

1999

Refocusing the Burden of Proof in Criminal Cases: Some Doubt About Reasonable Doubt

Lawrence Solan

Follow this and additional works at: <https://brooklynworks.brooklaw.edu/faculty>

 Part of the [Criminal Law Commons](#), and the [Criminal Procedure Commons](#)

Recommended Citation

78 Tex. L. Rev. 105 (1999)

This Article is brought to you for free and open access by BrooklynWorks. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of BrooklynWorks.

Refocusing the Burden of Proof in Criminal Cases: Some Doubt About Reasonable Doubt

Lawrence M. Solan*

I. Introduction

This Article asks whether the proof beyond a reasonable doubt standard is the best way to promote the values our system of criminal justice claims to venerate. The results of the inquiry point to a single conclusion: standard reasonable doubt instructions focus the jury on the defendant's ability to produce alternatives to the government's case, and thereby shift the burden of proof to the defendant. In fact, empirical studies and linguistic analysis strongly suggest that it is more difficult to establish proof by clear and convincing evidence than it is to establish proof beyond a reasonable doubt, even though our system regards the former as reflecting a lighter burden.

Proof beyond a reasonable doubt has been the government's burden in criminal cases for virtually this country's entire history.¹ Although the Supreme Court did not hold this standard to be a constitutional imperative until 1970,² its use has not been a matter of any serious controversy. The phrase developed in the late eighteenth and early nineteenth centuries not as a value in itself, but rather as a fair way to express the sentiment that only when the members of a jury are very certain of a defendant's guilt should they convict. Some two hundred years later, we still use the same expression, and we profess to use it for the same reason.

Most debate in judicial opinions and in the scholarly literature has focused on whether reasonable doubt should be defined for the jury, and,

* Associate Professor of Law, Brooklyn Law School. An earlier draft of this Article was presented at the Rutgers-Newark School of Law. I wish to thank the members of the Rutgers faculty for their valuable suggestions. Sherry Colb, Carlos Gonzalez and George Thomas will quickly notice their contributions to this project upon reading the text. I would also like to thank Judges Prentice Marshall and Richard Posner, Joe Cecil of the Federal Judicial Center, and Professors Ursula Bentele, Margaret Berger, Charles Clifton, Jill Fisch, Susan Herman, Claire Hill, Bruce Kobayashi, Phil Johnson-Laird, Gary Minda, Bob Pitler, Peter Tiersma, Edwin Williams and Steve Winter for valuable discussion. Finally, thanks to Cori Browne, Brendan Dowd, Antonella Gallizzi, Paul Leroy, Alex Rohan, Robyn Schneider, and Amy Sender for their assistance on this project. This work was supported in part by a research stipend from Brooklyn Law School.

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

2. See *id.* at 364.

if so, how it should be defined.³ Although there is great disagreement among those who engage in this debate, all players share an important assumption: Our system of criminal justice comes blessed with a ready-made standard for the government to meet, and all we have to do is to elucidate it a bit to make sure that jurors get it right. This Article challenges that assumption. It suggests that we revise our conception of the government's burden of proof to focus more clearly on what the government must establish. The Federal Judicial Center's model instruction, which tells the jury that the prosecution must leave them "firmly convinced" of a defendant's guilt before they may convict,⁴ comes close to the mark. By emphasizing the government's task—and not the defendant's—it goes a long way toward ensuring that the burden of proof in criminal cases remains where it belongs: on the government. New Jersey's instruction, which requires that the jury be "firmly convinced" of the defendant's guilt, does an even better job.⁵

The firmly convinced standard is virtually identical to the standard used in France,⁶ and is similar to that used by those English courts which, frustrated by a lack of universal understanding of reasonable doubt, have abandoned that term altogether.⁷ Nonetheless, despite the availability of instructions that appropriately place the burden on the government to prove its case with strong evidence, very few jurisdictions use them. I argue that this situation should be corrected—whether by appellate courts, legislatures, or jury reform commissions.

This approach to the burden of proof in criminal cases is seemingly at odds with prevailing popular thought about the criminal justice system. The sense that the system is soft on criminal defendants is pervasive in our

3. Compare Note, *Reasonable Doubt: An Argument Against Definition*, 108 HARV. L. REV. 1955, 1955 (1995) (concluding that "courts should not attempt to define the term [reasonable doubt] in conveying the reasonable doubt concept to juries"), with Jessica N. Cohen, *The Reasonable Doubt Jury Instruction: Giving Meaning to a Critical Concept*, 22 AM. J. CRIM. L. 677, 678 (1995) (arguing that "because reasonable doubt is a term of art it should be defined for the jury"), and Henry A. Diamond, Note, *Reasonable Doubt: To Define, or Not to Define*, 90 COLUM. L. REV. 1716, 1716 (1990) ("[J]ury instructions defining reasonable doubt should always be given in criminal trials and are constitutionally required when requested by the defendant or the jury."). See also Thomas V. Mulrine, *Reasonable Doubt: How in the World is it Defined?*, 12 AM. U. J. INT'L. L. & POL'Y 195 (1997) (explaining various approaches to and definitions of reasonable doubt).

4. See FEDERAL JUDICIAL CTR., PATTERN CRIMINAL JURY INSTRUCTIONS No. 21, at 28 (1988).

5. See *State v. Medina*, 685 A.2d 1242, 1251 (N.J. 1996) (instructing courts to define reasonable doubt as that proof which "leaves you firmly convinced of the defendant's guilt"). This case is discussed more fully *infra* note 53.

6. See THE AMERICAN SERIES OF FOREIGN PENAL CODES: THE FRENCH CODE OF CRIMINAL PROCEDURE, art.353, at 172 (Gerald L. Kock & Richard S. Frase trans., rev. ed. 1988) (requiring the court to ask the jury: "Are you thoroughly convinced?").

7. See *Walters v. Queen* [1969] 2 App. Cas. 26, 30 (P.C. 1969) (appeal taken from Jam.) (upholding the validity of a jury charge given by a lower court that had jettisoned the reasonable doubt standard).

society. A substantial majority of those polled believe the courts to be responsible in part for high crime rates,⁸ even though such rates have been falling.⁹ O.J. Simpson's acquittal, moreover, still lingers in the minds of many as evidence that the judicial system is too lenient.¹⁰ The perceived stringency of the reasonable doubt standard is considered one of the reasons for Simpson's exoneration.¹¹ I suggest that this critique is also misfocused. The problem is not that the reasonable doubt standard is generally too hard for the government to meet when its case is strong. On the contrary, much of this Article will show that the standard generally requires less of the government in criminal cases than is commonly believed.

This Article will also address how the reasonable doubt standard can lead, in certain circumstances, to the acquittal of guilty defendants. The reasonable doubt standard does encourage defense attorneys to direct a jury's attention to remote contingencies when the government has presented a strong case. Sometimes, especially when they have substantial resources, defense attorneys will be successful in using this strategy to gain the acquittal of a defendant even when the government's case is convincing. The approach espoused here should help to remedy that problem as well.

In Part II, I discuss the values, both cultural and historical, that underlie the reasonable doubt standard. A criminal justice system that risks freeing the guilty in order to avoid convicting the innocent is generally regarded as a hallmark of a decent civilization. The requirement that the government prove a defendant guilty beyond a reasonable doubt is one way our society promotes this ideal. I summarize the history of the reasonable

8. See, e.g., Laura B. Meyers, *Bringing the Offender to Heel: Views of the Criminal Courts*, in AMERICANS VIEW CRIME AND JUSTICE: A NATIONAL PUBLIC OPINION SURVEY 46, 53 (Timothy J. Flanagan & Dennis R. Longmire eds., 1996) (finding that 67% of those polled were "concerned about courts that do not reduce crime"). One recent study showed barely half the people surveyed (56%) to agree with Blackstone's famous statement that it is better to allow ten guilty people to go free than to convict one innocent person by mistake. See Robert J. Boeckmann & Tom R. Tyler, *Commonsense Justice and Inclusion Within the Moral Community: When Do People Receive Procedural Protections From Others?*, 3 PSYCHOL. PUB. POL. & L. 362, 366 (1997).

9. See Marlene Cimon, *Study Finds Murder Rate at 30-Year Low Trends*, L.A. TIMES, Jan. 3, 1999, available in 1999 WL 2117172 (citing statistics indicating that the national murder rate was at a three decade low and that the rate of violent crime overall was at the lowest level in 24 years).

10. For a more thorough discussion, see Reid Hastie & Nancy Pennington, *The O.J. Simpson Stories: Behavioral Scientists' Reflections on The People of the State of California v. Orendhal James Simpson*, 67 U. COLO. L. REV. 957, 958 (1996) (offering a "psychological post-mortem on the Simpson trial verdict that explicates the jurors' decision-making processes") and Mulrine, *supra* note 3, at 195-96, as well as the voluminous discussions in the popular press.

11. See, e.g., ALAN M. DERSHOWITZ, REASONABLE DOUBTS: THE O.J. SIMPSON CASE AND THE CRIMINAL JUSTICE SYSTEM 69 (1996) (suggesting that the "exacting" nature of the reasonable doubt standard was a possible explanation for the jury's verdict in the Simpson case); Steven M. Cooper, *People v. Simpson: Some Thoughts on Aberrance*, 25 AM. J. CRIM. L. 165, 167 (1997) ("If there was reasonable doubt in the Simpson trial, how could there ever be a criminal case without it?").

doubt standard in this light. I then present three approaches that courts use today. Some courts define what a reasonable doubt is, others present the reasonable doubt standard to the jury but leave it unexplained, and other courts tell the jury of the government's burden in a manner that focuses more on what the government must prove and less on the types of doubts that require acquittal. The approach that focuses more on the government's burden has, I argue, the best chance of leading to convictions when the government's case is strong and to acquittals when the government's case is weak.

In Part III, I discuss both real-world data and a number of experimental studies designed to test how jurors apply various burden of proof instructions. These studies indicate that typical reasonable doubt instructions are likely to lead to conviction of defendants in cases in which favorable evidence is evenly divided between the government and the defendant, and that a rather low probability of guilt is needed to meet the standard for most jurors. In one such study, about one-third of jurors who had just participated in a criminal trial actually agreed that reasonable doubt requires that the defendant prove her innocence once the government has come forward with evidence of guilt.¹² When different instructions are compared in some of these experiments, the "firmly convinced" standard that focuses on the government's burden alone results in high conviction rates for cases in which the government's evidence is strong and much lower conviction rates when the evidence against a defendant is equivocal. Moreover, some of these studies suggest that the supposedly weaker clear and convincing evidence standard is actually tougher to meet than proof beyond a reasonable doubt.

Using analytical tools developed by linguists, cognitive psychologists, and philosophers, I examine in Part IV the expression "proof beyond a reasonable doubt" with the goals of explaining the empirical results described in Part III and determining whether that expression is likely to foster the values discussed in Part II. Close analysis reveals several problems. First, the expression misfocuses the jury on the extent to which the defense has created doubt. It says absolutely nothing about how strong the government's case has to be. This increases the likelihood that the government will win when it brings a case which is weak, but for which the defendant presents no answer—possibly because she knows nothing about the crime. Second, the reasonable doubt standard cannot overcome prototype effects. People tend to categorize new things and events with respect to mental models that they have already formed. If conviction of a crime fits the facts better than acquittal, it is extremely difficult to overcome the desire to match the facts with the better of the two models,

12. See *infra* notes 62-63 and accompanying text.

even if the case is not very strong. Third, “doubt” is what linguists and philosophers call a verb of propositional attitude. Such verbs, which include “think” and “believe,” are best understood in terms of “possible world” semantics. This means that the process of doubting is an imaginative one. Therefore, any instructions that tell jurors not to consider “speculative” doubts are really telling the jurors not to consider any doubts at all. In contrast, telling the jurors that the government has the burden of leaving them firmly convinced that the defendant has committed every element of the crime goes a long way to remedy these problems. Part V is a brief conclusion.

II. The Standard of Proof In Criminal Cases

A. *Who Should Be Convicted of Crimes?*

The answer to this question is simple: Those who commit crimes should be convicted of them. Those who do not commit crimes should not be charged at all. If they are charged, then they should be acquitted. The problem, of course, is that we are not always certain in advance about the defendant’s guilt, and often are never certain.¹³ As Justice Harlan explained in his famous concurrence in *In re Winship*:¹⁴

[I]n a judicial proceeding in which there is a dispute about the facts of some earlier event, the fact finder cannot acquire unassailably accurate knowledge of what happened. Instead, all the fact finder can acquire is a belief of what *probably* happened. The intensity of this belief—the degree to which a fact finder is convinced that a given act actually occurred—can, of course, vary. In this regard, a standard of proof represents an attempt to instruct the fact finder concerning the degree of confidence our society thinks he should have in the correctness of factual conclusions for a particular type of adjudication. Although the phrases “preponderance of the evidence” and “proof beyond a reasonable doubt” are quantitatively imprecise, they do communicate to the finder of fact different notions concerning the degree of confidence he is expected to have in the correctness of his factual conclusions.¹⁵

Harlan regarded the much more onerous reasonable doubt standard as the correct one for criminal cases because of a “fundamental value determination of our society that it is far worse to convict an innocent man than

13. For a discussion of this point in the context of the Fourth Amendment, see Sherry F. Colb, *Innocence, Privacy and Targeting in Fourth Amendment Jurisprudence*, 96 COLUM. L. REV. 1456, 1462 (1996) (arguing that “the Fourth Amendment is a right of the innocent and the guilty, but it is not exactly the same right for both”).

14. 397 U.S. 358 (1970).

15. *Id.* at 370.

to let a guilty man go free."¹⁶ Putting it somewhat differently, Justice Brennan said in his majority opinion in the same case: "It is also important in our free society that every individual going about his ordinary affairs have confidence that his government cannot adjudge him guilty of a criminal offense without convincing a proper factfinder of his guilt with utmost certainty."¹⁷ This, according to Brennan, is because our society values "the good name and freedom of every individual."¹⁸ Eight years later, writing for a unanimous court in *Addington v. Texas*,¹⁹ Chief Justice Burger expressed the same sentiment: "In the administration of criminal justice, our society imposes almost the entire risk of error upon itself. This is accomplished by requiring under the Due Process Clause that the state prove the guilt of an accused beyond a reasonable doubt."²⁰

These ideas are old: they derive from seventeenth century thought.²¹ Most widely quoted is William Blackstone's famous maxim, "it is better that ten guilty persons escape, than that one innocent suffer."²² The expression "proof beyond reasonable doubt," however, did not become widely used until the late eighteenth or early nineteenth century, and until the latter part of the nineteenth century courts had not reached agreement on proof beyond a reasonable doubt as the universal expression of the concept.²³ While some of the history is

16. *Id.* at 372.

17. *Id.* at 364.

18. *Id.*

19. 441 U.S. 418 (1979).

20. *Id.* at 423-24.

21. See BARBARA J. SHAPIRO, *BEYOND REASONABLE DOUBT AND PROBABLE CAUSE: HISTORICAL PERSPECTIVES ON THE ANGLO-AMERICAN LAW OF EVIDENCE* 40 (1991); Anthony A. Morano, *A Reexamination of the Development of the Reasonable Doubt Rule*, 55 B.U. L. REV. 507, 513-15 (1975) (both tracing the rule's philosophical origins to late seventeenth century England); see also Theodore Waldman, *Origins of the Legal Doctrine of Reasonable Doubt*, 20 J. HIST. IDEAS 299, 299-301 (1959) (tracing the origins of the reasonable doubt rule to the resolution of religious conflicts during and after the Reformation).

22. 4 WILLIAM BLACKSTONE, *COMMENTARIES* 358 (photo. reprint 1978) (9th ed. 1783). Other versions of this maxim have claimed various ideal numbers of false acquittals to avoid convicting a single innocent person. This number ranges from 5 (Hale) to 20 (Fortescue). See John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065, 1077 n.11 (1968).

23. At the beginning of the nineteenth century, proof beyond all reasonable doubt was the most popular version. See *Ogden v. Saunders*, 25 U.S. 213, 270 (1827) (stating that a state law should be presumed valid until a violation of the Constitution is proved beyond all reasonable doubt); *The Santissima Trinidad*, 20 U.S. 283, 340 (1822) (noting that the burden of proof rests on the claimant, and that an absence of proof gives rise to the presumption that the claimant's assertions are untrue); *Spurr v. Pearson*, 22 F. Cas. 1011, 1015 (C.C.D. Mass. 1816) (No. 13,268) (noting that the guilt of parties is to be established beyond all reasonable doubt). One variant used proof beyond any reasonable doubt. See *White v. Nicholls*, 44 U.S. 266, 290 (1845); *Philadelphia & Trenton R.R. Co. v. Stimpson*, 39 U.S. 448, 460 (1840); *Maisonnaire v. Keating*, 16 F. Cas. 513, 517 (C.C.D. Mass. 1815) (No. 8978) (each applying a "beyond any reasonable doubt" standard). Other courts used the naked, "proof beyond reasonable doubt." See *United States v. The Brig Burdett*, 34 U.S. 682, 690 (1835); *The Amiable Isabella*, 19 U.S. 1, 78 (1821); *United States v. Haswell*, 26 F. Cas. 218, 219

controversial,²⁴ Justice O'Connor summarizes the standard account in her majority opinion in *Victor v. Nebraska*.²⁵ In essence, seventeenth century thinkers, influenced heavily by John Locke, developed an epistemology that differentiated among various kinds of evidence. Demonstrative evidence, experienced first hand through sensory experience, was considered the most reliable. Indirect evidence, from the reports of witnesses, was called "moral evidence," and was considered less reliable.²⁶ The expression "moral certainty," which we still hear in some reasonable doubt instructions today, essentially meant "the highest degree of certitude based on such evidence."²⁷ Proof beyond a reasonable doubt was equated with moral certainty.²⁸ In more recent times, the Supreme Court has described the reasonable doubt standard itself in similar terms: "[B]y impressing upon the factfinder the need to reach a subjective state of near certitude of the guilt of the accused, the standard symbolizes the significance that our society attaches to the criminal sanction and thus to liberty itself."²⁹ A unanimous Supreme Court reiterated its commitment to these values in 1993, when it held that a constitutionally deficient instruction on the meaning of reasonable doubt cannot be subjected to harmless error analysis.³⁰

For the present, I would like to accept all of these statements as given. That is, we use the expression "proof beyond a reasonable doubt" because

(C.C.D. Vt. 1800) (No. 15,324) (all requiring a "beyond reasonable doubt" standard). Only occasionally do we see proof beyond a reasonable doubt during this time period. See *Prevost v. Gratz*, 19 U.S. 481, 498 (1821); *United States v. Hewson*, 26 F. Cas. 303, 309 (C.C.D. Mass. 1844) (No. 15,360) (both requiring proof "beyond a reasonable doubt"). For reasons discussed in Part IV, *infra* note 158 and accompanying text, these differences reflect significant changes between early conceptualizations of the government's burden of proof and modern ones.

24. Compare SHAPIRO, *supra* note 21 (arguing that "beyond a reasonable doubt" was intended to require the highest degree of certainty), with Morano, *supra* note 21, at 519, 527 (arguing that "beyond a reasonable doubt" replaced the tougher standard, "beyond all doubt," in order to make convictions easier). It is not entirely clear that these positions are inconsistent. First, Shapiro relies on English cases, Morano on American cases. See Morano, *supra* note 21, at 519 (discussing the "development of the reasonable doubt rule in America"); Shapiro, *supra* note 21, at xii ("My primary focus is on early modern England . . ."). Second, if the goal is to convict only when one is as certain as one can expect oneself to be in a world in which nothing is certain, then "proof beyond all doubt" really does not accurately reflect the then-current epistemology.

25. 511 U.S. 1 (1994).

26. See *Boyd v. Gosser*, 82 So. 758, 761 (Fla. 1918) (categorizing demonstrative evidence as "evidence of the highest rank" and noting that the value of moral evidence depends "upon veracity on the one hand and credulity upon the other").

27. *Victor*, 511 U.S. at 11.

28. See *id.* at 12 ("Proof 'beyond a reasonable doubt' . . . is proof 'to a moral certainty,' as distinguished from an absolute certainty. As applied to a judicial trial for crime, the two phrases are synonymous and equivalent; . . . each signifies such proof as satisfies the judgment and consciences of the jury, as reasonable men . . .") (quoting *Commonwealth v. Costley*, 118 Mass. 1, 24 (1875)).

29. *Jackson v. Virginia*, 443 U.S. 307, 315 (1979) (citing *In re Winship*, 397 U.S. 358, 372 (1970) (Harlan, J., concurring)).

30. See *Sullivan v. Louisiana*, 508 U.S. 275, 277-82 (1993).

we believe that the government should be required to prove its case so strongly that the evidence leaves the jury with the highest degree of certitude based on such evidence.³¹ The rest of this Article is devoted to an important question too seldom asked: Does the expression "proof beyond a reasonable doubt" accomplish that goal?

B. *Approaches to Burden of Proof in Criminal Cases*

Courts tell jurors of the government's burden in three different ways. Some courts attempt to define what constitutes a doubt that is reasonable. This is the most common approach, and is a mistake for many reasons discussed later in this Article. Others use the reasonable doubt standard without defining it. This, I believe, is also a mistake. A third group attempts to define the government's burden of proof affirmatively. It is this approach that I espouse here.

1. *Explaining What a Doubt Is.*—The most typical approach to instructing juries of the government's burden is to tell them that the government must prove guilt beyond a reasonable doubt, and then to tell the jurors what it means to have a reasonable doubt. This approach is also

31. Despite this sentiment, some people are convicted of crimes that they did not commit. Press accounts of wrongful convictions are not uncommon. See Pam Belluck, *Convict Freed After 16 Years on Death Row*, N.Y. TIMES, Feb. 6, 1999, at A7, available in LEXIS, News Library, NYT File (noting the release of a man who was sentenced to death for a crime he did not commit); Adam Cohen, *Innocent, After Proven Guilty*, TIME, Sept. 13, 1999, at 26 (describing a junior high school science teacher's wrongful conviction for rape and murder). While there is substantial literature documenting the conviction of innocent defendants in today's system of justice, including some celebrated cases, I can find no analysis that quantifies the phenomenon convincingly. See, e.g., EDWARD CONNORS ET AL., CONVICTED BY JURIES, EXONERATED BY SCIENCE: CASE STUDIES IN THE USE OF DNA EVIDENCE TO ESTABLISH INNOCENCE AFTER TRIAL (1996); C. RONALD HUFF ET AL., CONVICTED BUT INNOCENT: WRONGFUL CONVICTION AND PUBLIC POLICY (1996); MARTIN YANT, PRESUMED GUILTY: WHEN INNOCENT PEOPLE ARE WRONGLY CONVICTED (1991). In *Convicted But Innocent*, people in the judicial system were asked for their impressions of the extent of the problem, but it is hard to draw any reliable inferences from the results. See HUFF ET AL., *supra*, at 61. Somewhat more convincing, but still attenuated, are inferences that Huff *et al.* draw from one well-known study. See HUFF ET AL., *supra*, at 60 (speculating, based on Kalven and Zeisel's study, that one percent of all convictions are actually erroneously decided against innocent defendants). Using real cases, Kalven and Zeisel found that judges tend to be harder on defendants than are jurors. See HARRY KALVEN, JR. & HANS ZEISEL, THE AMERICAN JURY 59 (1966) (representing that "jury trials show on balance a net leniency of 16 percent"). By questioning judges at the conclusion of jury trials, Kalven and Zeisel found that in 16.9% of the cases the judge would have convicted when a jury acquitted, but in only 2.2% of the cases would the judge have acquitted when the jury convicted. Thus, Kalven and Zeisel concluded that the judge and jury disagreed 19.1% of the time and agreed 75.4% of the time. In the remaining 5.5% of cases there was a hung jury. See *id.* at 56 (providing additional interpretation of these statistics). Under almost any set of assumptions about who is right in the disputed cases, judge or jury, the rate of false acquittals is greater than the rate of false convictions, but not by a margin even approaching Blackstone's 10 to 1 ideal. See HUFF ET AL., *supra*, at 59-61. In a society that prosecutes such huge numbers of individuals, even a small rate of error must be taken seriously, since it will cause the false conviction of substantial numbers of people.

prevalent in the Canadian courts.³² In a notorious example, the New York Court of Appeals ruled in 1947 that the following definition of “reasonable doubt” in a criminal jury instruction violated a defendant’s right to due process:

It is not a doubt based upon sympathy or a whim or prejudice or bias or a caprice, or a sentimentality, or upon a reluctance of a weak-kneed, timid, jellyfish of a juror who is seeking to avoid the performance of a disagreeable duty, namely, to convict another human being of the commission of a serious crime.³³

This instruction is embarrassing. Presumably, no judge would say such a thing today. But the fear that drove the trial judge to issue the instruction—the fear that proof beyond a reasonable doubt offers too much protection to the guilty—is very much alive in modern courts’ subtler but equally effective approaches to minimizing the government’s burden of proof in criminal cases.

Some courts remind jurors that for a doubt to be reasonable, it must be “an actual and substantial doubt, and not a mere possible doubt.”³⁴ Other courts try to rein in the jurors’ imaginations: “A reasonable doubt is a doubt which is something more than a guess or surmise. It is not a conjecture or a fanciful doubt. A reasonable doubt is not a doubt which is raised by someone simply for the sake of raising doubts.”³⁵ Courts that issue instructions like these have learned to stay away from jellyfish metaphors, but are still trying to dissuade jurors from acquitting the guilty based on far-fetched alternative explanations of the evidence.³⁶

32. See *Brydon v. The Queen*, [1995] 129 D.L.R. 4th 1 (Can.) (evaluating the trial judge’s explanation of the meaning of reasonable doubt and ordering a new trial based on his suggestion that the jury must grant the accused the benefit of a “unanimous” doubt).

33. *People v. Feldman*, 71 N.E. 2d 433, 439 (N.Y. 1947) (noting that such a definition “was not conducive to a fair and impartial consideration of the evidence”).

34. *Ex parte Adkins*, 600 So. 2d 1067, 1070 (Ala. 1992); see also *Ramirez v. Hatcher*, 136 F.3d 1209, 1211 (9th Cir. 1998) (affirming a jury instruction that defined reasonable doubt as doubt that is “actual and substantial” rather than doubt based on “mere possibility or speculation”); *Truesdale v. Moore*, 142 F.3d 749, 757 n.5 (4th Cir. 1998) (“The term reasonable doubt as used in these instructions is defined as a substantial doubt or a well-founded doubt arising out of the evidence or the lack of evidence.”).

35. *State v. Taylor*, 687 A.2d 489, 501 n.12 (Conn. 1996); see also *Truesdale*, 142 F.3d at 757 n.5 (“A reasonable doubt is not a weak doubt or a slight doubt, nor is a reasonable doubt a whimsical, fanciful or imaginary doubt, for a person may of course have these kinds of doubts about any proposition.”); *United States v. Isaac*, 134 F.3d 199, 202 (3d Cir. 1998) (“It is not a mere possible or imaginary doubt, because as you well know, everything relating to human affairs, and depending on oral testimony, is open to some possible or imaginary doubt.”).

36. That is not to say that courts cannot go too far with these instructions. In *Cage v. Louisiana*, the Supreme Court held that an instruction telling jurors that “[i]t must be such a doubt as would give rise to a grave uncertainty” did not pass constitutional muster. *Cage v. Louisiana*, 498 U.S. 39, 40-41 (1990). The Court also criticized the use of the expression “substantial doubt” elsewhere in the instruction. See *id.* at 41.

In contrast, some courts attempt to define reasonable doubt by likening it to doubts encountered in real life experience. The most popular of these instructions tells jurors that a reasonable doubt is "a kind of doubt that would cause a reasonable and sensible person to hesitate before acting upon a matter of importance in his or her own affairs."³⁷ A variant says that proof beyond a reasonable doubt "is such proof as ordinary prudent men and women would act upon in their most important affairs."³⁸

Many instructions give mixed messages:

A reasonable doubt is not a possible doubt, a speculative, imaginary or forced doubt. Such a doubt must not influence you to return a verdict of not guilty if you have an abiding conviction of guilt. On the other hand, if, after carefully considering, comparing and weighing all the evidence, there is not an abiding conviction of guilt, or, if, having a conviction, it is one which is not stable but one which wavers and vacillates, then the charge is not proved beyond every reasonable doubt and you must find the defendant not guilty because the doubt is reasonable.³⁹

This sampling does not cover the entire range of "reasonable doubt" instructions.⁴⁰ There are a host of these definitions, some of which contain a grab bag of warnings about which doubts are not reasonable.⁴¹

37. *Commonwealth v. Hall*, 701 A.2d 190, 208 (Pa. 1997); *see also United States v. Williams*, 20 F.3d 125, 129 n.2 (5th Cir. 1994) (approving an instruction that contains similar language along with some language from the Federal Judicial Center model instruction).

38. *State v. Mitchell*, 577 N.W.2d 481, 485 (Minn. 1998). This language tracks the Minnesota Jury Instruction Guide. *See MINNESOTA DISTRICT JUDGES ASS'N COMM. ON CRIMINAL JURY INSTRUCTION GUIDES*, 10 MINNESOTA JURY INSTRUCTION GUIDES—CRIMINAL No. 3.03 (3d ed. 1990). Some courts give jurors both versions. *See Isaac*, 134 F.3d at 199. The Third Circuit reviewed the following instruction in *Isaac*:

A reasonable doubt is a fair doubt, based upon reason and common sense—the kind of doubt that would make a reasonable person hesitate to act. Proof beyond a reasonable doubt must, therefore, be proof of such a convincing character that you would be willing to rely and act upon it, unhesitatingly, in the most important of your own affairs.

Id. at 202.

39. *Woods v. State*, 596 So. 2d 156, 158 (Fla. 1992). This tracks the language of the Florida Standard Jury Instructions. *See FLORIDA STANDARD JURY INSTRUCTIONS IN CRIMINAL CASES* § 2.03 (1981) (amended 1998).

40. I will not present an encyclopedia of reasonable doubt instructions, but for some examples, see *Chambers v. State*, 944 P.2d 805, 810 (Nev. 1997) ("such doubt as would govern or control a person in the more weighty affairs of life"); *State v. Darby*, 477 S.E.2d 710, 710 (S.C. 1996) ("the kind of doubt that would cause a reasonable person to hesitate to act"); COMMITTEE ON MODEL JURY INSTRUCTIONS WITHIN THE EIGHTH CIRCUIT, *MANUAL OF MODEL CRIMINAL JURY INSTRUCTIONS FOR THE DISTRICT COURTS OF THE EIGHTH CIRCUIT* § 3.11 (1996) ("proof of such a convincing character that a reasonable person would not hesitate to rely and act") and U.S. SIXTH CIRCUIT DISTRICT JUDGES ASS'N, *PATTERN CRIMINAL JURY INSTRUCTIONS* §§ 1.03(4) & (5) (1991) ("Possible doubts or doubts based purely on speculation are not reasonable doubts. . . . [What is required is] proof which is so convincing that you would not hesitate to rely and act on it in making the most important decisions in your own lives.").

41. An especially rich example is the instruction rejected by the Supreme Court of New Jersey in *State v. Medina*, 685 A.2d 1242, 1244-45 (N.J. 1996), discussed *infra* note 162. The confusion

Even a cursory review shows that the main concern of many of these instructions is that jurors not get carried away with their imaginations. Speculation is not permitted—doubts must have an articulable reason. In Part IV of this Article, I will explain why even much tamer versions of this approach will often lead to conviction when the strength of the government's evidence is questionable.

2. *Leaving Reasonable Doubt Undefined.*—Some courts inform the jury that the government must prove its case beyond a reasonable doubt, but offer no explanation of that term on the theory that any possible definition would only make the instructions more complicated and confusing. The Seventh Circuit takes this approach,⁴² as does the Fourth Circuit.⁴³ English courts not embracing the “satisfied that you are sure” standard also take this position.⁴⁴ The Seventh Circuit justifies its approach by accepting the following propositions:

[T]hat definition is impossible, that the phrase is self-defining, that there is no equivalent phrase more easily understood, that every attempt to explain renders an explanation of the explanation necessary, that the better practice is not to attempt the definition, and that any effort at further elucidation tends to misleading refinements.⁴⁵

associated with these “definitions” has not gone unnoticed. See PETER M. TIERSMA, *LEGAL LANGUAGE* 194-96 (1999) (describing both the difficulty that California jurors face in comprehending that state's antiquated definition of “reasonable doubt” and the risk of reversal that judges who try to explain the term face); Patricia M. Wald, *Beyond a Reasonable Doubt: A Norm Gives Way to the Numbers*, 1993 U. CHI. LEGAL F. 101, 111-12 (noting a Supreme Court decision rejecting a state court definition of reasonable doubt as “grave uncertainty,” “actual substantive doubt,” or “moral certainty” as too restrictive (citing *Cage v. Louisiana*, 498 U.S. 39, 41 (1990))).

42. See *United States v. Blackburn*, 992 F.2d 666, 668 (7th Cir. 1993) (“We have reiterated time and again our admonition that district courts should not attempt to define reasonable doubt.”); *United States v. Hall*, 854 F.2d 1036, 1039 (7th Cir. 1988) (asserting that further definition of “reasonable doubt” risks the defendant's constitutional rights and offers no appreciable benefit); *United States v. Marquart*, 786 F.2d 771, 784 (7th Cir. 1986) (“We advise against defining ‘reasonable doubt’ because often the definition engenders more confusion than does the term itself.”); *United States v. Shaffner*, 524 F.2d 1021, 1023 (7th Cir. 1975) (“It is our opinion that any use of an instruction defining reasonable doubt presents a situation equivalent to playing with fire. The problem is that such an instruction is generally employed to favor one side.”); *United States v. Lawson*, 507 F.2d 433, 439-42 (7th Cir. 1974) (observing that attempts to clarify such common words generally result in further confusion).

43. See, e.g., *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.”).

44. In a celebrated case, the English Court of Appeal has recently endorsed both approaches as acceptable. See *R. v. Bentley*, [1999] Crim. L. Rev. 330 (finding reversible error in a jury instruction because it failed to set forth either the “reasonable doubt” standard or the “satisfied that you are sure” standard by which the jury was to have determined guilt).

45. *Lawson*, 507 F.2d at 443.

Judge Posner states the problem in more practical terms:

Undefined and unelaborated, the expression “proof beyond a reasonable doubt” requires, and is (I believe) understood to require, that the jury be certain of the defendants’ guilt, with this proviso: complete certainty—the certainty of such propositions as that cats do not grow on trees and that I have never set foot on Mars—is never attainable with respect to the question whether a criminal defendant is guilty of the crime for which he is being tried, and the jury should set aside doubts that it would be unreasonable to entertain given the practical limitations on attaining certainty in the trial setting. . . . The principal significance of the requirement of proving guilt beyond a reasonable doubt comes in cases where, because the screening has been careless or the prosecutor is improperly motivated, an innocent person has been indicted (the grand jury is no real screen). Unfortunately there are such cases—and we don’t want a quarter of them to end in conviction.⁴⁶

Posner’s views reflect a continued belief in the venerable values discussed earlier in this Article. There is good reason to believe, however, that use of the proof beyond a reasonable doubt standard without definition does not effectively prevent convictions when the government’s case is weak.⁴⁷

3. *Focusing on the Strength of the Government’s Case.*—The principal argument of this Article is that the proof beyond a reasonable doubt standard misfocuses the jury’s attention on whether the defense has proffered reasonable alternative explanations to the prosecution’s case, a burden that a defendant should not be required to meet. But not all explanations of the government’s burden of proof exhibit this problem. The Federal Judicial Center (“FJC”), in its pattern instruction, tells jurors that they must be “firmly convinced” of the defendant’s guilt before they may convict. While the instruction also uses the phrase “proof beyond a reasonable doubt,” it is almost entirely oriented toward explaining to the jury how strongly the government must prove its case.⁴⁸ This kind of approach has

46. *Hall*, 854 F.2d at 1044 (Posner, J., concurring). Posner’s reference to “a quarter of them” is an allusion to probabilistic thresholds for conviction associated with reasonable doubt. Quantification of reasonable doubt is discussed more fully *infra* Part III.

47. See *infra* Parts III, IV.

48. The Federal Judicial Center instruction reads:

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

been accepted in a few states,⁴⁹ and by some federal circuits.⁵⁰ Justice Ginsburg has specifically endorsed it in her concurring opinion in *Victor v. Nebraska*,⁵¹ as has Judge Newman of the United States Court of Appeals for the Second Circuit, in a much-cited law review article.⁵²

An instruction that both maintains “proof beyond a reasonable doubt,” and directs the jury to focus on the strength of the government’s case was adopted by the Supreme Court of New Jersey in *State v. Medina*.⁵³ The court adopted a modified version of the FJC instruction:

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 necessary to prove only 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.

A reasonable doubt is an honest and reasonable uncertainty in your minds about the guilt of the defendant after you have given full and impartial consideration to all of the evidence. A reasonable doubt may arise from the evidence itself or from a lack of evidence. It is a doubt that a reasonable person hearing the same evidence would have.

Proof beyond a reasonable doubt is proof, for example, that leaves you firmly convinced of the defendant’s guilt. In this world, we know very few things with absolute certainty. In criminal cases

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.

FEDERAL JUDICIAL CTR., *supra* note 4, No. 21.

49. *See, e.g.*, *State v. Portillo*, 898 P.2d 970, 974 (Ariz. 1995) (requiring that Arizona trial courts adopt the FJC “firmly convinced” standard in all criminal cases); *Winegeart v. State*, 665 N.E.2d 893, 902 (Ind. 1996) (endorsing the FJC “firmly convinced” instruction and recommending Indiana courts to use the instruction without embellishment); *Joyner-Pitts v. State*, 647 A.2d 116, 122-23 (Md. 1994) (holding that the trial court’s instructions were invalid because they failed to communicate clearly that the jury’s decision must be made “without reservation”).

50. *See, e.g.*, COMMITTEE ON MODEL JURY INSTRUCTIONS FOR THE NINTH CIRCUIT, NINTH CIRCUIT MANUAL OF MODEL JURY INSTRUCTIONS—CRIMINAL § 3.3 (1997) (defining proof beyond a reasonable doubt as “proof that leaves you firmly convinced that the defendant is guilty”); *United States v. Williams*, 20 F.3d 125, 128-32 (5th Cir. 1994) (approving of the district court’s use of the “firmly convinced” standard).

51. 511 U.S. 1, 26-27 (1994).

52. *See* Jon O. Newman, *Beyond “Reasonable Doubt,”* 68 N.Y.U. L. REV. 979, 980-81 (1993). Similarly, Justice Fortunato of Rhode Island recommends an instruction that tells jurors only to convict when “you are convinced in your mind that it is just about certain—or nearly certain—that the defendant committed the crime.” Stephen J. Fortunato, Jr., *Structuring on Reasonable Doubt After Victor v. Nebraska: A Trial Judge’s Certain Thoughts on Certainty*, 41 VILL. L. REV. 365, 427 (1996).

53. 685 A.2d 1242 (N.J. 1996).

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 are not firmly convinced of defendant's guilt, you must give defendant the benefit of the doubt and find him not guilty.⁵⁴

The instruction reminds jurors that not every doubt is a reasonable one. But its main focus is to tell the jurors that the real issue is whether the government has made a strong, convincing case. It is marginally superior to the FJC's instruction, which ends by indicating that there should be an acquittal "if there is a real possibility that [the defendant] is not guilty."⁵⁵ This last line in the FJC instruction tends to shift the burden of proof, as acknowledged by the Second Circuit.⁵⁶ The New Jersey version corrects this problem.⁵⁷

The FJC approach is by no means limited to American jurisdictions. Some English courts gave up on the reasonable doubt standard decades ago, and now instruct jurors that they should convict only if they are satisfied that they are sure.⁵⁸ Article 353 of the French Code of Criminal Procedure requires that written instructions be posted in the jury room, telling jurors that the only standard is that they be "thoroughly convinced" of the defendant's guilt.⁵⁹ For reasons discussed in Part IV of this

54. *Id.* at 1251-52.

55. *See supra* note 48.

56. *See United States v. McBride*, 786 F.2d 45, 51-52 (2d Cir. 1986) ("As for the district court's use of the 'real possibility' language . . . we suggest caution . . . as it may provide a basis for confusion and may be misinterpreted by jurors as unwarrantedly shifting the burden of proof to the defense."). *See infra* Part III for a discussion of how reasonable doubt instructions tend to shift the burden of proof. *See also* *Scott v. Class*, 532 N.W.2d 399, 405 (S.D. 1995) (holding that the "real possibility" language does not render the FJC instruction unconstitutional). My point is not whether the FJC instruction falls within a range of constitutionally acceptable jury instructions, which it clearly does. Rather, the point is that the problem can be easily corrected, as the New Jersey court has done.

57. In contrast to New Jersey, some courts begin with the FJC instruction and then add to it other language that reverts to the approach of defining which doubts are acceptable. The District of Columbia Court of Appeals, for example, after using most of the FJC instruction, goes on to tell the jury: "Reasonable doubt is the kind of doubt that would cause a reasonable person, after careful and thoughtful reflection, to hesitate to act in the graver or more important matters in life. However, it is not an imaginary doubt, nor a doubt based on speculation or guesswork; it is a doubt based upon reason." *Smith v. United States*, 709 A.2d 78, 82 (D.C. Cir. 1998).

58. *See Walters v. Queen* [1969] 2 App. Cas. 26, 30 (P.C. 1969) (appeal taken from Jam.) (holding that instructions should convey to the jury an understanding that, to find against the defendant, the jury must be sure of his guilt).

59. *See THE AMERICAN SERIES OF FOREIGN PENAL CODES: THE FRENCH CODE OF CRIMINAL PROCEDURE* art.353, at 172 (Gerald L. Kock & Richard S. Frase trans., rev. ed. 1988).

Before the assize court retires, the president shall read the following instruction, which, in addition, shall be posted in large letters in the most prominent place in the conference room: "The law does not ask an accounting from judges of the grounds by which they became convinced; it does not prescribe for them rules on which they must make the

Article, these approaches, which focus on what the government must establish, are less likely to lead to convictions when the government's case is weak.

III. Empirical Research: The Problems with Reasonable Doubt Instructions

Empirical evidence supports my suggestion that the expression "proof beyond a reasonable doubt" is likely to lead to conviction in some cases in which the government's evidence is unimpressive. This part of the Article will first examine a number of studies using both real and mock jurors reported over the past two decades. The experimental literature contains two types of studies: some compare conviction rates for different jury instructions, while others place probability values on different standards of proof. Both types of study strongly suggest that jurors are more likely to convict defendants on highly equivocal evidence when the instruction either attempts to explain which doubts are reasonable, or leaves reasonable doubt undefined. This part will then discuss criminal justice data from New Jersey, noting suggestive statistical changes following the 1996 *Medina* decision in which the Supreme Court of New Jersey ordered a radical change in burden of proof instructions.

A. *Experimental Studies Using Real and Mock Jurors*

1. *The Effects of Different Instructions.*—Many empirical studies support the claim that standard reasonable doubt instructions tend to shift the burden of proof to the defendant. A number of these further support the position that the FJC's firmly convinced instruction is less likely than standard reasonable doubt instructions to lead to conviction when the government's case is weak. One recent study questioned actual jurors in Wyoming after their participation in trials.⁶⁰ Wyoming judges use standard reasonable doubt instructions that do not define that term.⁶¹ One of the questions asked jurors whether they agreed with the statement that once the state has come forward with evidence of a defendant's guilt, it becomes

fullness and sufficiency of a proof particularly depend; it requires of them that they ask themselves, in silence and reflection to seek out, in the sincerity of the conscience, what impression the evidence reported against the accused and the ground of his defense have made on their reason. The law asks them only the single question, which encompasses the full measure of their duties: "Are you thoroughly convinced?"

60. See Bradley Saxton, *How Well Do Jurors Understand Jury Instructions? A Field Test Using Real Juries and Real Trials in Wyoming*, 33 LAND & WATER L. REV. 59 (1998).

61. The Wyoming revised criminal pattern jury instruction reads: "In order to convict the defendant of the crime charged, every material and necessary element to constitute such crime must be proved beyond a reasonable doubt and if the jury has a reasonable doubt on any necessary element, it is your duty to give the benefit of such doubt to the defendant and acquit him." *Id.* at 96 n.73 (quoting WYOMING PATTERN CRIMINAL JURY INSTRUCTIONS No. 1.03).

the defendant's responsibility to persuade the jury of his innocence.⁶² Of the jurors participating in the study, 30.5% said they were either very sure or pretty sure that this statement was true.⁶³ That is, when asked, almost one-third of the jurors consciously believed that the reasonable doubt standard does cause a shift in the burden of proof from the government to the defendant, despite instructions by the court explaining the presumption of innocence.⁶⁴

In another study, Irwin Horowitz, a psychology professor, and Laird C. Kirkpatrick, a law professor, compared juror reactions to various burden of proof instructions, including the FJC's firmly convinced instruction. Jury-eligible adults recruited from the community watched a mock murder trial presented on tape and slides.⁶⁵ Half of the subjects were presented with a "weak" case, in which the evidence was prescreened to be at the midpoint on a certainty of guilt scale. The other half were presented with a "strong" version, in which the evidence was predetermined to favor the prosecution by an 85-15 margin.⁶⁶ After hearing one of the stories, the jurors were divided into six-juror panels for

62. See *id.* at 141. The question read:

According to the instructions the judge gave you, is the following statement true or false: In a criminal trial, the state is responsible for producing evidence for the jury that tends to show that the defendant may have committed the crime—once the state has made this showing, it is the defendant's responsibility to produce witnesses or other evidence to persuade the jury that the defendant did not commit the crime.

Id. Jurors could respond with, "I'm VERY SURE that this statement is TRUE"; "I'm PRETTY SURE that this statement is TRUE"; "I DON'T KNOW"; "I'm PRETTY SURE that this statement is FALSE"; or "I'm VERY SURE that this statement is FALSE." *Id.*

63. See *id.* at 97.

64. For another study using actual jurors, see Geoffrey P. Kramer & Dorean M. Koenig, *Do Jurors Understand Criminal Jury Instructions? Analyzing the Results of the Michigan Juror Comprehension Project*, 23 U. MICH. J.L. REFORM 401 (1990). The instructions that these participants heard defined reasonable doubt as "not merely an imaginary doubt or a flimsy, fanciful doubt or a doubt based upon the mere possibility of the innocence of the defendant or a doubt based upon sympathy . . ." *Id.* at 436. But the instruction ends with: "[T]here must be such evidence that causes you to have a firm conviction to a moral certainty of the truth of the charge here made against the defendant." *Id.* The instruction is a confusing grab-bag, and the results of the study reflect this. Of the instructed jurors, 25.2% wrongly thought that they should acquit if they could see "any possibility, no matter how slight, that the defendant is innocent." *Id.* at 414. But 31.8% of the instructed jurors thought that "[a] reasonable doubt must be based only on the evidence that was presented in the courtroom, not on any conclusion that you draw from the evidence." *Id.* One cannot tell the extent to which the same jurors believed both of these erroneous propositions. The strongest conclusion one can draw is that the instructions left jurors confused.

65. See 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 LAW & HUM. BEHAV. 655, 659-61 (1996). This work is further presented in Irwin A. Horowitz, *Reasonable Doubt Instructions: Commonsense Justice and Standard of Proof*, 3 PSYCHOL. PUB. POL'Y & L. 285 (1997). The methodology was developed in part in Irwin A. Horowitz & David G. Seguin, *The Effects of Bifurcation and Death Qualification on Assignment of Penalty in Capital Crimes*, 16 J. APPLIED SOC. PSYCHOL. 165 (1986).

66. Horowitz & Kirkpatrick, *supra* note 65, at 661.

deliberation.⁶⁷ At that point the forty jury panels were divided into five separate groups of eight juries.⁶⁸ Each eight-jury group heard one of the following five reasonable doubt instructions: (1) moral certainty; (2) does not cause you to waver and vacillate;⁶⁹ (3) real doubt; (4) reasonable doubt undefined; and (5) firmly convinced, the Federal Judicial Center’s instruction.⁷⁰

Each of the three types of instruction discussed in Part II of this Article is represented in the Horowitz and Kirkpatrick study. The first three attempt to define reasonable doubt, the fourth leaves reasonable doubt undefined and the fifth is the Federal Judicial Center’s instruction. The results are presented below in Table 1:

TABLE 1
Guilty Verdicts as a Function of Evidentiary Strength
and Reasonable Doubt Instructions (From Horowitz and Kirkpatrick)⁷¹

	Definitions of Reasonable Doubt			Reasonable Doubt Undefined Instruction	FJC
Weak Case					
	Moral Certainty	Waiver or Vacillate	Real Doubt	Undefined	Firmly Convinced
Guilty	4	4	3	4	0
Not Guilty	4	3	3	2	8
Hung Jury	0	1	2	2	0

67. See *id.* at 659.

68. See *id.* at 664.

69. Horowitz & Kirkpatrick call this instruction, “WV,” since it says that “[p]roof beyond a reasonable doubt is proof that means that you do not *waver* or *vacillate* concerning the defendant’s guilt.” *Id.* at 660. But the instruction then goes on at length explaining that not all doubt is reasonable doubt and ends by saying: “If . . . you do not *waver* or *vacillate* that the defendant is guilty of the crime charged, you must find him guilty. If . . . 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.’” *Id.* at 660. This instruction obviously gives the jury mixed and confusing messages.

70. *Id.* at 660-61.

71. *Id.* at 663 tbl.1. The data are from Table 1 of their Article, but has been reorganized here for presentational purposes. The data are repeated in Horowitz, *supra* note 65, at 296.

Strong Case					
	Moral Certainty	Waiver or Vacillate	Real Doubt	Undefined	Firmly Convinced
Guilty	8	5	6	5	6
Not Guilty	0	1	1	1	2
Hung Jury	0	2	1	2	0

For the strong case, the category of instruction did not matter significantly. Jurors convicted most of the time across the board, the actual rate of conviction varying slightly as a function of the number of hung juries for each condition. But for the weak case, the only instruction for which there were no convictions was the Federal Judicial Center's. For all others, there were convictions about half the time. This is startling—it means that the prosecutor had a fifty percent chance of getting a conviction even when a case was designed to be as neutral as possible with respect to a defendant's guilt. Even if one is skeptical about the actual percentages, the comparisons are striking. Jurors perform differently in response to the firmly convinced standard, which keeps the focus on the government's burden of proof.

A study by Robert J. MacCoun and Norbert L. Kerr provides further support for this position.⁷² Using students as subjects, they set out to investigate whether deliberation makes jurors more lenient.⁷³ Harry Kalven, Jr. and Hans Zeisel had found in their 1966 book that juries acquitted some defendants that judges would have convicted had there been a bench trial.⁷⁴ MacCoun and Kerr attempted to investigate the source of this phenomenon. They presented subjects with an audiotape of an auto theft trial along with a transcript and various photographs and exhibits.⁷⁵ The structure of the study was to have subjects put into four-person juries in which half of the jurors believed the defendant to be guilty prior to deliberation, and the other half believed him to be innocent prior to

72. Robert J. MacCoun & Norbert L. Kerr, *Asymmetric Influence in Mock Jury Deliberation: Jurors' Bias for Leniency*, 54 J. PERSONALITY & SOC. PSYCHOL. 21 (1988).

73. See *id.* at 21-22.

74. KALVEN & ZEISEL, *supra* note 31, at 55-59.

75. See MacCoun & Kerr, *supra* note 72, at 27.

deliberation.⁷⁶ If significantly more of the juries acquitted, then the study would support the leniency effect.

MacCoun and Kerr quickly encountered a design difficulty. If the case was too strong, it was difficult to find enough jurors who initially believed that the defendant was innocent. They dealt with this problem as follows: "In order to facilitate the formation of 2:2 mock juries, the trial transcript was modified until pretesting indicated that it was as equivocal as possible regarding the guilt of the defendant. The overall 55% conviction rate in the present study indicates that this effort was successful."⁷⁷ In other words, MacCoun and Kerr intentionally used a weak case, a tactic frequently employed in mock jury experiments to avoid "swamp[ing] the (usually more subtle) variables that are the focus of the study."⁷⁸ The reasonable doubt instruction contained the statement that "[a] reasonable doubt about the defendant's guilt must be a fair one, one based on reason, and one for which reasons can be given."⁷⁹ This instruction attempts to define which doubts are reasonable. Unlike the FJC instruction, it focuses the jury on the defendant's evidence, not on the government's burden.

Half of MacCoun and Kerr's mock juries were given a preponderance of the evidence instruction and half a reasonable doubt instruction. The results are presented below in Table 2.

TABLE 2
Group Outcomes for Each Standard of Proof
(From MacCoun and Kerr)⁸⁰

Outcome	Standard of Proof	
	Reasonable Doubt	Perponderance
Guilty	10	15
Not Guilty	28	21
Hung Jury	9	6

The twenty-six percent conviction rate (10 of 38 juries reaching a verdict) is consistent with the proposition that such instructions lead to unacceptably

76. *See id.* at 28.

77. *Id.* at 27.

78. Michael J. Saks, *What Do Jury Experiments Tell Us About How Juries (Should) Make Decisions?*, 6 S. CAL. INTERDIS. L.J. 1, 19 (1997).

79. MacCoun & Kerr, *supra* note 72, at 28.

80. *See id.* at 29 tbl.6.

high rates of conviction in weak cases. In contrast, these results support a "leniency hypothesis" only in the most ironic sense: When a case is so weak that important values underlying our legal system tell us there should be no conviction, the acquittal rate is about three quarters.⁸¹

Two other studies suggest that the proof beyond a reasonable doubt standard leads to more convictions in weak cases than a standard of proof focused on what the government must establish. In a study conducted at the London School of Economics, mock jurors selected from the community were asked to deliberate after hearing simulated trials in a rape case with two defendants and a theft case.⁸² The trials were "selected for their brevity and uncertainty,"⁸³ meaning that they were not strong cases. Some jurors received definitional "reasonable doubt" instructions of the sort criticized earlier in this Article.⁸⁴ Others were told not to convict unless "you . . . feel sure and certain on the evidence you have heard that the accused is guilty."⁸⁵ A third group was given a preponderance of the evidence instruction.

The results for the rape case are just as we should expect. There were significantly fewer convictions for each of the two defendants from juries that heard the "sure and certain" instruction (55% and 18%, respectively, for the two defendants) than from juries who heard the "proof beyond a reasonable doubt" instruction (66% and 32%, respectively, for the two defendants).⁸⁶ Again, the London School of Economics experiment supports the notion that instructions that focus on the government's proof lead to fewer convictions in weak cases than instructions that focus on how strongly the defendant has raised doubt.

81. MacCoun & Kerr's subjects were students, who tend to be more lenient than members of jury pools. See *infra* note 110. This may explain, at least in part, the fact that their experiment shows lower conviction rates than do Horowitz & Kirkpatrick's experiments. See *supra* note 71 and the accompanying table. I do not claim that MacCoun & Kerr's study as reinterpreted here disconfirms Kalven & Zeisel's observation that jurors tend to be more lenient than judges. See KALVEN & ZEISEL, *supra* note 31, at 68. Nothing here or in MacCoun and Kerr's experiments argues against that observation.

82. See L.S.E. Jury Project, *Juries and the Rules of Evidence*, 1973 CRIM. L. REV. 208, 210-11.

83. *Id.* at 210.

84. The instruction for the rape cases said in part: "you should be sure beyond reasonable doubt and by reasonable doubt I mean not a fanciful doubt that you might use to avoid an unpleasant decision, but a doubt for which reasons can be given." *Id.* at 213.

85. *Id.* This instruction was used as an intermediate burden of proof, similar to the "clear and convincing" standard used in the United States. See *Addington v. Texas*, 441 U.S. 418, 424 (1978) (observing that the "clear and convincing" standard of proof is generally used in civil cases where the "interests at stake are . . . deemed to be more substantial than the mere loss of money"); see also MCCORMICK, EVIDENCE § 340, 441-45 (4th ed. 1992) (listing the kinds of cases in which the "clear and convincing" standard is applied).

86. See L.S.E. Jury Project, *supra* note 82, at 216-17.

Finally, in another study by Norbert Kerr and his colleagues,⁸⁷ 645 subjects (again psychology students) were shown a fifty minute simulation of a rape trial which had produced "strong, predeliberation disagreement."⁸⁸ Subjects were subsequently divided into six-person juries. One-third of the juries received a "lax" reasonable doubt instruction, which included the requirement that for a doubt to be reasonable, it must be possible to state a reason for it.⁸⁹ The "stringent" reasonable doubt instruction, also given to one-third of the juries, told the jurors that they must acquit if they can come up with an alternative model of the facts.⁹⁰ The remaining jurors received no definition at all.

The results are dramatic. While only 2 of 19 juries hearing the stringent condition (alternative model of the facts) voted unanimously for conviction (10.5%), 8 of 19 juries hearing the lax condition (possible to state a reason for doubt) voted unanimously to convict (42.1%).⁹¹ Of those juries that received no definition, 3 of 11 (27.3%) voted for conviction.⁹²

The conditions in Kerr *et al.*'s study correspond to the three approaches discussed in Part II of this Article, although their lax instruction is more lax than those used in most jurisdictions today, and their stringent instruction is more stringent than the Federal Judicial Center's. Nonetheless, the results are consistent with the results of all the other studies: Juries convict less often when they are asked to focus on how well the government has proven its case instead of how well the defense has established doubt.

2. *Probability Studies.*—A second type of experimental data supports the view that instructions that focus on doubt instead of the government's burden increase conviction rates. These studies ask jurors how certain they think they should be on a numerical scale before they convict. If we accept Blackstone's 10 to 1 error rate in favor of false acquittals,⁹³ then

87. See Norbert L. Kerr et al., *Guilt Beyond a Reasonable Doubt: Effects of Concept Definition and Assigned Decision Rule on the Judgments of Mock Jurors*, 34 J. PERSONALITY & SOC. PSYCHOL. 282 (1976).

88. *Id.* at 286.

89. The lax instruction read in part: "A reasonable doubt about the defendant's guilt must be a substantial one, a fair one, one based on reason, and one for which reasons can be given. In summary you need not be absolutely sure that the defendant is guilty to find him guilty." *Id.*

90. The stringent instruction read in part: "That is, before you can return a verdict of guilty you must be sure and certain that the defendant is guilty. . . . If you feel that the facts of this case are compatible with any other theory besides the one in which the defendant is guilty, then you have a reasonable doubt about his guilt and must find him not guilty." *Id.*

91. See *id.* at 287 th.1.

92. See *id.* The experiment contained other conditions that concerned the effects of jury instructions on juries when unanimous votes are not required. I do not discuss those results here.

93. See BLACKSTONE, *supra* note 22, at 358.

we might expect reasonable doubt to require about 91% certainty. Three published surveys have asked sitting judges how certain they think jurors should be before convicting beyond a reasonable doubt. One was conducted by Judge Weinstein with federal judges sitting in the Eastern District of New York, and is reported in his opinion in *United States v. Fatico*.⁹⁴ Ten judges were polled. Nine gave a numerical probability associated with reasonable doubt. Of the nine judges, one gave a probability of 76%, one gave 80%, four gave 85%, two gave 90%, and one gave 95%.⁹⁵ In other words, the probabilities hovered around 85%-90%.

A second study asked federal judges throughout the United States to quantify the conviction threshold for reasonable doubt.⁹⁶ Of the 171 judges who responded, 126 had thresholds that were 90% or higher.⁹⁷ Eleven judges had thresholds of 75% or below, one of whom was satisfied with a 50% probability.⁹⁸ The other study was conducted among Illinois state court judges.⁹⁹ On a scale of 1 to 10, the mean level of certainty in this study was 8.9, with a median of 8.8; 63% of the judges responded with a level of 9.0 or higher.¹⁰⁰ Most (but not all) judges, then, tend to see the government's burden much the way Blackstone did.

Our courts do not use jury instructions based on percentages of certainty, and I do not suggest here that they begin doing so.¹⁰¹ Nonetheless, studies that present jurors with different burden of proof instructions and then, prior to deliberation, ask them to quantify the threshold of certainty needed for a conviction can act as a window into how different approaches to burden of proof influence jurors' understandings of their task. Similarly, studies that compare rates of conviction between

94. 458 F. Supp. 388 (E.D.N.Y. 1978).

95. *See id.* at 410.

96. *See C.M.A. McCauliff, Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees?*, 35 VAND. L. REV. 1293, 1325 (1982).

97. *See id.*

98. *See id.*

99. *See 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 (1971). Judges were asked: "What would the likelihood or probability have to be that a defendant committed the act for you to decide that he is guilty?" *Id.* at 324.

100. *See id.* at 324. Potential jurors responding to the same question displayed somewhat lower standards for the government. The median was 8.6, the mean, 7.9; 54% of the jurors responded with 9.0 or higher. *See id.*

101. For arguments against creating a quantitative approach to burden of proof, see generally Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971). In this regard I disagree with Saks, *supra* note 78, at 31, who suggests that juror comprehension of burden of proof instructions be improved by providing them with quantified definitions. To support his claim that these quantified definitions help, Saks relies on Dorothy K. Kagehiro & W. Clark Stanton, *Legal vs. Quantified Definitions of Standards of Proof*, 9 LAW & HUM. BEHAV. 159, 169 (1985) (finding that "[i]n terms of their intended effects on verdicts, the quantified definition worked"), which is discussed *infra* note 111.

jurors given probability thresholds on the one hand, and jurors given reasonable doubt instructions on the other, can be instructive. Both kinds of studies exist, and both suggest that instructions focusing on the government's burden instead of on definitions of reasonable doubt decrease the rate of conviction in weak cases.

For example, Horowitz and Kirkpatrick asked their subjects (prospective jurors), after hearing the trial and the instructions but before deliberating, to quantify the government's burden of proof required for conviction. Recall that Horowitz and Kirkpatrick presented both strong and weak cases to their subjects. Here, I focus only on responses in the weak cases since those are the cases in which we should not see convictions. The average response was an astonishingly low 61.122%.¹⁰² When asked again after deliberation, subjects responded similarly, at 61.574%.¹⁰³ Of all the instructions, only the firmly convinced standard produced a higher threshold. Pre-deliberation, the average response was 68.87%, and post-deliberation, the average response was 80.75%.¹⁰⁴ In contrast, when the reasonable doubt instruction told jurors that they must have a "real doubt," the pre- and post-deliberation probability thresholds were 68.25% and 61.62%, respectively.¹⁰⁵ Instructions leaving reasonable doubt undefined produced a pre-deliberation threshold of 52.87%, and a post-deliberation threshold of 55.00%,¹⁰⁶ well below any level of certainty that our system claims to tolerate.

Kerr *et al.* conducted a similar survey. On an eleven-point scale, subjects were asked for the "probability of guilt necessary for conviction beyond a reasonable doubt."¹⁰⁷ The lax (reasonable doubt requires a reason that can be stated) and undefined reasonable doubt instructions both produced threshold probabilities of 82%.¹⁰⁸ The stringent, government-oriented instruction (jurors must acquit if they can come up with an alternative model of the facts) produced a threshold of 87%, significantly higher.¹⁰⁹ One can speculate about why the absolute numbers vary so much between the two sets of studies. One factor, no doubt, is that Kerr *et al.* used student subjects, while Horowitz and Kirkpatrick used prospective jurors. Students tend to be more lenient than the population as a whole.¹¹⁰ Regardless of the absolute numbers, the relationship between

102. See Horowitz & Kirkpatrick, *supra* note 65, at 664.

103. See *id.*

104. See *id.*

105. See *id.*

106. See *id.*

107. Kerr *et al.*, *supra* note 87, at 290.

108. See *id.*

109. See *id.*

110. See Simon & Mahan, *supra* note 99, at 324 tbl.4 (finding the mean probability threshold of students at 8.9, and the mean probability threshold of jurors at 7.9, based on a 10-point scale).

threshold probability and the type of reasonable doubt instruction remains constant.

In a set of experiments by Dorothy K. Kagehiro and W. Clark Stanton, psychology students were presented with a summary of a civil trial.¹¹¹ At the end of the trial, some were given a reasonable doubt instruction that defined reasonable doubt in terms of moral certainty, and others were told that they should render a verdict for the plaintiff only if they had at least a 91% degree of certainty.¹¹² In the first experiment, subjects then stated what they thought the verdict should be, without deliberation. The results were plaintiff's verdicts 43% of the time with the reasonable doubt instruction, but only 31% of the time with the quantified instruction, a difference of 12%.¹¹³ A second experiment was conducted the same way, but using the Federal Judicial Center's firmly convinced instruction instead of the moral certainty instruction. The rate of plaintiff's verdicts for those hearing the quantified instruction (requiring 91% certainty) was 36%, and for those hearing the firmly convinced instruction the rate was 41%.¹¹⁴ The difference between the two was largely eliminated. This adds further support to the notion that the firmly convinced standard comes closest to meeting the stated values underlying the notion of proof beyond a reasonable doubt.

In yet a third experiment, Kagehiro and Stanton gave their subjects a variety of reasonable doubt instructions.¹¹⁵ Among other tasks, subjects were asked to say how hard it should be on a 1-9 scale to convict under the standard presented to them.¹¹⁶ The quantified reasonable doubt instruction of 91% certainty produced a score of 6.1.¹¹⁷ The firmly convinced instruction produced a score of 6.9.¹¹⁸ The others produced scores ranging from 3.8 to 5.1,¹¹⁹ reflecting a significantly lower burden on the government. In these experiments, only the firmly convinced standard even approaches the assumed burden of proof.

Interestingly, Kagehiro and Stanton also looked at instructions telling the jury what it means to decide a case based on clear and convincing evidence. Although this standard is supposed to be an intermediate burden of proof,¹²⁰ I would predict that defendant-oriented reasonable doubt instructions actually place less burden on the government than do clear and

111. See Kagehiro & Stanton, *supra* note 101, at 162.

112. See *id.* at 163.

113. See *id.* at 164 tbl.1.

114. See *id.* at 169 tbl.3.

115. See *id.* at 170.

116. See *id.* at 171.

117. See *id.* at 173 tbl.4.

118. See *id.*

119. See *id.*

120. See *supra* note 85.

convincing instructions. The latter focus the jury on the government's burden, while the former focus the jury on the defendant's ability to come up with alternative explanations. The Colorado jury instructions tested by Kagehiro and Stanton have these characteristics. The clear and convincing instruction focuses on the government's burden,¹²¹ while the reasonable doubt instruction focuses on the quality of doubts acceptable to avoid conviction.¹²² Subjects hearing these instructions gave the clear and convincing evidence a 6.3 probability threshold on the 1-9 scale, but gave the reasonable doubt instruction only a 3.8 probability threshold.¹²³ The Federal Judicial Center's firmly convinced instruction, in contrast, was considered more demanding than the corresponding federal clear and convincing evidence standard (6.9 vs. 5.6).¹²⁴

B. *Real World Data: The New Jersey Experience*

As discussed above, in 1996 the Supreme Court of New Jersey rejected a reasonable doubt instruction that focused on defining which doubts are reasonable¹²⁵ and replaced it with a modified version of the Federal Judicial Center's instruction.¹²⁶ The questions asked here are what changes, if any, have there been in New Jersey's criminal justice statistics since that time, and whether we can attribute any such changes to the increase in the government's burden of proof.

Prior to 1996, the conviction rate in New Jersey criminal cases was about two-thirds. Since the switch, it has remained about the same¹²⁷—the small fluctuation that has occurred is not statistically significant.¹²⁸

121. The instruction reads: "[Proof by clear and convincing evidence is] proof which is stronger than a 'preponderance of the evidence' and which is highly probable and free from serious or substantial doubt." *Id.* at 177.

122. The reasonable doubt instruction reads in part: "It is doubt which is not vague, speculative or imaginary doubt, but such a doubt as would cause reasonable persons to hesitate to act in matters of importance to themselves." *Id.*

123. *See id.* at 173 tbl.4.

124. The federal clear and convincing evidence instruction used in the experiment required "proof which is 'clear' in the sense that it is certain, plain to the understanding, unambiguous, and 'convincing' in the sense that it is so reasonable and persuasive as to cause you to believe it." *Id.*

125. *See State v. Medina*, 685 A.2d 1242, 1246-48 (N.J. 1996).

126. *See id.* at 1251-52.

127. In the year ending June 1996, there were 651 acquittals and 1305 convictions in criminal trials for a 33.28% acquittal rate. *See* N.J. ADMIN. OFFICE OF THE COURTS, CRIMINAL DIV., 1995-1996 ANN. REP. [hereinafter 1995-1996 ANN. REP.]. In the year ending June 1997, the rate was 32.71% (616 acquittals and 1267 convictions). *See* N.J. ADMIN. OFFICE OF THE COURTS, CRIMINAL DIV., 1996-1997 ANN. REP. [hereinafter 1996-1997 ANN. REP.]. In the year ending June 1998, the rate was 35.18% (635 acquittals and 1170 convictions). *See* N.J. ADMIN. OFFICE OF THE COURTS, CRIMINAL DIV., 1997-1998 ANN. REP. [hereinafter 1997-1998 ANN. REP.].

128. A chi-square analysis was performed: $X^2 = 2.1165$, $df = 2$, $p > .25$. The chi-square compares actual frequencies with expected frequencies, given a null hypothesis. Here, the null hypothesis is that the rate of conviction would remain constant after the *Medina* decision. The actual

To see why the rate has remained the same, let us hypothesize that the typical New Jersey prosecutor is willing to take a case to trial if she believes that she has a two out of three chance of getting a conviction. Assuming that this level of risk taking remains constant even after the New Jersey Supreme Court increased the government's burden of proof, we would not expect a lower rate of conviction. Rather, all other things being equal, fewer cases would go to trial, with more dismissals and more guilty pleas to reduced charges.¹²⁹ Some cases that a prosecutor might have been willing to take to trial before 1996 have become less attractive after *Medina*. This is exactly what the data suggest. Although the total number of criminal cases increased 5.7% between 1996 and 1998, 7.8% fewer cases went to trial.¹³⁰ This decrease is statistically significant.¹³¹ In contrast, more cases were terminated after guilty pleas to downgraded charges.¹³² Dismissals remained about constant during these years.¹³³ Although there are too many variables to permit strong causative claims,¹³⁴ the data certainly do not provide disconfirming evidence—

conviction rates did not vary enough from the expected rates to require that the null hypothesis be rejected. The chi-square is explained in virtually every text on statistics. See, e.g., JAMES BROOK, *A LAWYER'S GUIDE TO PROBABILITY AND STATISTICS* 197-208 (1990).

129. For a discussion of the relationship between case selection and win rates in the context of civil cases, see Kevin M. Clermont & Theodore Eisenberg, *Do Case Outcomes Really Reveal Anything About the Legal System? Win Rates and Removal Jurisdiction*, 83 CORNELL L. REV. 581 (1998).

130. While the number of cases in 1997 increased to 115,877 from 110,051 in 1996, the total number of trials decreased slightly from 1,956 to 1,883. See 1995-1996 ANN. REP., *supra* note 127; 1996-1997 ANN. REP., *supra* note 127. For 1998, the total number of cases increased further to 116,269, and the number of trials had decreased further, to 1805. See 1997-1998 ANN. REP., *supra* note 127. Given the 5.7% increase in the number of cases, the null hypothesis would be a commensurate increase in the number of trials. But the number of trials fell 7.8% (a decrease of 151 from 1,956 trials).

131. $\chi^2=48.4178$, $df=2$, $p<.001$. The null hypothesis for the analysis is that the number of trials will vary with the number of cases.

132. In the year ending June 1998, 47,995 cases were terminated after guilty pleas to downgraded charges. See 1997-1998 ANN. REP., *supra* note 127. In the year ending June 1997, 48,375 cases were terminated after guilty pleas to downgraded charges. See 1996-1997 ANN. REP., *supra* note 127. In the year ending June 1996, 43,570 cases were terminated after guilty pleas to downgraded charges. See 1995-1996 ANN. REP., *supra* note 127.

133. Pre-indictment dismissals numbered 15,264 in 1996 (13.87% of the total cases), see 1995-1996 ANN. REP., *supra* note 127; 14,989 in 1997 (12.99% of the total cases), see 1996-1997 ANN. REP., *supra* note 127; and 14,582 in 1998 (12.54% of the total cases), see 1997-1998 ANN. REP., *supra* note 127. Post-indictment dismissals numbered 6,899 in 1996 (6.27% of the total), see 1995-1996 ANN. REP., *supra* note 127; 6,956 in 1997 (6.0% of the total), see 1996-1997 ANN. REP., *supra* note 127; and 6,937 in 1998 (5.99% of the total), see 1997-1998 ANN. REP., *supra* note 127.

134. For a discussion of some of the dangers of drawing strong conclusions from such data, see Neil Vidmar, *Pap and Circumstance: What Jury Verdict Statistics Can Tell Us About Jury Behavior and the Tort System*, 28 SUFFOLK U. L. REV. 1205 (1994). Here, for example, the number of criminal cases taken to trial had fallen in some years prior to the New Jersey Supreme Court's *Medina* decision, demonstrating that there can be many reasons for changes in this statistic. Moreover, we do not know in any detail whether the range of negotiated pleas has changed, as would be predicted.

neither conviction rates nor the number of cases going to trial have increased following the *Medina* decision.¹³⁵

C. Summary

Although some of these studies are better designed than others, there are simply too many data points to ignore. Most problems with the experiments, such as the sophistication of the trial materials, the population of subjects, and the fact that subjects knew they were participating in an experiment, go to the reliability of the absolute scores. Nonetheless, within each experimental setting, the relationships among the reactions to various instructions reflect differences in how the subjects understand the instructions with respect to the experimental task. The studies show the same relationships among the different instructions regardless of the task.¹³⁶ One study after another, when analyzed in terms of the observations made earlier in this Article, shows that jurors convict more often in weak cases when the burden of proof centers around the meaning of doubt. The studies further show that jurors set unacceptably low conviction thresholds when the instruction focuses on the defendant, and not on the government. The conclusion is clear: courts should abandon all efforts to tell jurors which doubts are legitimate and which ones are not. Instead, they should tell jurors not to convict unless the government has proven its case to a very high degree of certainty. This is the standard that earlier

135. Criminal justice statistics for the federal courts are more plentiful. See, e.g., DIR. OF THE ADMIN. OF THE U.S. COURTS ANN. REP. 204 tbl.D-7 (1990). Differences in the demographics among the districts, combined with the fact that the courts of appeals and the Supreme Court accept a wide range of reasonable doubt instructions from the trial courts, make it particularly difficult to draw sound conclusions from the federal data. Moreover, these results are analogous to other empirical evidence of what happens when the burden on the government is increased in the criminal process. For example, studies suggest that the number of criminal suspects who confess has not changed as a result of the Supreme Court's 1966 *Miranda* decision. See *Miranda v. Arizona*, 384 U.S. 436 (1966) (requiring police officers to warn suspects about their right to remain silent). For a discussion of these studies, see George C. Thomas III, *Is Miranda a Real-World Failure? A Plea for More (and Better) Empirical Evidence*, 43 UCLA L. REV. 821 (1996).

In an interesting article, George Thomas suggests that while the rate of confession may not have changed, police attitude and conduct have. See Thomas, *supra*, at 832-33 (citing a study concluding that when *Miranda* warnings are given, police tend to be "more serious about the interrogation"). If so, then the New Jersey experience discussed in this section may be a special case of a larger phenomenon by which government officials adjust their conduct in reaction to new assessments of risk. Without question, more work in this area is called for, and is occurring. See, e.g., Paul G. Cassell & Bret S. Hayman, *Police Interrogation in the 1990s: An Empirical Study of the Effects of Miranda*, 43 UCLA L. REV. 839 (1996) (analyzing the results of a study done in Utah regarding police questioning and confessions).

136. "Every type of study, and every individual study, inevitably will be imperfect. The basic solutions to those imperfections are replication and triangulation." Saks, *supra* note 78, at 5. "Triangulation" occurs "[i]f studies employing different methodologies produce similar results, [and therefore] we gain increasing confidence in the conclusions." *Id.* (footnote omitted). In the various experiments I have examined we have both replication and triangulation.

generations in our history intended, and this is the standard to which criminal defendants are entitled.

IV. Why We Should Expect the Reasonable Doubt Standard to Lead to Convictions in Weak Cases: A Linguistic Analysis

Recent advances in linguistics, philosophy of language, and cognitive psychology explain the empirical results presented above. For several reasons, the expression "proof beyond a reasonable doubt" is not likely to accomplish the policy goal articulated in Part II that conviction occur only when the jury has the "highest degree of certitude"¹³⁷ about the defendant's guilt.

A. *Doubting and the Creation of Mental Models*

The verb "to doubt" is what linguists and philosophers call a verb of propositional attitude. It describes the speaker's attitude toward a proposition. Loosely, it means something like, "to think a proposition is not true," as in: "Do you think it will rain tomorrow?" "I doubt it."¹³⁸ Justice Scalia's recent characterization of the cognate noun, "an uncertain, tentative, or provisional disbelief,"¹³⁹ gleaned from various dictionaries, describes this sense of the concept well.¹⁴⁰ The word is a label for a cognitive process that occurs innately. That is, if our language had no verb, "to doubt," we would know how to doubt anyway, and would just have to use different language to describe the mental process.¹⁴¹

Philosophers of language and some linguists have, over the past several decades, explained verbs of propositional attitude through "possible

137. *Victor v. Nebraska*, 511 U.S. 1, 11 (1994).

138. The verb "doubt" has 3 senses in WordNet:

1. doubt—(consider unlikely or have doubts about)
2. suspect, distrust, doubt—(suspect to be false; 'I distrust that man')
3. doubt—(lack confidence in; 'I doubt these reports').

Search of WordNet 1.6, Cognitive Science Laboratory of Princeton University (visited Sept. 9, 1999) <<http://www.cogsci.princeton.edu/cgi-bin/webwn/?stage=1&word=doubt>>.

139. *Allentown Mack Sales & Service, Inc. v. NLRB*, 118 S.Ct. 818, 823 (1998). That case concerned the interpretation of the phrase, "good faith reasonable doubt." *See id.* at 820. The NLRB requires that an employer have a good faith reasonable doubt that a union representing its workers continues to enjoy majority support before it is permitted to conduct an election to determine whether that support continues to exist. *See id.* The applicability of the reasonable doubt standard in labor law is outside the scope of this Article. However, the analysis presented here suggests that the standard is inappropriate. It would make more sense for the NLRB to state the standard in terms of a level of certainty rather than in terms of doubt. My agreement with Justice Scalia's characterization of "doubt" does not extend to agreement with his interpretation of the standard's applicability in NLRB proceedings.

140. "Doubt" as a noun is actually ambiguous. It also is used to refer to the things and events that lead one to have doubt. *See infra* text accompanying notes 159-61.

141. For a recent discussion of why such concepts should be considered innate regardless of their linguistic manifestation, see JERRY A. FODOR, *CONCEPTS: WHERE COGNITIVE SCIENCE WENT WRONG* 120-45 (1998).

world” semantics.¹⁴² According to this perspective, the meaning of a proposition is the set of conditions in which it is true. Thus, “it is raining” is true if it is raining, and the meaning of that sentence, therefore, is a description of the state of affairs in the world when it is raining. But what are the truth conditions for counterfactual statements such as “If it weren’t raining, then I would be outside”? Such sentences are true if I am outside in some possible world that is just like the real world except for the fact that it is not raining in this possible world. Verbs of propositional attitude, like *know*, *believe*, *doubt*, *think*, and so on, can be similarly described. A sentence “B doubts p” is true if B thinks that p is not true in the real world, even though it might be true in possible worlds, which are just like the real world, except for p. When I doubt that it is raining, I usually have in mind that it might be raining, but I think it really isn’t raining.

In the past two decades, linguists and psychologists have constructed theories that capture and claim psychological reality to the insights of the philosophical literature. Philip Johnson-Laird’s “mental models,”¹⁴³ Ray Jackendoff’s “conceptual structure,”¹⁴⁴ Gilles Fauconnier’s “mental space,”¹⁴⁵ and George Lakoff’s “idealized cognitive models,”¹⁴⁶ all attempt to capture the notion that in thinking and understanding language we construct models into which we fit new information and draw inferences. If I doubt it is raining, I have created a model of the world in which it is raining, but I think that this model does not fit the reality of the world at the time.¹⁴⁷

Seen this way, the burden of proof in criminal cases should require the government to do two things. First, the government must create a model of the events in question that strongly supports its claim of guilt. The literature on jury behavior speaks of jurors being presented with conflicting stories, which is consistent with the mental model approach.¹⁴⁸ Second, the government must prove its case to the point that the jurors cannot reasonably create alternative mental models of the evidence inconsistent with the government’s case. In other words, the juror has to be able to say to herself, “the government’s proof is so strong that I can’t reasonably

142. For a brief discussion of this perspective, see PHILIP N. JOHNSON-LAIRD, *MENTAL MODELS* 56-61 (1983). For a more technical discussion, see SAUL A. KRIPKE, *NAMING AND NECESSITY* (1980); DAVID LEWIS, *COUNTERFACTUALS* 84-91 (1973). For a readable explanation of model-theoretic semantics, see EMMON BACH, *INFORMAL LECTURES ON FORMAL SEMANTICS* (1989).

143. See JOHNSON-LAIRD, *supra* note 142, at 53.

144. See RAY S. JACKENDOFF, *SEMANTICS AND COGNITION* 17 (1983).

145. See GILLES FAUCONNIER, *MENTAL SPACES* 16 (1994).

146. See GEORGE LAKOFF, *WOMEN, FIRE AND DANGEROUS THINGS* (1986).

147. For an interesting article that discusses judicial reasoning in terms of the formation of mental models, see Dan Simon, *A Psychological Model of Judicial Decision Making*, 30 *RUTGERS L.J.* 1, 77-85 (1998).

148. See REID HASTIE ET AL., *INSIDE THE JURY* 15-36 (1983).

imagine that he didn't do it." Some older reasonable doubt instructions, no longer in use, put the burden of proof in just those terms.¹⁴⁹

Of course, these two tasks are not unrelated. The stronger the government's case, the harder it is to come up with reasonable alternative accounts. However, they are not exactly the same thing. It is entirely possible for the government to put on a weak case for which an innocent defendant is unable to produce an alternative story that is not entirely speculative. One problem with "proof beyond a reasonable doubt" is that it focuses too much on the second of these tasks, at the expense of the first. The phrase encourages jurors to take the government's case as a given and then challenges them to find alternatives. None of this is a problem when the government's proof is strong and convincing, but when the government's case is weak, this focus asks too little of the government and does not support the goals the reasonable doubt standard is intended to promote.¹⁵⁰

Once the jury has found the facts by creating whatever models it does, it must decide how to match the model it has accepted with the elements of the crime of which the defendant is accused. In criminal cases, jurors are not limited to fact-finding. They also determine whether the facts that they find constitute a crime as the judge defines it.¹⁵¹ The legal categories are themselves mental models, and the jurors must form "a classification in which the best match between the story features and verdict category features is determined."¹⁵² For reasons I discuss below, this task, too, is problematic. The tendency is to find the closest match, rather than the match in which the jurors are certain that the government has proven every element of the crime.¹⁵³

B. Conceptual Problems with the Reasonable Doubt Standard

1. *The Government's Burden Remains Unquantified.*—One problem with the expression "proof beyond a reasonable doubt" is that it fails to address how much proof is needed before we start thinking about doubts. To see the problem as a special case of a more general cognitive

149. See *Hopt v. Utah*, 120 U.S. 430, 439 (1887) (approving an instruction to the jury "[t]hat if you can reconcile the evidence before you upon any reasonable hypothesis consistent with the defendant's innocence, you should do so, and in that case find him not guilty.>").

150. The goals of the standard are articulated in Judge Posner's concurrence in *United States v. Hall*, 854 F.2d 1036, 1043 (7th Cir. 1988), discussed *supra* note 46.

151. See *HASTIE ET AL.*, *supra* note 148, at 22.

152. *Id.*

153. For an interesting article that argues that juries engage in a process of statutory interpretation much like that used by judges in matching the facts of a case to the law as presented by the judge, see generally Darryl K. Brown, *Plain Meaning, Practical Reason, and Culpability: Toward a Theory of Jury Interpretation of Criminal Statutes*, 96 MICH. L. REV. 1199 (1998).

phenomenon, imagine a laboratory manual that contains the following instructions for mixing an important chemical agent:

1. Carefully pour chemical AAA into a beaker.
2. Then carefully pour twice as much of chemical BBB into the beaker.
3. Finally, test for purity. If the beaker contains more than 300 parts of contaminant, destroy the batch. Otherwise, go to step 4. . . .

The problem with this instruction is step 3. It specifies a purity threshold, but does not say how big the beaker must be. If, for example, we always use a one liter beaker, and always fill it completely (1/3 AAA and 2/3 BBB), then we know that the substance must be pure to the extent of containing fewer than 300 parts of contaminant per liter. But as the instructions are stated, technicians can use any sized beaker from a milliliter to a decaliter, and still pass the purity test if there are fewer than 300 parts of contaminant.

“Proof beyond a reasonable doubt” is subject to the same criticism. We focus our attention on the expression, “reasonable doubt.” But what about “proof”? If I am right about what we do when we doubt, proof beyond reasonable doubt requires the government to tell a story consistent with the facts that the defendant cannot rebut by showing reasonable alternatives. This places much more of the burden on the defendant than the system acknowledges. To continue the metaphor, the standard does not say how large the proof beaker must be before we measure reasonable doubt contaminants in the government’s story. The risk is that cases based on questionable evidence can result in convictions.

Consider the following example: An innocent defendant with a criminal record is accused of holding up a liquor store. At the time of the robbery, he was home with his wife. The defendant does not testify, because if he does his record will come out and he will surely be convicted. His wife does testify, but she is not believed. The prosecution’s case is based on circumstantial evidence, such as the defendant’s car being identified as being near the scene. The jurors are not thoroughly convinced of the defendant’s guilt, but the defendant has given them no reason to doubt the government’s story. I predict that defendants like this one will be convicted much of the time. I also believe that the problem is with the reasonable doubt standard, regardless of how “reasonable doubt” is defined, or whether it is defined at all. The empirical studies discussed in the previous section suggest that I am right.¹⁵⁴

154. Judge Posner would reduce the magnitude of this problem by amending the rules of evidence to exclude impeachment by prior offenses. See FED. R. EVID. 609(a) (allowing evidence of prior convictions to be introduced to impeach a defendant who takes the stand); Richard A. Posner, *An*

Assuming the values discussed in Part II of this Article, the proof beyond a reasonable doubt standard was probably fine in the nineteenth century. It was associated with moral certainty, jury instructions typically contained the expression "moral certainty,"¹⁵⁵ and people knew what moral certainty meant.¹⁵⁶ But we rarely speak of moral certainty now except in legal situations. We certainly cannot count on jurors understanding the term in the Lockean sense when they do hear it in a courtroom.¹⁵⁷

A subtle change in the language of reasonable doubt instructions in the middle of the nineteenth century further supports the point that people understand proof beyond a reasonable doubt as imposing a burden on defendants. In the beginning of the nineteenth century, federal courts generally spoke of a burden of proof "beyond *all* reasonable doubt," "beyond *any* reasonable doubt," or simply "beyond reasonable doubt." "Beyond *a* reasonable doubt" could be seen occasionally, but did not become the norm until the second half of the nineteenth century.¹⁵⁸

When used as a noun, "doubt" refers either to the state of doubting, or to evidence that causes one to be in the state of doubting.¹⁵⁹ Compare the following pair:

I have some doubt about that

I have some doubts about that

The first of these sentences uses "doubt" as a mass noun. Mass nouns, like *water*, generally occur only in the singular, and do not take the article,

Economic Approach to the Law of Evidence, 51 STAN. L. REV. 1477, 1526-27, 1544-45 (1999). This would make it easier for defendants with criminal records to take the stand to tell their stories, and in this sense Judge Posner is right. But his suggestion does not, I believe, go far enough. Posner's idea makes it more plausible for defendants to prove their innocence, but does not address the larger problem, which is that criminal defendants are not supposed to bear the burden of proving their innocence in the first place.

155. The most prominent of these instructions, the *Webster* instruction of Chief Justice Shaw of the Supreme Judicial Court of Massachusetts, *Commonwealth v. Webster*, 59 Mass. (5 Cush.) 295, 320 (1850), is quoted and discussed in Justice O'Connor's majority opinion in *Victor v. Nebraska*, 511 U.S. 1, 8-12 (1994) (recognizing that the phrase "moral certainty" has fallen out of common usage and stating that "beyond a reasonable doubt" is an equivalent and more commonly used phrase).

156. See *Victor*, 511 U.S. at 12-13 (asserting that the phrase "moral certainty," which was generally understood in 1850, might confuse a modern juror, even though, in *Fidelity Mutual Life Ass'n v. Mettler*, 185 U.S. 308, 317 (1902), the Court held that "[p]roof to a 'moral certainty'" was an acceptable substitute for "beyond a reasonable doubt").

157. The Supreme Court has criticized the expression "moral certainty" as a distraction from evidentiary certainty. *Cage v. Louisiana*, 498 U.S. 39, 41 (1990) (per curiam) (holding that the use of "moral certainty" rather than "evidentiary certainty" could lead to a finding of guilt based on a degree of proof below that which is required by the Due Process Clause).

158. See *supra* note 23. This pattern describes federal courts' use of these expressions in published opinions. Not all states followed this pattern. See, e.g., *Dawson v. Thruston*, 12 Va. 132, 138 note, § XII (1808) (using the phrase "proved beyond *a* reasonable doubt" (emphasis added)). These variations should not be surprising since there is no reason to believe that they reflect conscious decisions to use one form over another.

159. I am grateful to Edwin Williams for his analysis of this distinction.

a; rather, they take *some*.¹⁶⁰ *Doubt* in this sense conforms to Justice Scalia's definition, "an uncertain, tentative, or provisional disbelief."¹⁶¹ The second sentence uses *doubts* as a count noun. Count nouns, like *book*, do occur freely in the plural, and take *a* in the singular and *some* in the plural. *Doubts* as a count noun refers to the facts that cause one to be in a state of doubting. It would be appropriate to respond to the second sentence (but not the first) by asking, "What are they?" The expression "proof beyond a reasonable doubt" uses "doubt" as a count noun. Its focus, therefore, is even more on evidence the defendant has raised to create doubt in the juror's mind rather than the strength of the government's evidence.

The most obvious response to all of these arguments is that jurors are instructed in no uncertain terms that the defendant is presumed innocent until proven guilty, and that the entire burden of proof rests on the government. In fact, the presumption of innocence is considered even more basic to notions of fair play than is the proof beyond a reasonable doubt standard.¹⁶² This response is correct as far as it goes. But once again, we should not confuse good intentions with good execution. An instruction that tells the jury that the government bears the burden of proving its case does very little unless it says how strong that proof must be. If the only measure of strength is stated as a function of alternative stories that one can imagine, then the government's burden is very low indeed. This explains why, when asked to quantify the minimum level of certainty needed for a conviction, people respond with much lower thresholds than we believe appropriate.¹⁶³ We would like to think that we do not convict unless the members of the jury are as certain as they can reasonably expect themselves to be of a defendant's guilt. But traditional jury instructions do not require that level of certainty.

2. *Prototype Effects and Mental Models: Playing "The Price Is Right"*.—In a television game show called "The Price Is Right," a panel of participants is presented with a consumer item, such as a washing machine. Each must guess the manufacturer's suggested retail price. This might be

160. We understand "I want a water" to mean I want a glass of water, or a bottle of water. That is, we only use *a* with water when we are speaking elliptically about a count noun, generally some sort of container.

161. *Allentown Mack Sales & Service, Inc. v. NLRB*, 118 S.Ct. 818, 832 (1998). For further discussion, see *supra* note 139.

162. Even those instructions that I criticize here contain perfectly clear statements on the presumption of innocence and the government bearing the burden of proving the defendant's guilt. See *State v. Medina*, 685 A.2d 1242, 1244 (N.J. 1996) (noting in instructions to the jury that "the presumption [of innocence] does not fade or extinguish until 12 of you agree that he is guilty of something[]").

163. See *supra* note 101 for an analysis of quantification studies on burdens of proof.

hard enough, since manufacturers' suggested prices are inflated to allow retailers to appear to be giving discounts when they charge normal market prices. What makes the game even harder is that the winner is defined as the contestant who comes closest to the suggested retail price without exceeding it. If the price of the washing machine is \$398, then \$150 beats \$400.

We do not generally conceptualize in terms of maximums or minimums. Rather, we build models that contain stereotypical information, sometimes called "prototypes" in the psychological literature.¹⁶⁴ Thus, we think of a \$400 washing machine as a washing machine that costs \$400, plus or minus some degree of error that we are willing to tolerate—not a washing machine whose price cannot exceed \$400. Yet, playing *The Price Is Right* is exactly what we want jurors to do. We want jurors to convict the defendant of the most serious charged crime that they are sure the defendant committed. They should not convict a defendant of a crime they are not sure he committed, even if the defendant's conduct looks more like that crime than it does any other choices that the jury has.

To see how this works, it is useful to conduct a thought experiment. Pretend that you are participating in a study in which the experimenter shows you a set of pictures, and you have to label each of them either as a rabbit or a fish. You have no other choices. In addition to showing real rabbits and real fish, the experimenter shows you something that looks a little like a rabbit, but is clearly not a rabbit. Nonetheless, it certainly looks more like a rabbit than a fish. Given the experimental choices, you will no doubt call this a rabbit because it is closer to being a rabbit than a fish.

In everyday life, we make these sorts of choices all the time. If, for example, the red traffic lights in a town you are driving through are a slightly different shade from the ones that you are used to seeing, you will generalize the off-red signals to the red ones that you know so well, and stop your car. Our ability to fit new situations into categories characterized by prototypical situations is not only commonplace but a human strength. It is what allows us to apply our knowledge of the world to unfamiliar problems.

Now let us alter our experiment by adding one new instruction. In addition to telling you that you will be asked to call each of the pictures either a rabbit or a fish, you are told that if you have any reasonable doubts about the picture being a rabbit, then you should call it a fish. Now when you are shown the same picture—the picture of the fictional animal

164. See JOHNSON-LAIRD, *supra* note 142, at 190-95 (noting that "many natural categories are mentally represented by *prototypes*, i.e., schemata of their most characteristic members"); Elenor Rosch, *Cognitive Representations of Semantic Categories*, 104 J. EXPERIMENTAL PSYCHOL.: GEN. 192, 193 (1975) (arguing from experimental evidence that we categorize in terms of prototypes rather than definitions).

that resembles a rabbit but isn't a rabbit—you become uncomfortable. You have in front of you something that looks a lot like a rabbit, but have been told to call it a fish.

The problem has two solutions. The first is to call it a fish. This is harmless enough in our thought experiment. But if your answer really mattered, as it does if you are voting as a member of a jury, you might gradually convince yourself to disregard those features of the picture that made you think it was not a rabbit. Over time, you may become willing enough to call it a rabbit, which certainly seems closer to reality than calling it a fish. After all, you come to the experiment not only with a model of a rabbit, but also with a model of a fish. The temptation to call it a rabbit is increased by the unattractiveness of calling it a fish.

The fact that we assimilate new information into mental models based on its fit with a category's prototype has been exploited in the legal literature. For example, Steven Winter has employed prototypically motivated models of the sort used in everyday conceptualization to explain a host of legal phenomena, from standing doctrine to the well-known "no vehicles in the park" puzzle.¹⁶⁵ A recent article uses Winter's prototype analysis to explain difficulties in police brutality prosecutions.¹⁶⁶ Prototype analysis has also been recently used to deal with issues of statutory interpretation¹⁶⁷ and to explain the law of boycotts.¹⁶⁸

Most pertinent here is the substantial evidence that jurors do this as well. In a very interesting set of studies, psychologist Vicki L. Smith showed that psychology students participating in a mock jury study were more likely to convict when the defendant's acts reflected typical attributes of a crime, even when these attributes did not match the legal definition of the crime as reflected in jury instructions.¹⁶⁹ For example, subjects

165. H.L.A. Hart first described this puzzle in his 1958 Harvard Law Review article *Positivism and the Separation of Law and Morals*. H.L.A. Hart, *Positivism and the Separation of Law and Morals*, 71 HARV. L. REV. 593, 606-07 (1958). For some recent examples, see STEVEN L. WINTER, A CLEARING IN THE FOREST: HOW THE STUDY OF THE MIND CHANGES OUR UNDERSTANDING OF LIFE AND LAW 100-02 (1999) (explaining that a legal rule cannot be understood without the use of "prototypicality, function-based categorization, and normatively infuse judgement"); Steven L. Winter, *The Cognitive Dimension of the Agon Between Legal Power and Narrative Meaning*, 87 MICH. L. REV. 2225, 2256-67 (1989) (discussing the role of narrative, consisting of idealized cognitive models, and its "prototype effect" on lawmaking) and Steven L. Winter, *The Metaphor of Standing and the Problem of Self-Governance*, 40 STAN. L. REV. 1371 (1988) (analyzing the negative effects of the role of metaphor as related to the doctrine of standing).

166. See David Dante Trout, *Screws, Koon and Routine Aberrations: The Use of Fictional Narratives in Federal Police Brutality Prosecutions*, 74 N.Y.U. L. REV. 18, 85-117 (1999).

167. See Lawrence M. Solan, *Law, Language and Lenity*, 40 WM. & MARY L. REV. 57 (1998) (employing in part linguistic prototype theory to explain problems in criminal statutory interpretation).

168. See GARY MINDA, *BOYCOTT IN AMERICA: HOW IMAGINATION AND IDEOLOGY SHAPE THE LEGAL MIND* 17-75 (1999) (discussing opposing metaphors of boycotts and the implications of each cognitive metaphor).

169. See Vicki L. Smith, *When Prior Knowledge and Law Collide: Helping Jurors Use the Law*, 17 L. & HUM. BEHAV. 507, 509 (1993) [hereinafter Smith, *Prior Knowledge*] ("[V]erdict decisions are

typically identified a burglary as a crime in which something of value is taken from a home or apartment after a break-in with a purpose to steal. In contrast, the legal definition requires that the defendant, "without authority, knowingly enters a building with intent to commit a felony therein."¹⁷⁰ Smith found it harder to get convictions when some of the jurors' preconceived attributes were missing, and easy to get convictions when those attributes were present even though some of the legal elements were absent. This suggests that jurors will tend to associate facts with legal categories based on the closest conceptual match to mental models of the crime that they already possess,¹⁷¹ as opposed to the jury considering whether the government has proven each element of the crime beyond a reasonable doubt. The result, if Smith's experiments reflect reality, is that it should be difficult for the government to convict a guilty defendant of a crime that is remote from jurors' mental model of the offense with which the defendant is charged. At the same time, it should be relatively easy for the government to convict an innocent defendant when the defendant's conduct, while not meeting all the elements of the crime, comes close to the jurors' model.

This, I believe, is why, in a celebrated case, the English *au pair* Louise Woodward was convicted of second degree murder for having caused the death of an infant.¹⁷² Although the government's case was not compelling, jurors believed she had done something wrong, and her conduct came closer to their model of second degree murder than it did to their model of innocence. The most significant legal event in the case was defense counsel's decision not to allow the jury to be charged on involuntary manslaughter.¹⁷³ Not recognizing the difficulties the jury would have with the Price Is Right effect, defense counsel assumed that no jury could possibly convict Woodward of second degree murder based on the government's seriously flawed case. After the conviction, the trial judge relieved defense counsel of the consequence of the gamble by reducing the

influenced by people's prototypes of crime categories when they should be based on a set of specific necessary and sufficient conditions."); Vicki L. Smith, *Prototypes in the Courtroom: Lay Representations of Legal Concepts*, 61 J. PERSONALITY & SOC. PSYCHOL. 857, 866 (1991) [hereinafter Smith, *Prototypes*] (concluding that many jurors are influenced by their own erroneous concepts of what the prototype of the charged crime should be).

170. This is the definition that Smith uses, see Smith, *Prior Knowledge*, *supra* note 169, at 512, which is taken from the Illinois Pattern Jury Instructions. See ILLINOIS SUPREME COURT COMM. ON PATTERN JURY INSTRUCTIONS IN CRIMINAL CASES, ILLINOIS PATTERN JURY INSTRUCTIONS—CRIMINAL §§ 14.05-.06 (2d ed. 1981).

171. Hastie et al., *supra* note 148, at 22, suggest that this is just what jurors do. They also find, quite disturbingly, that during deliberations jurors spend very little time discussing burden of proof issues. See *id.* at 86-87.

172. See *Commonwealth v. Woodward*, 694 N.E.2d 1277, 1277 (Mass. 1998).

173. See *id.* at 1281.

verdict to involuntary manslaughter anyway. In a 4-3 decision, the Supreme Judicial Court of Massachusetts affirmed.¹⁷⁴

Moreover, jurors bring to the courthouse more than their prototypes of individual crimes. They also have prototypes for an innocent defendant. Films, mystery novels, and television programs frequently show innocent defendants on trial, only to be acquitted when a hero discovers who really committed the crime, sometimes getting him to confess on the witness stand. Real trials, in contrast, are often about the relative weight that should be given to evidence that suggests guilt and evidence that suggests innocence. Trials end without the defendant looking like the prototypical innocent defendant.¹⁷⁵

Prototype effects also explain what is wrong with approaches that suggest that reasonable doubt be left unexplained, since the words are simple and jurors (supposedly) know what they mean. Jurors know what the expression means from folk knowledge of the judicial system. The expression is not used outside of that context with any frequency. A review of many instances in which the expression is used in the popular press shows that it is typically used without any explanation of what the standard means.¹⁷⁶ Sometimes, the expression is used in a way that implies that the defense has the burden to create reasonable doubt by poking holes in the government's case.¹⁷⁷ On occasion, prosecutors are quoted defining reasonable doubt in a way that is unacceptable to the

174. *See id.* at 1298.

175. As Carlos Gonzalez has brought to my attention, there is a tension between this fact and the fact that jurors respond differently to various instructions. If these prototype-based models were the whole story about conceptualization, then it really should make no difference what instruction is used. Jurors will simply vote consistent with the model that is the best match. Much recent work in the psychology of conceptualization, however, suggests that prototypes are only part of how we conceptualize, and should not be seen as anything like a complete explanation. *See FODOR, supra* note 141. Yet there is also virtually complete consensus that prototype effects are part of our psychology. For a recent description of current thinking in cognitive psychology, see Edward E. Smith, *Concepts and Categories*, in 3 AN INVITATION TO COGNITIVE SCIENCE: THINKING (Edward E. Smith & Daniel N. Osherson eds., 2d ed. 1995). For my views on this issue, see Solan, *supra* note 167, at 65-75. If, as I suggest here, prototype effects reflect only part of how we conceptualize, then one might expect prototype effects to play some role in jurors' thinking, but not to trump the instructions entirely. And that seems to be what happens. This makes it all the more important to be concerned about what instructions the jury actually receives.

176. Typical is the following sort of statement: "They need to prove beyond a reasonable doubt who committed this crime and then seek a death sentence for the killer." Editorial, *When a Time to Kill Has Come*, THE INDIANAPOLIS NEWS, May 28, 1998, at A12, available in 1998 WL 8331621.

177. "What the Nichols defense tried to do, really, is to get jurors asking questions, which might lead them to doubts about the government's evidence, which in turn might lead to reasonable doubt, which might lead to a hung jury or an acquittal." Andrew Cohen, *Defense Stressed Doubts: But It May Be Enough to Get Jury to Acquit Nichols*, DENVER POST, Dec. 12, 1997, at A29, available in 1999 WL 13886132; John Dickerson, *Four for Four*, TIME, Mar. 14, 1994, at 33 (explaining that in the World Trade Center bombing trial, "defense attorneys cooperated in raising doubts about each part of the prosecution's reconstruction in hopes of raising reasonable doubt about the overall story").

courts.¹⁷⁸ Often the expression appears in the press with the implication that the government has a difficult burden to meet, but without much more.¹⁷⁹

Thus, while jurors do have a prior model of what reasonable doubt means, they necessarily built that model the way we all build our own models: from newspapers, magazines, television, books, and so on. But most of these sources contain incomplete or inaccurate reports of what courts demand of the government. Jurors' prototypes for reasonable doubt may vary significantly from what we really want of our criminal justice system, just as their prototypes for what it means to have committed various crimes differ from legal definitions, as Smith demonstrates.¹⁸⁰ For this reason, I do not agree with those jurisdictions that are satisfied to instruct on reasonable doubt and leave the term to the jurors' prior understanding.

3. *Gricean Implicature and the Pressure to Convict.*—In a short article that has generated enormous literature, the philosopher H. Paul Grice wrote that we routinely and unselfconsciously rely on a standard set of inferences in understanding speech and, when speaking, assume that the hearer will rely on this same set of inferences.¹⁸¹ The core inference is Grice's cooperative principle: "Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged."¹⁸² Jurors hearing instructions, like the rest of us, apply this principle. They assume that the judge has a purpose in presenting the instructions in the form she does.

Some of the definitions of reasonable doubt discussed earlier show how Gricean implicatures can affect the ways in which jurors understand their jobs. Many instructions tell the jury not to consider "speculative" or "imaginary" doubts.¹⁸³ Such an instruction, while perhaps well-

178. "William Wise, assistant Boulder district attorney, recently likened a 'probable cause' finding to a 51 percent certainty, while a 'reasonable doubt' threshold would be at least 75 percent." Guy Kelly & Kevin McCullen, *Risk Seen In Grand Jury Probe of JonBenet Case*, CHI. SUN-TIMES, Mar. 15, 1998, available in 1998 WL 5572184.

179. "[Acting New York Supreme Court Justice] Tejada said the circumstantial case must be proved beyond a reasonable doubt and 'to a moral certainty.'" Barbara Ross, *Jury Finds 3 Cops Not Guilty in Wash. Heights Burglaries*, N.Y. DAILY NEWS, Nov. 4, 1997, available in 1997 WL 16051189.

180. See *supra* note 169.

181. See H. P. Grice, *Logic and Conversation*, in 3 SYNTAX AND SEMANTICS 41 (Peter Cole & Jerry L. Morgan eds., 1975).

182. *Id.* at 45.

183. See *Walker v. State*, 369 So.2d 814, 823-24 (Ala. Crim. App. 1978), *rev'd on other grounds*, 369 So.2d 825 (Ala. 1979); *State v. Medina*, 685 A.2d 1242, 1244-45 (N.J. 1996) (discussed *infra* note 187).

intentioned, is tantamount to telling the jury not to consider any doubts at all. Doubting, after all, *is* a matter of speculation and imagination. It requires one to imagine alternative models consistent with the evidence. A juror hearing such an instruction is likely to draw the inference that the judge is trying to tell her not to take her doubts too seriously unless they are extremely strong. Otherwise, the two instructions together would make no sense, and the judge would be violating the cooperative principle.

Of course, this particular problem can be cured. Jury instructions that tell jurors not to speculate or imagine can and should be eliminated by trial judges and jury reform panels, or, if necessary, struck down by appellate courts.¹⁸⁴ Jurors might still be told that their ability to form outlandish or very improbable alternative models does not constitute “reasonable” doubt. Such an instruction would be a far cry from the jury instructions that I criticize here.

The same problem exists for the instruction that a reasonable doubt “is proof of such a convincing character that you would be willing to rely and act upon it without hesitation in the most important of your own affairs.”¹⁸⁵ If typical jurors are anything like me, they always hesitate before acting in the most important aspects of life. But the judge could not possibly mean that jurors should never convict because they always hesitate to act when they make important decisions in their own lives. Thus, a reasonable juror might infer that the judge must mean that one should not convict if, after deciding that the defendant is guilty, one still hesitates. The FJC recognizes this problem with “hesitate to act” instructions in the commentary to its “firmly convinced” instruction.¹⁸⁶

A related problem comes from instructions that contain many warnings for jurors not to acquit unless the doubts are very serious ones.¹⁸⁷ Some contain several statements in which the court tells jurors why not to take a particular sort of doubt seriously. One of Grice’s principles is the maxim of quantity: An agent will do as much as is required for the achievement

184. Of course, courts that use an undefined reasonable doubt instruction do not have any of the problems that result from unwanted inferences from the definitions themselves.

185. *United States v. Williams*, 20 F.3d 125, 129 n.2 (5th Cir. 1994). The *Williams* court approved the use of Instruction 1.06 of the Fifth Circuit’s pattern jury instructions. *See id.* at 19 n.2; *see also* DISTRICT JUDGES ASSOCIATION, FIFTH CIRCUIT, PATTERN JURY INSTRUCTIONS (CRIMINAL CASES) 16 (1990).

186. 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.

FEDERAL JUDICIAL CTR., *supra* note 4, No. 21, at 29 (1998).

187. *See, for example, the instruction reviewed by the Supreme Court of New Jersey in State v. Medina*, 685 A.2d 1242, 1244-45 (N.J. 1996).

of the current goal; an agent will not do more than is required.¹⁸⁸ The weight of the instruction conveys a message to the jurors: The judge would not have presented so many ways in which the juror's doubts can be used improperly if this were not the main problem to avoid. Such a message is likely to focus jurors on the strength of the defendant's case as a criterion for acquittal rather than on whether the government has proven its case with near certitude.

4. *Why the Firmly Convinced Standard Should Work Better.*—I have argued that the expression “proof beyond a reasonable doubt” tends to shift the burden of proof to the defendant, leaves the government's burden unquantified, and is subject to prototype effects that can help the government win cases based on weak, but uncontested, evidence. When the proof shows that a defendant is clearly guilty, none of this matters—there will generally be a conviction no matter what the standard.¹⁸⁹ It is when the government's case is less strong that we need to be concerned about these issues.

Whether stated as a definition of proof beyond a reasonable doubt, or simply as its own standard, expressions like “firmly convinced,” “satisfied that you are sure,” and “thoroughly convinced” eliminate some of the problems and reduce others. What does it mean to convince someone thoroughly? It means that the speaker has caused the hearer to construct a mental model that is powerful, and that no alternative models are suitable alternatives. In other words, it means exactly what the government should have to do to get a conviction. It must put on a case in which the evidence favoring its position is strong and in which there are no other reasonable alternatives. Moreover, the words themselves focus on what the government must do—not on what the defendant has been able to accomplish. It tells jurors not to accept the government's model unless they are convinced that it is the right one.¹⁹⁰ The FJC and New Jersey instructions tell jurors to keep both sides of the equation in mind, whereas standard reasonable doubt instructions do not.

As for prototype effects, expressions like “firmly convinced” will still lead to assimilation of the closest match between the facts and a legal category, as opposed to matches in which all of the elements are proven.

188. See Grice, *supra* note 181, at 45-46. For discussion of this maxim and of Grice's work generally, see GEORGIA M. GREEN, PRAGMATICS AND NATURAL LANGUAGE UNDERSTANDING 89-129 (2d ed. 1996).

189. See generally L.S.E. Jury Project, *supra* note 82, at 219 (noting that the nature of a jury instruction only takes on significance in borderline cases).

190. “Convince” is what Austin called a “perlocutionary” verb. Its meaning includes the notion that an utterance has effected a change in the state of mind of the hearer. See J.L. AUSTIN, HOW TO DO THINGS WITH WORDS 101-02 (2d ed. 1975).

But by focusing the jury on the government's obligation there should be some improvement in this regard as well. Moreover, the Gricean problems endemic in instructions that specify which doubts are reasonable ones are absent in both the FJC instruction and New Jersey's modified firmly convinced instruction in *Medina*.¹⁹¹ I do not claim that the firmly convinced standard is preferable because it is a "good" definition of proof beyond a reasonable doubt. It is not a definition in the sense that it uses simpler words to convey the meaning of a more complex concept. If it is a definition at all, it is a stipulated definition. The superiority of the firmly convinced instruction comes not from its semantic fidelity to the reasonable doubt standard but from its greater success in promoting important values. To lose sight of this point is to accept a formulaic approach to our system of criminal justice rather than an approach that recognizes the values that led to the adoption of the reasonable doubt standard in the first place.

Nor do I argue here that courts must eliminate the words "reasonable doubt" entirely from jury instructions, as some English courts¹⁹² and the French Code of Criminal Procedure have done.¹⁹³ The empirical studies using the FJC instruction, which contains the expression "proof beyond a reasonable doubt," show that it reduces the rate of conviction when the evidence is highly equivocal. My argument has been that the standard instructions misfocus the jury, not that doubting is irrelevant to whether the government has met its burden of convincing the jury sufficiently. The concept of reasonable doubt does highlight part of what a jury should consider in its deliberations. Without further empirical evidence, I have no reason to conclude at this point that instructions currently used in England and France are superior to ones that I espouse here. Other possible instructions, such as "proof so convincing that it leaves no reasonable doubt of the defendant's guilt,"¹⁹⁴ may also accomplish the same goals. In this regard, future experimental research directed at these subtle differences may be useful, even if only to show that at some point, once the appropriate concepts are conveyed, the precise language begins to matter less.

A related problem is the status of the intermediate "clear and convincing evidence" standard. Two of the empirical studies discussed in Part III

191. Again, these problems are also absent from the approach used by the Fourth and Seventh Circuits, in which jurors are told that they must find guilt beyond a reasonable doubt, but are not given definitions of that term. See *supra* notes 42-46 and accompanying text.

192. See *Walters v. Quenn* [1969] 2 App. Cas. 26, 30 (P.C. 1969) (appeal taken from Jam.) (requiring a jury to feel "quite sure" that the prosecution has established the guilt of the accused).

193. See FRENCH CODE OF CRIMINAL PROCEDURE art. 353; *supra* text accompanying note 59; see also a proposal by Judge Fortunato to this effect, *supra* note 52.

194. This language is similar to that found in *Peter v. Beverly*, 35 U.S. 532, 568 (1836) ("[B]ut the evidence must certainly be so clear and satisfactory, as to leave no reasonable doubt . . .").

suggest that it is harder to meet this standard than it is to meet the proof beyond a reasonable doubt standard.¹⁹⁵ The linguistic analysis presented in this Part predicts that this should be the case. Here, too, further empirical work would be useful. It appears at this point, however, that our system is very likely mistaken about the relative burdens that it imposes on parties in different circumstances. Fortunately, the empirical research also suggests that changing the burden of proof instructions in criminal cases along the lines suggested here should correct this problem.

Finally, the approach to burden of proof in criminal cases espoused here should make it more difficult for defense counsel to argue that the burden has not been met when the case against the defendant is strong. Trial advocacy texts traditionally instruct lawyers to focus their summations on the concept of reasonable doubt.¹⁹⁶ The empirical literature has not tested the success of this strategy, perhaps because experimental studies focus on juror reactions to equivocal evidence and reasonable doubt arguments occur during defense counsel's summation. Nonetheless, just as the reasonable doubt standard seriously disadvantages the innocent defendant who cannot meet the burden of raising doubts, it may also aid the wealthy defendant whose counsel is able to raise remote theories of the case that are entirely unconvincing but create some small level of doubt. The "firmly convinced" standard should make this more difficult to accomplish. Particular "doubts" that the defense has raised would be properly integrated into the jury's analysis of how convincing the case against the defendant actually is. I advance this conclusion tentatively, since it is not supported by empirical research. It would not be surprising, however, to find that both parties in criminal prosecutions take advantage of the distortions that the proof beyond a reasonable doubt standard imposes on the system.

V. Conclusion

The burden that we impose on the government in criminal cases reflects our society's abhorrence of the conviction of the innocent. The phrase typically used to ensure that there are very few such convictions,

195. See L.S.E. Jury Project, *supra* note 82; Kagehiro & Stanton, *supra* note 101; *supra* notes 121-24 and accompanying text.

196. See, e.g., ALFRED S. JULIEN, JULIEN ON SUMMATION 82-83 (1986) ("The defense has so many imaginative fundamentals working for it, such as 'My opponent must prove the prosecution's case beyond a reasonable doubt.' The emphasis should be on the words *beyond* and *doubt*." (emphasis in original)); RICHARD H. LUCAS & K. BYRON MCCOY, THE WINNING EDGE: EFFECTIVE COMMUNICATION AND PERSUASION TECHNIQUES FOR LAWYERS 124-25 (1993) (suggesting the use of analogies to help the jury better understand reasonable doubt); HENRY B. ROTHBLATT, SUCCESSFUL TECHNIQUES IN THE TRIAL OF CRIMINAL CASES § 6.6 (1961) ("Your entire summation is based upon one premise: the failure of the prosecutor to establish the guilt of the defendant beyond a reasonable doubt. Your argument has been pointed toward raising that doubt in the minds of the jurors. Therefore, your summation must finish on that note.").

“proof beyond a reasonable doubt,” came into widespread use in the nineteenth century. However, these may not be the best words to accomplish this important jurisprudential goal, at least not unless they are explained in terms of the government’s burden to present convincing evidence. The biggest problem is that the phrase focuses too much on the defendant’s ability to create alternative mental models and not enough on the government’s obligation to prove its case. Linguistic analysis of the language supports this proposition, as does a close examination of the experimental literature.

Most courts in this country continue to seek definitions of which doubts are properly considered reasonable ones. I recommend that the Federal Judicial Center’s “firmly convinced” standard, as modified by the Supreme Court of New Jersey,¹⁹⁷ be adopted by those courts that have not yet done so. While “firmly convinced” is not really a definition of “beyond a reasonable doubt,” it best reflects the idea that defendants should not be convicted unless the government has proven guilt to near certitude. Surprisingly few courts have adopted this approach, even though the FJC’s instructions were first published more than fifteen years ago. Such a change may well be needed to correct a potential imbalance in our system: Empirical research suggests that it is easier to prove a case beyond a reasonable doubt than it is to prove a case by clear and convincing evidence. Empirical research also suggests that this imbalance disappears when an instruction like the FJC’s is used in criminal cases.

It is time for both the states and the federal government to eliminate unfairness in their criminal justice systems by reforming jury instructions to reflect the deeply held values that we claim to cherish. France and England have taken such steps. So should the United States.

197. See *supra* text accompanying note 54.

