. However, in jettisoning this approach also, one ought to be mindful of the dearth of reasonable alternatives.

We arrive finally at the question—if we are willing to assume a social cost interpretation of Blackstonian ratios and if we are willing to assume that Formula 1 embeds opportunity costs—what is the efficient value of β ? Clearly, answering this would require a great deal of empirical research. However, it is a practicable task to determine the relevant values that would allow good approximations of efficient β . That value, plugged into Formula 1 would return the standard of proof which would be “efficient” in criminal trials if the evidence were sufficient to determine guilt probabilities and the wedge between subjective and objective probabilities were sufficiently small.

4. Heterogeneity and Tailoring the Standard

We should now mention an important corollary to this point. Consider that $U(AG)$ and $U(CI)$ could vary by offense. For example, the social cost of wrongly acquitting a rapist is surely greater than the social cost of wrongly acquitting a vandal; and the social cost of wrongly convicting a defendant of rape is surely greater than the social cost of wrongly convicting a defendant of vandalism. Given that these costs vary by offense, it would be somewhat surprising if the efficient ratio remained constant over offense types.

Indeed, a colorable argument may be made that the ratio of disutilities varies not only from offense to offense, but from case to case.

It is upon this insight that Lillquist argues *in favor of* “confusing” verbal jury instructions and *against* probabilistic doubt.¹⁴⁷ He reasons that if the efficient β varies from case to case, then the efficient standard of proof will also vary from case to case.¹⁴⁸ And if the efficient standard of proof varies from case to case, then the jury instructions in any particular trial will require a *tailored* probabilistic threshold.¹⁴⁹ This, Lillquist argues, is unfeasible.¹⁵⁰

Although Lillquist does not state his argument explicitly, we can reasonably reconstruct the gist of his justification: adopting efficient standards of proof—specifying in each case the efficient probabilistic threshold for conviction—would incur litigation costs of such magnitude as to render the policy net social welfare reducing. Lillquist argues therefore that the use of vague instructions licenses the jury to determine efficient β for themselves.¹⁵¹ If this is a task for which jurors are comparatively less skilled, Lillquist seems to maintain, it is still an acceptable second-best solution.¹⁵²

We should be deeply skeptical of whether “confusing language” will have the effect Lillquist assumes,¹⁵³ and it does not seem to be borne out in experimental observations. As we remarked earlier, mock jury experiments have found that jurors seem to vary wildly in their estimation of the standard of proof when given vague and confusing instructions.¹⁵⁴ Suppose we grant, *arguendo*, Lillquist’s claim that jurors *should* vary the standard of proof across cases. If jurors spontaneously discover efficient β in each case, we should expect their treatment of β to remain consistent, from juror to juror, *within the same case*. Yet mock jury experiments demonstrate precisely the opposite: that β varies dramatically from juror to juror even when they are looking at the *same set of facts*.¹⁵⁵ This seriously undermines his claim that vague jury instructions on the meaning of reasonable doubt would tend to cause jurors to gravitate to whatever standard of proof happens to be efficient in a particular case.¹⁵⁶

147. Lillquist, *supra* note 79, at 175–76.

148. *Id.* at 171.

149. *Id.* at 170–71.

150. *Id.*

151. *Id.* at 162 (“[A]t least in the abstract, we might prefer a reasonable doubt standard that varies depending upon the case. . . . [T]he existing reasonable doubt standard, with its confusing language, is well-designed to allow this to happen.”).

152. *See id.*

153. *Id.* at 162.

154. *See* discussion *supra* Section III.C.1.

155. *See, e.g.,* Kagehiro & Stanton, *supra* note 11, at 164–65; Kagehiro, *supra* note 101, at 195; Lieberman & Sales, *supra* note 102, at 597–99.

156. Lillquist, *supra* note 79, at 171.

Some further objections may be raised to Lillquist's premises. First, although it is implausible that the ratio of the disutility of type-I and type-II errors remains constant across offense types, there is reason to suppose that such variations might be bounded within an acceptably narrow range. If sanction severity were proportional to harm across offense types, then it is a reasonable hypothesis that the cost of false acquittals would tend to increase with the cost of false convictions. Thus, efficient β 's, though varying from case to case, may well tend to converge upon some modal value.

Second, to the extent that criminal law's objective is deterrence, the case-to-case variation that motivates Lillquist's argument may simply be irrelevant. The expected sanction attached to a proscribed activity is likely to be understood by a prospective offender in terms of offense categories and not in terms of the particular manner or context in which the offender commits the offense. Moreover, even if marginal variations in the distribution of disutilities existed within offense types, it is far from obvious that these would be significant. Even if they were, and even if juries *did* spontaneously apply an efficient β (or some approximation of it), this adjustment would be invisible to prospective criminals, undermining any potential deterrent effect. To the extent that the criminal law should be responsive to the eccentricities of each case, it seems the better tool for that task would be to adopt flexible sentencing rules rather than a variable standard of proof.

Third, if we want the standard of proof to vary depending upon the facts of each case, it is curious why we should not simply *instruct* the jury to make a finding of the efficient standard of proof on those facts. In other words, it is unclear why we should favor a *confusing* standard of proof instruction if the goal is to encourage jurors to determine the efficient standard for themselves. Surely it would be better to tell jurors *expressly* that they should undertake this task. Lillquist anticipates this objection, even offering a facially acceptable instruction for such a charge.¹⁵⁷ But Lillquist rejects his own hypothetical instruction, arguing that it would invite attorneys to litigate the issue.¹⁵⁸ However, it is unclear why this would be an undesirable effect. If the law treated the standard of proof

157. *Id.* at 187 ("The government, represented by the prosecutor in this case, has the burden of establishing and proving the guilt of the defendant beyond a reasonable doubt. If, after you have heard all of the evidence and applied to it the rules of law on which I have instructed you, you are not convinced of the defendant's guilt beyond a reasonable doubt, you must acquit the defendant. If, on the other hand, you are convinced of the defendant's guilt beyond a reasonable doubt, you must convict the defendant. The question, naturally, is what is a reasonable doubt? It represents the degree of certainty that you believe should be required of the government in this case.")

158. *Id.* at 188.

as a question of fact, it would be strangely inconsistent to design a system where *that crucial fact* could not be the subject of litigation. Disallowing litigation over that fact question undermines the fact-finder's ability to weigh the relevant evidence.

5. Merging Doctrinal β and Efficient β

Efficiency is not the only possible source of a normative β . Indeed, from a doctrinal perspective, the efficient value of β would not matter if it were already determined as a matter of law. If the most recent opinion of the highest court in a jurisdiction held, for example, that $\beta = 1/999$ or $\beta = 1/3.1415$. . . , then so it would be until the day the holding were overturned.

Yet we may reasonably wonder how seriously such pronouncements ought to be taken. Are declarations that it is better to acquit x guilty than to convict one innocent truly *holdings* or merely *obiter dicta*?

We think they are clearly holdings. Recall the definitions that students are taught in their first year of law school. Translated literally, "obiter dictum" is any part of the judicial opinion which is "said in passing."¹⁵⁹ It is "unnecessary to the decision in the case and therefore not precedential."¹⁶⁰ By contrast, the "ratio decidendi" is that part of the judicial opinion which is "the reason for deciding."¹⁶¹ It is "[t]he principle or rule of law on which a court's decision is founded."¹⁶²

When a question relating to the standard of proof is raised, and a court recites or paraphrases a Blackstonian ratio, it is typically to identify the principle as a *justification* for adopting a relatively high burden of proof.¹⁶³ It is both in form and in use the kind of thing which ought not be regarded as mere *dictum*, but properly *ratio decidendi*.

Yet some readers may still insist that we are taking too seriously the offhand remarks of grandiloquent judges, which are *dicta* in spirit, if not strictly *dicta* by the dictionary definition. It is far from clear that judges do not intend their pronouncements of Blackstonian ratios to be taken seriously.¹⁶⁴ But let us suppose,

159. *Obiter Dictum*, BLACK'S LAW DICTIONARY (10th ed. 2014).

160. *Id.*

161. *Ratio Decidendi*, BLACK'S LAW DICTIONARY (10th ed. 2014).

162. *Id.*

163. For an illustrative example, see *State v. Thoss*, 120 N.E.3d 1274, 1284 (Ohio Ct. App. 2018) (citing *Lamprecht v. State*, 95 N.E. 656, 660 (Ohio 1911)) (providing authority to support $\beta = 1/99$ as the reason for ordering a retrial).

164. See, e.g., *State v. Adkins*, 96 So. 3d 412, 434 (Fla. 2012) (Perry, J., dissenting) ("The majority opinion breaks that sacred law [the Blackstonian maxim] and, as discussed below, threatens bedrock principles of the presumption of innocence and burden of proof in contexts well beyond the one at hand.").

arguendo, that we are taking them more seriously than they intended.¹⁶⁵ So what?

As we have discussed, it is well established in experimental researches that the probabilistic approach to reasonable doubt enjoys several important advantages. Regardless of their intention, judicial articulations of reasonable doubt can provide a tractable mechanism for taking probabilistic doubt seriously in practice. And the establishment of *formal* imprimatur is precisely the point. It does not matter whether judges *really* meant them to be taken seriously but, rather, whether they can be persuaded that the *ought* to have meant them seriously.

The reader who would insist upon the disingenuousness of judges quoting Blackstone, Hale, or Fortescue misses the point. The practice of law is ever changing to better fit the poetry of our rhetoric than the reality of its intention. Our treatment of the Equal Protection and Due Process Clauses has evolved, not to maximize fidelity to their original meaning, but rather to maximize social welfare. It is a bizarre contention that the practice of law should not conform to our rhetoric because the rhetoric was intended merely to express an ideal. To take seriously what was meant as patriotic puffery or judicial bombast is neither misguided nor trifling, but rather the very hallmark of legal creativity.

The interpretation of legal language is a kind of game, extracting rules from linguistic expressions. If a better rule can be extracted than the utterer of the expression contemplated, why should anyone object? It is truly a humorless and unimaginative grumbler who would complain that judges did not intend articulations of Blackstonian ratios to be taken so seriously.

Yet we may still wonder, when a judge declares it better that x guilty go free than that one innocent suffer punishment unjustly, from what source does he divine x ? Linking together our earlier discussion on efficient β and our present discussion of doctrinal β , it is a plausible interpretation of such proclamations as judicial “best guesses” as to the value of *efficient* β . There of course exists a substantial—albeit controversial—literature arguing for the proposition that judges deliberately and self-consciously seek efficient rules.¹⁶⁶ Even to the extent that judges are guided by conceptions of morality, this too may tend toward efficiency, as it is frequently urged that moral principles are simply social norms which evolved via processes of natural selection to

165. Although, at least some judges clearly do understand the mathematical implications of Blackstonian ratios. See *United States v. Fatico*, 458 F. Supp. 388, 411 (E.D.N.Y. 1978).

166. See RICHARD POSNER, *ECONOMIC ANALYSIS OF LAW* 26 (9th ed. 2014).

combat implicit market failures.¹⁶⁷ In this case, the distinction between a recital of moral principle and a best guess at efficiency reduces to a merely intensional—i.e., not extensional—distinction.¹⁶⁸

Interpreting judicial articulations of Blackstonian ratios as *judicial estimates* of the efficient β reconciles some theoretical tensions latent in the probabilistic doubt analysis. Though the residue of legal formalism remains soaked in the fabric of legal practice, legal scholarship has, since the rise of American legal realism, rightly regarded purely legalistic reasoning with disfavor. Our present understanding of the law is that it *does* and *should* look to policy rationales—the incentive effects of the law and not the “internal logic” inherent in a set of doctrines—for its justification. If we regard judicial articulations of Blackstonian ratios as estimates of efficient β , and the determination of efficient β relies upon contact with the real world, then realist objections to taking judicial articulations of β seriously may be somewhat quieted. After all, it is a plausible point of departure to suppose that judges would have some sense of what an efficient β might be, inferring inductively from their acquaintance with many like cases over many years of experience.

However, there are some obstacles to accepting the merger of efficient z and doctrinal z which are worth highlighting. First, there exist alternative explanations why judges might specify one Blackstonian ratio rather than another. A judge may simply be recalling the formulation he was taught as a law student—whether Blackstone’s, Fortescue’s, Hale’s, or that of some other author. He may have simply picked a large-seeming integer arbitrarily. Or he may have been influenced by the score of his favorite football team’s last game.¹⁶⁹ More research is warranted, yet we think differences in judicial articulations may plausibly be found to correlate with those factors relevant to determining efficient β .

Additionally, it is disputable whether judges’ acquaintance with many like cases over many years of experience would allow them

167. A representative example is Edward O. Wilson, *The Biological Basis of Morality*, ATLANTIC MONTHLY, April 1998, at 53, 53. See generally EDWARD O. WILSON, CONSILIENCE: THE UNITY OF KNOWLEDGE (1998) (developing a wide-ranging general theory of culture grounded in evolutionary biology); EDWARD O. WILSON, SOCIOBIOLOGY: THE NEW SYNTHESIS (1975) (a systematic defense of sociobiology from critics in the social sciences).

168. On the evolution of social norms, see, e.g., Elinor Ostrom, *Collective Action and the Evolution of Social Norms*, 14 J. ECON. PERSP. 137, 144 (2000); Francesco Parisi, *The Genesis of Liability in Ancient Law*, 3 AM. L. & ECON. REV. 82, 82 (2001); Richard A. Posner, *A Theory of Primitive Society, with Special Reference to Law*, 23 J.L. & ECON. 1, 4–5 (1980).

169. A disturbing yet evidently genuine possibility. See Naci Mocan & Ozkan Eren, *Emotional Judges and Unlucky Juveniles*, 10 AM. ECON. J.: APPLIED ECON. 171, 172 (2018).

any better access to estimating the efficient value of β than a lay person. Obviously, the most useful information would be the *actual* guilt or innocence of defendants in trials he has overseen, however judges have no better access to this information than the public at large. Moreover, even if trial court judges were imbued with such instinct, the *binding* estimates of efficient β would not be theirs, but rather those of appeals court judges who would not have necessarily had the exposure to the trial-level data to develop that instinct.

It cannot be said whether probabilistic doubt is the *best* analysis of reasonable doubt. However, it seems to be the best among the alternatives yet conceived. The several conceptual and theoretical issues one might raise against the probabilistic doubt approach are serious but addressable. At a minimum, probabilistic articulations of the reasonable doubt standard will tend to effect the desired separation between the different standards of proof (i.e., reasonable doubt, clear and convincing evidence, and preponderance of the evidence). The experimental studies suggest moreover that probabilistic analyses would tend to reduce juror confusion and ensure better consistency over like cases.¹⁷⁰ We have remarked on how probabilistic articulations of reasonable doubt could even prove to be a useful policy lever for achieving social welfare maximizing incentives.

Critically, there exist precedential bases for resisting proscriptions of probabilistic doubt instructions. As we have discussed, one plausible interpretation of judicial pronouncements of Blackstonian ratios—i.e., as statements of relative social costs—implies precise numerical thresholds for the standard of proof. Read in this light, the law does not unequivocally reject a probabilistic doubt analysis. Rather, it contradicts itself.

Faced with the contradiction, it is unclear whether judges would affirm a stronger commitment to their Blackstonian ratio pronouncements or whether they would double down on a policy of systematic opacity. For all the reasons we have heretofore discussed, the former seems to be the wiser alternative. Yet it suffices for the purpose of legal argument that quantifying doubt probabilistically in jury instructions does have a precedential ground; and this is enough—as against other analytical approaches—to accord the probabilistic doubt analysis special consideration.

The present state of reasonable doubt analysis is severely problematic. At the level of utmost generality, there is broad agreement. Apart from Lillquist,¹⁷¹ we have located no other scholarship defending

170. Kagehiro & Stanton, *supra* note 11, at 169 tbl.3.

171. Lillquist, *supra* note 79, at 87–88.

the current state of reasonable doubt jurisprudence. Change is surely wanted. Among the proposed solutions, the social cost interpretation of probabilistic doubt seems the most tractable analysis of the reasonable doubt standard yet proposed. We posit that taking seriously judicial articulations of Blackstonian ratios as precedential estimates of the efficient probabilistic doubt standard represents a doctrinally acceptable path to realizing such change.

Nevertheless, it should be observed that cases might arise for which a numerically precise probabilistic standard would not prove especially helpful. Therefore, even if probabilistic articulations should be preferred generally, there may still be space for soft definitions in idiosyncratic cases where probabilistic articulations would be more likely to confuse than clarify the jurors' task. We see no reason why an informal equivalent, stated in terms of the reasonable person for example, could not be deployed as an alternative formulation, if a non-circular alternative could be devised.

Of course, operationalizing a concept with multiple equivalent analyses is not a novel conceptual move. An obvious exemplar may be seen in the law of torts, where the reasonable person standard is given dually as $B < PL$,¹⁷² and also as the hypothetical behavior of the reasonable person. The two articulations, though conceptually distinct, are understood as representing a single standard. To introduce a similar bifurcation for the reasonable doubt standard would thus not be an unprecedented maneuver in the design of legal standards.

IV. THE REASONABLE DOUBT STANDARD BY JURISDICTION

If judicial pronouncements of Blackstonian ratios are to be taken seriously, then we might naturally inquire what exactly those pronouncements are. In 1997, Professor Volokh undertook a survey of cases in the fifty states and federal courts, locating judicial statements of Blackstonian ratios in each jurisdiction.¹⁷³ Now more than twenty years-old, Volokh's data warrants an update. Table 3 proves a handy starting point for attorneys and judges interested in attempting to utilize such an approach in future criminal adjudication.

172. The "Hand Formula" was articulated by Judge Learned Hand in *United States v. Carroll Towing Co.* See 159 F.2d 169, 173 (2d Cir. 1947).

173. Volokh, *supra* note 14, at 201.

JURISDICTION	LOWER BOUND	REASONABLE DOUBT
Alabama ¹⁷⁴	$\beta \leq 1/5$	83%
Alaska ¹⁷⁵	$\beta \leq 1$	50%
Arizona ¹⁷⁶	$\beta \leq 1$	50%
Arkansas ¹⁷⁷	$\beta \leq 1/5$	83%
California ¹⁷⁸	$\beta \leq 1/10$	90%
Colorado ¹⁷⁹	$\beta \leq 1$	50%

174. *Ex parte* Mauricio, 523 So. 2d 87, 92 (Ala. 1987) (quoting *People v. Galbo*, 112 N.E. 1041, 1044 (N.Y. 1916) (Cardozo, J.)). However, Alabama law has not historically been clear on this point. See Volokh, *supra* note 14, at 202 nn.251–54 (discussing Alabama’s historical treatment of the Blackstonian ratio); see also *Morris v. State*, 60 So. 3d 326, 373–74 (Ala. Crim. App. 2010) (holding that failure to include a statement of $\beta \leq 1/100$ in a jury instruction was not error); *Jackson v. State*, 432 So. 2d 504, 508 (Ala. Crim. App. 1983) (holding the Blackstonian formulation to be “merely” an abstract principle of law without connection to the case).

175. Alaskan courts have consistently stated $\beta \leq 1$. See *Smart v. State*, 146 P.3d 15, 30–31 (Alaska 2006) (citing *In re Winship*, 397 U.S. 358, 362–64 (1970)); *Shaw v. Dep’t of Admin.*, 861 P.2d 566, 570 (Alaska 1993) (citing *In re Winship*, 397 U.S. at 372 (Harlan, J., concurring)); *State v. Alto*, 589 P.2d 402, 406 n.16 (Alaska 1979) (citing *In re Winship*, 397 U.S. at 370–72 (Harlan, J., concurring)).

176. *Am. Pepper Supply Co. v. Fed. Ins. Co.*, 93 P.3d 507, 509 (Ariz. 2004) (en banc) (citing *In re Winship*, 397 U.S. at 372–73 (Harlan, J., concurring)).

177. *Dunaway v. Troutt*, 339 S.W.2d 613, 620 (Ark. 1960) (stating “several”); *Jones v. State*, 320 S.W.2d 645, 649 n.6 (Ark. 1959) (citing *Coffin v. United States*, 156 U.S. 432 (1895)) (making a possible distinction in cases of double jeopardy, when “some” is used).

178. California courts have given a variety of mixed signals with respect to the Blackstonian ratio. For example, *Salisbury v. County of Orange* states $\beta \leq 1$, but *In re Sodersten*, quotes Blackstone approvingly with $\beta \leq 1/10$. See 31 Cal. Rptr. 3d 831, 836 (Ct. App. 2005); 53 Cal. Rptr. 3d 572, 611 n.35 (Ct. App. 2007). A number of cases mention Blackstonian ratios disapprovingly in cases of pedophilia. See, e.g., *In re April C.*, 31 Cal. Rptr. 3d 804, 812 (Ct. App. 2005) (first citing *In re Kailee B.*, 22 Cal. Rptr. 2d 485 (Ct. App. 1993); then citing *In re Carmen O.*, 33 Cal. Rptr. 2d 848 (Ct. App. 1994); *In re Kailee B.*, 22 Cal. Rptr. 2d 485, 489–90 (Ct. App. 1993); *In re Carmen O.*, 33 Cal. Rptr. 2d 848, 856 n.7 (Ct. App. 1994) (citing *In re Kailee B.*, 22 Cal. Rptr. 2d 485 (Ct. App. 1993)). These seem to be carving out an exception rather than repudiating the formulation. “While one may accept [Blackstone’s ratio] in a criminal setting, though its exact statistical basis has not been precisely defined nor universally accepted, we trust that few, if any, would agree it is better that 10 pedophiles be permitted to continue molesting children than that 1 innocent parent be required to attend therapy sessions in order to discover why his infant daughter was falsely making such appalling accusations against him.” *In re Kailee B.*, 22 Cal. Rptr. 2d at 489–90. We think the articulation in *In re Sodersten* to be most authoritative for present purposes.

179. *People v. Dunway*, 88 P.3d 619, 637 (Colo. 2004) (citing *In re Winship*, 397 U.S. at 364).

Connecticut ¹⁸⁰	$\beta \leq 1$	50%
Delaware ¹⁸¹	No ruling	
Florida ¹⁸²	$\beta \leq 1/10^*$	90%
Georgia ¹⁸³	$\beta \leq 1$	50%
Hawai'i ¹⁸⁴	$\beta \leq 1$	50%
Idaho ¹⁸⁵	$\beta \leq 1/$ "Many"	98%
Illinois ¹⁸⁶	$\beta \leq 1$	50%
Indiana ¹⁸⁷	$\beta \leq 1$	50%

180. Connecticut is very consistent in following $\beta \leq 1$ articulated in *In re Winship*. *E.g.*, *Wiseman v. Armstrong*, 989 A.2d 1027, 1041 (Conn. 2010); *State v. Valinski*, 731 A.2d 311, 321 (Conn. App. Ct. 1999); *State v. Gerardi*, 677 A.2d 937, 941 (Conn. 1996). Curiously, *Miller v. Comm'r of Corrections* cites both *In re Winship*'s $\beta \leq 1$ and Blackstone's (1769) $\beta \leq 1/5$ approvingly. 700 A.2d 1108, 1141 (Conn. 1997) (Berdon, J., concurring).

181. Sadly, there does not seem to exist any clear articulation of the Blackstonian ratio in Delaware. The nearest Delaware courts have come seems to be in *Hughes v. State* implying $\beta \leq 1$. *See* 437 A.2d 559, 567 (1981) (quoting *Berger v. United States*, 295 U.S. 78, 88 (1935) ("[The prosecutor] is in a peculiar and very definite sense the servant of the law, the twofold aim of which is that guilt shall not escape or innocence suffer.")).

182. Florida courts have vacillated between $\beta \leq 1$ and $\beta \leq 1/10$. *See, e.g.*, *State v. Adkins*, 96 So. 3d 412, 433 (Fla. 2012) (Perry, J., dissenting) (citing *Coffin v. United States*, 156 U.S. 432, 456 (1895)) (supporting $\beta \leq 1/10$); *Nixon v. State*, 857 So. 2d 172, 178 (Fla. 2003) (Anstead, J., concurring) (citing *Nixon v. Singletary*, 758 So. 2d 618, 625–26 (Fla. 2000)) (approving of the language in *Singletary*); *Singletary*, 758 So. 2d at 626 (Harding, J., concurring) (first citing *In re Winship*, 397 U.S. at 372 (Harlan, J., concurring) (stating $\beta \leq 1$); then citing *Furman v. Georgia*, 408 U.S. 238, 367 n.158 (1972) (Marshall, J., concurring) (stating $\beta \leq 1/10$); *Adjmi v. State*, 154 So. 2d 812, 819 n.3 (Fla. 1963) (citing VOLTAIRE, ZADIG ch. 6 (1749) (supporting $\beta \leq 1$)). On balance, we think the balance of opinion seems to favor $\beta \leq 1/10$, though *Adjmi v. State* was the last majority statement by the Florida Supreme Court.

183. *Grace v. Hopper*, 217 S.E.2d 267, 269 (Ga. 1975) (citing *In re Winship*, 397 U.S. at 364).

184. *Masaki v. General Motors Corp.*, 780 P.2d 566, 574 (Haw. 1989).

185. *State v. Hester*, 760 P.2d 27, 41 (Idaho 1988). Note however that older cases have repudiated such formulations. *E.g.*, *State v. Reel*, 113 P. 721, 721 (Idaho 1911) (calling $\beta \leq 1/99$ "correct" as an abstract principle of the law while declining to find error in the trial court's refusal to allow it as a jury instruction); *State v. Crump*, 47 P. 814, 818 (Idaho 1897) (calling a variant of Blackstone's statement, where $\beta \leq 1/50$, a "heresy").

186. *People v. Bull*, 705 N.E.2d 824, 836–37 (Ill. 1998).

187. Indiana courts have been very consistent in stating $\beta \leq 1$. *E.g.*, *Orkin Exterminating Co. v. Traina* 486 N.E.2d 1019, 1023 (Ind. 1986) (citing *Travelers Indemnity Co. v. Armstrong*, 442 N.E.2d 349, 362 (Ind. 1982)); *Travelers*, 442 N.E.2d at 362 (citing *Tucker*, 408 N.E.2d at 820); *Walker v. State*, 779 N.E.2d 1158, 1161 (Ind. Ct. App. 2002) (citing *In re Winship*, 397 U.S. at 372 (Harlan, J., concurring)); *Indiana & Michigan Electric Co. v. Terre Haute Industries, Inc.*, 507 N.E.2d 588, 612 (Ind. Ct. App. 1987) (first citing *Travelers*, 442 N.E.2d at 362; then citing *Orkin*, 486 N.E.2d at 1023); *Tucker v. Marion Cty. Dep't of Pub. Welfare*, 408 N.E.2d 814, 820 (Ind. Ct. App. 1980).

Iowa ¹⁸⁸	$\beta \leq 1$	50%
Kansas ¹⁸⁹	No ruling	
Kentucky ¹⁹⁰	$\beta \leq 1$	50%
Louisiana ¹⁹¹	$\beta \leq 1$	50%
Maine ¹⁹²	$\beta \leq 1/10^*$	90%
Maryland	No ruling	
Massachusetts ¹⁹³	$\beta \leq 1$	50%
Michigan ¹⁹⁴	$\beta \leq 1/10$	90%
Minnesota ¹⁹⁵	$\beta \leq 1$	50%

188. *State v. Grattan*, 256 N.W. 273, 274 (Iowa 1934).

189. In *State v. Hoel*, the Supreme Court of Kansas rejected a jury instruction to the effect that $\beta \leq 1$, not because it is a mere abstract principle (the rationale often given in other states refusing such instructions), but more forcefully writing, “There is no such principle in the law of this state.” 243 P. 280, 280 (Kan. 1926).

190. *Brown v. Commonwealth*, 226 S.W.3d 74, 88 (Ky. 2007) (citing *In re Winship*, 397 U.S. at 372 (Harlan, J. concurring)). Historically, Kentucky has gone through several different standards. *E.g.*, *Lehrer v. Elmore*, 37 S.W. 292, 293 (Ky. 1896) (stating $\beta \leq 1/99$); *Fyffe v. Commonwealth*, 190 S.W.2d 674, 680 (Ky. 1945) (stating $\beta \leq 1/\text{“Many”}$).

191. *State v. Mussall*, 523 So. 2d 1305, 1308 (La. 1988) (citing *In re Winship*, 397 U.S. at 372 (Harlan, J., concurring)). Louisiana has been fairly consistent in describing $\beta \leq 1$. *See, e.g.*, *State v. Cage*, 583 So. 2d 1125, 1135 (La. 1991) (Dennis, J., dissenting) (though this opinion also forcefully states that the reasonable doubt standard must be more stringent than preponderance implying that $B \ll 1$); *State v. Hughes*, 900 So. 2d 168, 174 (La. Ct. App. 2005) (citing *In re Winship*, 397 U.S. at 368 (Harlan, J., concurring)), *rev’d*, 943 So. 2d 1047 (La. 2006); *State v. Mouton*, 653 So. 2d 1360, 1362 (La. Ct. App. 1995), *rev’d*, 653 So. 2d 1176 (La. 1995). However, there are some exceptions. *E.g.*, *State v. Furco*, 25 So. 951, 954 (La. 1899) (stating $\beta \leq 1/100$); *State v. Williams*, 591 So. 2d 404, 409 (La. Ct. App. 1991) (Bryan, J., concurring) (stating $\beta \leq 1/100$).

192. *Goddard v. Grand Trunk Ry. of Can.*, 57 Me. 202, 247 (1869) (quoting Blackstone possibly but without citation). This case, more than a century old, appears to be the only discussion in Maine’s courts about the relative value of type-I and type-II errors. Unfortunately, it states the principle in reference to punitive damages in a civil—not criminal—proceeding. It is therefore admittedly not quite on point.

193. *Commonwealth v. Johnson*, 650 N.E.2d 1257, 1263 (Mass. 1995) (citing *In re Winship*, 397 U.S. at 372 (Harlan, J., concurring)); *Commonwealth v. Walsh*, 378 N.E.2d 1378, 1383–84 (Mass. 1978) (citing *In re Winship*, 397 U.S. at 372 (Harlan, J., concurring)); and *In re Andrews*, 334 N.E.2d 15, 27 (Mass. 1975) (citing *In re Winship*, 397 U.S. at 371–72 (Harlan, J., concurring)).

194. *People v. Watkins*, 475 N.W.2d 727, 737 n.12 (Mich. 1991) (referring to $\beta \leq 1/10$ as a “historical principle” of the legal system); *People v. Allen*, 420 N.W.2d 499, 560 n.26 (Mich. 1986) (Boyle, J., dissenting) (referring to $\beta \leq 1/10$ as a “hallowed principle”).

195. *State v. Clausen*, 493 N.W.2d 113, 116 (Minn. 1992) (en banc) (citing *In re Winship*, 397 U.S. at 371–72 (Harlan, J., concurring)). *Clausen* replaced the $\beta \leq 1/\text{“Some”}$ statement from *State v. Butenhoff*. *See* 155 N.W.2d 894, 900 (Minn. 1968).

Mississippi ¹⁹⁶	$\beta \leq 1/\text{"Some"}$	98%
Missouri ¹⁹⁷	$\beta \leq 1/\text{"Some"}$	98%
Montana ¹⁹⁸	$\beta \leq 1$	50%
Nebraska ¹⁹⁹	$\beta \leq 1$	50%
Nevada	No ruling	
New Hampshire ²⁰⁰	$\beta \leq 1/10$	90%
New Jersey ²⁰¹	$\beta \leq 1$	50%

196. *McMillian v. State*, 67 So. 2d 290, 291 (Miss. 1953). Earlier cases show a wide variety of values. *E.g.*, *Herring v. State*, 84 So. 699, 705 (Miss. 1920) (Ethridge, J., concurring) (referring to $\beta \leq 1/\text{"Many"}$ as a "policy of the law"); *Jones v. State*, 30 So. 759, 762 (Miss. 1901) (stating $\beta \leq 1$ in reference to a prosecutor's responsibility to the law); *Jesse v. State*, 28 Miss. 100, 103 (1854) (stating $\beta \leq 1/99$).

197. It is somewhat difficult to determine β in Missouri. The best authoritative recent case seems to be *State v. Waller*, 163 S.W.3d 593, 597 (Mo. Ct. App. 2005) (per curiam) ("We accept the reality that sometimes the guilty will go free because of our great care to ensure that the innocent are not unjustly convicted."). However, the history is complicated. *See, e.g.*, *State v. Bonuchi*, 636 S.W.2d 338, 342 (Mo. 1982) (Donnelly, J., concurring) (strongly disapproving of $\beta \leq 1/99$); *State v. Benson*, 19 S.W. 213, 213 (Mo. 1892) (endorsing $\beta \leq 1/99$); *State v. Mayfield*, 879 S.W.2d 561, 565 (Mo. Ct. App. 1994) (citing *State v. Grim*, 854 S.W.2d 403, 425 (Mo. 1993) (en banc) (Robertson, C.J., dissenting) (referring to $\beta \leq 1$ as a rule of law)); *Town of Glenwood v. Roberts*, 59 Mo. App. 167, 171 (1894) (arguing for $\beta \leq 1/10$ to be applied equally between felonies and misdemeanors); *State v. Perkins*, 11 Mo. App. 82, 82 (1881) (disapproving of the inverse $\beta \leq 99/1$).

198. Montana courts are consistent in stating $\beta \leq 1$, though the cases are now quite old. *See State v. Ebel*, 15 P.2d 233, 237 (Mont. 1932); *State v. Riggs*, 201 P. 272, 282 (Mont. 1921); *State v. Rolla*, 55 P. 523, 526 (Mont. 1898).

199. *Rogers v. State*, 149 N.W. 318, 319 (Neb. 1914) (citing *McKay*, 132 N.W. at 774); *McKay v. State*, 132 N.W. 741, 745 (Neb. 1911) (in discussing prosecutor's duty to the law). Also, interesting is *Parrish v. State*, 15 N.W. 357, 358 (Neb. 1883) (holding that a jury instruction of $\beta \leq 1/99$ was properly refused, puzzlingly on the rationale that it would "confuse" jurors, with the caveat that it is a "maxim" that should inform judges in discharging their duties).

200. *State v. Mannion*, 136 A. 358, 363 (N.H. 1927) (calling it a "well-recognized maxim of our law").

201. *State v. Jimenez*, 908 A.2d 181, 193 (N.J. 2006). New Jersey Courts are fairly consistent about $\beta \leq 1$. *But see State v. Hill*, 631 A.2d 150, 157 (N.J. Super. Ct. App. Div. 1993) (first citing BLACKSTONE, *supra* note 12 in support of $\beta \leq 1/10$; then citing *State v. Haines*, 120 A.2d 118, 124 (N.J. 1956) (referring to $\beta \leq 1/\text{"Many"}$ approvingly as an "ancient view").

New Mexico ²⁰²	$\beta \leq 1/99$	99%
New York ²⁰³	$\beta \leq 1/5^*$	83%
North Carolina ²⁰⁴	$\beta \leq 1/10$	90%
North Dakota	No ruling	
Ohio ²⁰⁵	$\beta \leq 1/99$	99%

202. New Mexico courts are consistent in stating $\beta \leq 1/99$. See *State v. Bartlett*, 631 P.2d 321, 325 (N.M. Ct. App. 1981) (citing *Chambers*, 524 P.2d at 1002–03); *State v. Chambers*, 524 P.2d 999, 1002–03 (N.M. Ct. App. 1974).

203. New York courts have declared a large range of values for β . See *People v. Galbo*, 218 N.Y. 283, 291 (1916) (stating $\beta \leq 1/5$, citing *HALE*, *supra* note 14); *People v. Bennett*, 49 N.Y. 137, 143 (1872) (also stating $\beta \leq 1/\text{"Many"}$); *Ruloff v. People*, 18 N.Y. 179 (1858) (confusingly stating several values, $\beta \leq 1/\text{"Many"}$, *id.* at 184, $\beta \leq 1/5$, *id.* at 185 (citing *HALE*, *supra* note 14), and $\beta \leq 1/10$, *id.* at 187 (citing *BLACKSTONE*, *supra* note 12)); *People v. Lipsky*, 84 A.D. 2d 42, 47 (N.Y. App. Div. 1981) (stating $\beta \leq 1/\text{"Many"}$ and quoting the language in *Ruloff*, the relevant language pertaining to the Blackstonian ratio being overruled by *People v. Lipsky*); *People v. Edwards*, 236 N.Y.S.2d 84, 84 (N.Y. App. Div. 1962) (finding error in a trial court judge's description of $\beta \leq 1/99$ as "bunk" and "pious platitude of some old maid sop"); *People v. Larkman*, 259 A.D. 959, 962 (N.Y. App. Div. 1940) (Harris, J., dissenting) (stating $\beta \leq 1/5$ (citing *HALE*, *supra* note 14)); *People v. Cohen*, 191 N.Y.S. 831, 842 (N.Y. Sup. Ct. 1921) (calling $\beta \leq 1/10$ "a working principle which will probably abide with us as long as crime itself"); *Onderdonk v. State*, 648 N.Y.S.2d 214, 219 (N.Y. Ct. Cl. 1996) (identifying $\beta \leq 1/100$ as a due process guarantee); *In re Ralph M.*, 417 N.Y.S.2d 608, 611 (N.Y. Fam. Ct. 1979) (stating $\beta \leq 1$ (citing *In re Winship*, 397 U.S. 358, 372 (1970) (Harlan, J., concurring))); *In re X, Y and Z*, 43 N.Y.S.2d 361, 365 (N.Y. Fam. Ct. 1943) (calling $\beta \leq 1/99$ an "old adage in the law which has become embedded in our theory of jurisprudence"). We think it fair to infer $\beta \leq 1/5$ as it was stated in *Galbo*, the most recent Court of Appeals ruling, though we must admit the wide range of statements about the value of β , despite language purporting to take it seriously, somewhat undermines our hope that courts have taken the precise value seriously.

204. North Carolina courts are reasonably consistent. *State v. Smith*, 73 S.E.2d 901, 903 (N.C. 1953) (first citing *BLACKSTONE*, *supra* note 12; then citing *State v. Hendrick*, 61 S.E.2d 349 (N.C. 1950)); *Hendrick*, 61 S.E.2d at 356 (apparently quoting *BLACKSTONE*, *supra* note 12 without citation); *In re Spier*, 12 N.C. 491, 503 (1828) (referring to $\beta \leq 1/10$ as "the law"). *But see State v. Smith*, 24 N.C. 402, 407 (1842) (stating $\beta \leq 1/5$ (citing *HALE*, *supra* note 14)).

205. Ohio courts have been consistent that $\beta \leq 1/99$. See *Lamprecht v. State*, 95 N.E. 656, 660 (Ohio 1911) (calling it an "ancient and merciful rule of the common law"); *Jones Stranathan & Co. v. Greaves*, 26 Ohio St. 2, 4 (1874) (calling it a "humane principle"); *Silver v. State*, 17 Ohio 365, 369 (1848). But note that *State v. Wing* refers to $\beta \leq 1/99$ skeptically as an "old maxim, sometimes abused" and that subsequent cases (though from lower courts) have used looser language. 64 N.E. 514, 518 (Ohio 1902); *see also, e.g.*, *State v. Hill*, 317 N.E.2d 233, 237 (Ohio Ct. App. 1963) (stating $\beta \leq 1/\text{"A Few"}$); *Bixler v. State*, 18 Ohio Law Abs. 117, 120 (Ohio Ct. App. 1934) (stating $\beta \leq 1/\text{"Many"}$). *But see State v. Thoss*, 120 N.E.3d 1274, 1285 (Ohio Ct. App. 2018) (citing the $\beta \leq 1/99$ ratio articulated in *Lamprecht*, 84 Ohio St. at 49 as the reason for acquitting).

Oklahoma ²⁰⁶	$\beta \leq 1/10^*$	90%
Oregon ²⁰⁷	$\beta \leq 1$	50%
Pennsylvania ²⁰⁸	$\beta \leq 1$	50%
Rhode Island ²⁰⁹	$\beta \leq 1/10$	98%
South Carolina ²¹⁰	$\beta \leq 1/\text{"Many"}$	98%
South Dakota ²¹¹	$\beta \leq 1/\text{"Many"}$	98%
Tennessee ²¹²	$\beta \leq 1^*$	50%
Texas ²¹³	$\beta \leq 1^*$	50%

206. Oklahoma courts are inconsistent on the value of β . See *Stout v. State*, 130 P. 553, 554 (Okla. 1913) (mentioning $\beta \leq 1/10$ approvingly as a “maxim of English law”); *Abbott v. Territory*, 94 P. 179, 179 (Okla. 1908) (citing HALE, *supra* note 14) (stating $\beta \leq 1/5$). Also, from the lower courts, see *Brower v. State*, 221 P. 1050, 1052 (Okla. Crim. App. 1924) (stating $\beta \leq 1$ in reference to prosecutor’s duties to the law); *Pruitt v. State*, 270 P.2d 351, 362 (Okla. Crim. App. 1954) (stating $\beta \leq 1/100$). With reservations, we regard *Stout*, the most recent Oklahoma Supreme Court case, as authoritative.

207. *Stogsdill v. Bd. of Parole & Post-Prison Supervision*, 154 P.3d 91, 95 (Or. 2007) (quoting *In re Winship*, 397 U.S. at 370–71 (Harlan, J., concurring)).

208. *Appeal of Nicely*, 18 A. 737, 738 (Pa. 1889) (speaking in terms of a prosecutor’s duty as a “quasi judicial officer”). But subsequent lower court rulings have diverged widely. *E.g.*, *Butler v. Flo-Ron Vending Co.*, 557 A.2d 730, 742 (Pa. Super. Ct. 1989) (approvingly referring to $\beta \leq 1/99$ as a “proverb, become principle”); *In re McMullins*, 462 A.2d 718, 724 (Pa. Super. Ct. 1983) (approvingly referring to $\beta \leq 1/10$ as an “ancient maxim of the criminal law”); *Alberts v. Bradley*, 11 Pa. D. & C.2d 107, 114 (1958) (calling $\beta \leq 1/99$ a “fundamental concept”); *In re Myers & Brei*, 83 Pa. Super. 383, 394 (1924) (stating $\beta \leq 1/99$).

209. *Tempest v. State*, 141 A.3d 677, 695 (R.I. 2016) (quoting BLACKSTONE, *supra* note 12).

210. *State v. Fleming*, 89 S.E.2d 104, 106 (S.C. 1955); *State v. Baker*, 37 S.E.2d 525, 530 (S.C. 1946).

211. *State v. Brown*, 165 N.W. 987, 988 (S.D. 1917) (The court is somewhat ambiguous in approving of $\beta \leq 1/\text{"Many"}$).

212. *State v. Bishop*, 431 S.W.3d 22, 53 (Tenn. 2014) (citing *In re Winship*, 397 U.S. at 372 (Harlan, J., concurring)) (stating $\beta \leq 1$). *But see In re Benjamin M.*, 310 S.W.3d 844, 849 (Tenn. Ct. App. 2009) (citing *Schlup v. Delo*, 513 U.S. 298, 325 (1995)) (stating $\beta \leq 1/\text{"Several"}$). There is also the confusing language in a much older case. See *Rea v. State*, 76 Tenn. 356, 360 (1881) (“[I]t is better that ninety-nine guilty men *out of a hundred* should escape than that one innocent man should be convicted.” (emphasis added)).

213. Texas courts have been somewhat self-contradictory on the value of β . Courts have disapproved of various formulations. *E.g.*, *Delao v. State*, No. 10-05-00323-CR, 2006 WL 3317718, at *6 (Tex. App. Nov. 15, 2006), *aff’d*, 235 S.W.3d 235 (Tex. Crim. App. 2007) (finding no error in forbidding a voir dire inquiry about jurors’ beliefs as to $\beta \leq 1$); *Patterson v. State*, 598 S.W.2d 265, 272 (Tex. Crim. App. 1980) (finding no error in forbidding a voir dire inquiry about jurors’ beliefs as to $\beta \leq 1/10$); *Hudson Ins. Co. v. McKnight*, 58 S.W.2d 1088, 1091 (Tex. Civ. App. 1933) (referring to $\beta \leq 1/99$ as an

Utah ²¹⁴	$\beta \leq 1/10$	90%
Vermont	No ruling	
Virginia ²¹⁵	$\beta \leq 1$	50%
Washington ²¹⁶	$\beta \leq 1$	50%
West Virginia ²¹⁷	$\beta \leq 1/99$	99%
Wisconsin ²¹⁸	$\beta \leq 1/10^*$	90%

“erroneous proposition of law”). Cases that positively articulate a value of β seem to favor $\beta \leq 1$. *Pena v. State*, 226 S.W.3d 634, 651 (Tex. App. 2007) (first citing *Schlup v. Delo*, 513 U.S. 298, 325; then citing *In re Winship*, 397 U.S. at 372 (Harlan, J., concurring)); *Morrison v. State*, 845 S.W.2d 882, 884 n.7 (Tex. Crim. App. 1992); *Kelly v. State*, 483 S.W.2d 467, 480 (Tex. Crim. App. 1972) (Odom, J., concurring) (citing *In re Winship*, 397 U.S. at 372 (Harlan, J., concurring)) (stating $\beta \leq 1$).

214. *State v. Reyes*, 116 P.3d 305, 309 (Utah 2005) (citing BLACKSTONE, *supra* note 12); *State v. Sullivan*, 307 P.2d 212, 215 (Utah 1957) (referring to $\beta \leq 1/10$ as an “ancient and honored adage of our law”); *State v. Weldon*, 314 P.2d 353, 356 (Utah 1957) (referring to $\beta \leq 1/10$ as a “time honored and important precept of our law”). Utah courts have been fairly consistent, although there are a couple of exceptions. *See State v. Long*, 721 P.2d 483, 491 (Utah 1986) (citing *In re Winship*, 397 U.S. at 372 (Harlan, J., concurring)) (stating $\beta \leq 1$); *State v. Kourbelas*, 621 P.2d 1238, 1240 (Utah 1980) (citing *Sullivan*, which prescribes $\beta \leq 1/10$) (stating “some”).

215. Virginia has undergone a change in the articulation of β . Early cases advocated $\beta \leq 1/99$. *E.g.*, *McDaniel v. Commonwealth*, 181 S.E. 534, 538 (Va. 1935); *Finchim v. Commonwealth*, 3 S.E. 343, 344–45 (Va. 1887); *Cluverius v. Commonwealth*, 81 Va. 787, 878 (1886). *But see McCue v. Commonwealth*, 49 S.E. 623, 630 (Va. 1905) (“We have no fault to find with the [$\beta \leq 1/99$] as a rhetorical phrase, but as a guide to a jury in reaching a conclusion it is of no value.”). Subsequent cases have favored $\beta \leq 1$. *E.g.*, *Bateman v. Commonwealth*, 32 S.E.2d 134, 136 (Va. 1944) (citing *Mohler v. Commonwealth*, 111 S.E. 454, 454 (Va. 1922)); *Dingus v. Commonwealth*, 149 S.E. 414, 416 (Va. 1929) (citing *Mohler*, 111 S.E. at 460); *Fitzpatrick v. Commonwealth*, 115 S.E. 522, 523 (Va. 1923) (citing *Mohler*, 111 S.E. at 454); *Mohler*, 111 S.E. at 460 (on a prosecutor’s duty to the law); *Tuma v. Commonwealth*, 726 S.E.2d 365, 371 (Va. Ct. App. 2012) (citing *Appeal of Nicely*, 18 A. 737, 738 (Pa. 1889) (discussing a prosecutor’s duty to the law)); *Reedy v. Wright*, 60 Va. Cir. 18, 4 (2002) (citing *In re Winship*, 397 U.S. at 372 (Harlan, J., concurring)).

216. *State v. Rinaldo*, 655 P.2d 1141, 1144 (Wash. 1982) (en banc) (citing *In re Winship*, 397 U.S. at 372 (Harlan, J., concurring)); *State v. Wilcox*, 600 P.2d 561, 565 (Wash. 1979) (en banc) (citing *In re Winship*, 397 U.S. 358, 368 (1970)). There are also cases reinforcing this and stating the prosecutor’s duty to the law. *State v. Case*, 298 P.2d 500, 501, 503 (Wash. 1956); *State v. Montgomery*, 105 P. 1035, 1036 (Wash. 1909); *State v. Trout*, 105 P.3d 69, 83 (Wash. Ct. App. 2005).

217. *State v. Johnson*, 140 S.E. 532, 532 (W. Va. 1927) (referring to $\beta \leq 1/99$ as a “shibboleth of our criminal jurisprudence”); *Weeks v. Chesapeake & O. Ry. Co.*, 69 S.E. 805, 808 (W. Va. 1910) (Poffenbarger, J., dissenting) (stating $\beta \leq 1/99$); *Varner v. Martin*, 21 W. Va. 534, 542 (1883) (contrasting the rights of a criminal defendant from the “right” of a legislature). *But see State v. Michael*, 16 S.E. 803, 804 (W. Va. 1893) (stating $\beta \leq 1$).

218. Recent cases claim $\beta \leq 1/10$. *State v. McAlister*, 911 N.W.2d 77, 93 (Wis. 2018) (Bradley, J., dissenting) (“Our system of law has always operated under the theory that it

Wyoming ²¹⁹	$\beta \leq 1/100$	99%
U.S. Supreme Court ²²⁰	$\beta \leq 1$	50%

Table 3

V. CONCLUSION

The law's present treatment of the reasonable doubt standard is untenable. The courts refuse to characterize the standard in terms which might meaningfully guide a jury. They tell us merely that the threshold lies at some point *greater than 50% certainty and less than 100% certainty*.²²¹ Though courts proclaim with utmost solemnity the critical role of the reasonable doubt standard in the determination of justice, they have reduced its practical meaning to a guessing game. For the reasons we have discussed, the most practicable alternative—which avoids problems of burden-shifting and circularity—seems to be a probabilistic articulation of the standard. We point out that these probabilistic

is better for ten guilty people to go free than one innocent to languish in prison.”); *In re Torrance P., Jr.*, 725 N.W.2d 623, 636 (Wis. 2006) (Prosser, J., concurring) (citing BLACKSTONE, *supra* note 12); *State v. Dubose*, 699 N.W.2d 582, 600 n.1 (Wis. 2005) (Butler, J., concurring) (first citing BLACKSTONE, *supra* note 12; then citing *Furman v. Georgia*, 408 U.S. 238, 367 n.158 (1972)). In *Mayo v. Wisconsin Injured Patients & Families Compensation Fund*, the Wisconsin Supreme Court cites Blackstone's $\beta \leq 1/10$, albeit as *dicta*, contrasting it with the optimal distribution of type-I and type-II errors in constitutional cases. 214 N.W.2d 678, 703 (Wis. 2018). Earlier cases suggested $\beta \leq 1$. *State ex rel. Fitas v. Milwaukee*, 221 N.W.2d 902, 905 (Wis. 1974).

219. *State v. Peterson*, 194 P. 342, 349 (Wyo. 1920) (referring to $\beta \leq 1/100$ as an “almost universal doctrine” to contrast it with another proposition). This reference is tenuous but the only apparent statement of β in Wyoming courts.

220. *Francis v. Franklin*, 471 U.S. 307, 313 (1985).

221. In fact, few courts even do *this* much. Judge Newman observes, “A Westlaw search of all federal court opinions disclosed only two opinions in which a federal court of appeals explicitly stated that the evidence might be sufficient to satisfy the ‘preponderance’ standard but was insufficient to satisfy the higher ‘reasonable doubt’ standard.” Newman, *supra* note 47, at 990. A quarter of a century later, such explicit instruction remains uncommon, however Newman may have overstated how infrequently such statements occur. See *Addington v. Texas*, 441 U.S. 418, 423 (1979) (“Generally speaking, the evolution of this area of the law has produced across a continuum three standards or levels of proof for different types of cases. At one end of the spectrum is the typical civil case involving a monetary dispute between private parties. . . . In a criminal case, on the other hand, the interests of the defendant are of such magnitude that historically and without any explicit constitutional requirement they have been protected by standards of proof designed to exclude as nearly as possible the likelihood of an erroneous judgment.”). Cases citing *Addington* are also helpful. *E.g.*, *United States v. Restrepo*, 946 F.2d 654, 676 (9th Cir. 1991). It is also worth mentioning the Federal Judicial Center alludes to the distinction properly. FEDERAL JUDICIAL CENTER, PATTERN CRIMINAL JURY INSTRUCTIONS 28 (1987).

thresholds may be inferred from judicial pronouncements of Blackstonian ratios.

To those readers still skeptical of the value of mathematical precision in the law, we point out that the precisification of standards of proof is not a radical innovation. Courts have demonstrated little reluctance in stating with numerical precision that “preponderance of the evidence” means greater than 50% certainty. We find few courts or scholars protesting the precise articulation of *that* standard. Yet the courts have steadfastly refused to articulate a similarly precise threshold for reasonable doubt. It is worth bearing in mind that things were not always so. Prior to *In re Winship*, no less eminent a jurist than Learned Hand attempted to articulate a numerical threshold for reasonable doubt. His assertion, ironically, was that the reasonable doubt threshold should also be set at 50% certainty.²²² Hand evidently doubted whether any meaningful distinction existed between “the evidence which should satisfy reasonable men, and the evidence which should satisfy reasonable men beyond a reasonable doubt.”²²³ The consensus view of this seems to be that he lost sight of the objective.²²⁴

Regardless, the point remains that precisification of the threshold for certainty is *not* a foreign concept to our law generally. The steadfast refusal of courts to define it with precision is not an ancient principle handed down by the gods or Hammurabi or Justinian,²²⁵ but rather an aberration of the past several decades only.

In our review of the proposed definitions, theoretical arguments, and experimental research, we have braided several disparate threads in the development of the criminal standard of proof. We hope our arguments and caselaw survey of Blackstonian ratios might contribute to a resolution of this unnecessary semantic quandary which presently blights the criminal law.

222. *United States v. Feinberg*, 140 F.2d 592, 594 (2d Cir. 1944).

223. *Id.*

224. See Newman, *supra* note 47, at 985–86.

225. Justice O'Connor’s assertion that it is an “ancient and honored aspect of our criminal justice system” notwithstanding. *Victor v. Nebraska*, 511 U.S. 1, 5 (1994).