



Case Western Reserve Law Review

Volume 66 | Issue 2

2015

Trial by Traditional Probability, Relative Plausibility, or Belief Function?

Kevin M. Clermont

Follow this and additional works at: <https://scholarlycommons.law.case.edu/caselrev>

 Part of the [Law Commons](#)

Recommended Citation

Kevin M. Clermont, *Trial by Traditional Probability, Relative Plausibility, or Belief Function?*, 66 Case W. Res. L. Rev. 353 (2015)
Available at: <https://scholarlycommons.law.case.edu/caselrev/vol66/iss2/5>

This Article is brought to you for free and open access by the Student Journals at Case Western Reserve University School of Law Scholarly Commons. It has been accepted for inclusion in Case Western Reserve Law Review by an authorized administrator of Case Western Reserve University School of Law Scholarly Commons.

TRIAL BY
TRADITIONAL PROBABILITY,
RELATIVE PLAUSIBILITY,
OR BELIEF FUNCTION?

Kevin M. Clermont[†]

ABSTRACT

Almost incredible is that no one has ever formulated an adequate model for applying the standard of proof. What does the law call for? The usual formulation is that the factfinder must roughly test the finding on a scale of likelihood. So, the finding in a civil case must at least be more likely than not or, for the theoretically adventuresome, more than fifty percent probable. Yet everyone concedes that this formulation captures neither how human factfinders actually work nor, more surprisingly, how theory tells us that factfinders should work.

An emerging notion that the factfinder should compare the plaintiff's story to the defendant's story might be a step forward, but this relative plausibility conjecture has its problems. I contend instead that the mathematical theory of belief functions provides an alternative without those problems, and that the law in fact conforms to this theory. Under it, the standards of proof reveal themselves as instructions for the factfinder to compare the affirmative belief in the finding to any belief in its contradiction, but only after setting aside the range of belief that imperfect evidence leaves uncommitted. Accordingly, rather than requiring a civil case's elements to exceed fifty percent or comparing best stories, belief functions focus on whether the perhaps smallish imprecise belief exceeds its smallish imprecise contradiction. Belief functions extend easily to the other standards of proof. Moreover, belief functions nicely clarify the workings of burdens of persuasion and production.

CONTENTS

INTRODUCTION	354
I. PROOF BY FACTFINDERS' BELIEFS	360
<i>A. Belief Functions</i>	362

[†] Ziff Professor of Law, Cornell University. My sincere thanks to Zach Clopton and Mike Pardo for thoughtful and helpful reactions. Portions of this article are reprinted from Kevin M. Clermont, *Death of Paradox: The Killer Logic Beneath the Standards of Proof*, 88 NOTRE DAME L. REV. 1061 (2013), and from KEVIN M. CLERMONT, STANDARDS OF DECISION IN LAW: PSYCHOLOGICAL AND LOGICAL BASES FOR THE STANDARD OF PROOF, HERE AND ABROAD (2013).

1. Basics of Theory.....	364
2. Negation Operator	367
3. Lack of Proof	368
B. <i>Comparison of Beliefs</i>	369
II. BURDEN OF PERSUASION.....	369
A. <i>Traditional View</i>	370
B. <i>Reformulated View</i>	372
1. Preponderance of the Evidence	373
2. Clear and Convincing Evidence	375
3. Beyond a Reasonable Doubt	376
III. BURDEN OF PRODUCTION	377
A. <i>Traditional View</i>	377
B. <i>Reformulated View</i>	380
IV. OVERVIEW OF STANDARDS OF PROOF.....	382
A. <i>Compatibility of Reformulated and Current Standards</i>	382
B. <i>Application of Reformulated Standards to Multiple Elements</i>	385
CONCLUSION.....	391

INTRODUCTION

The different standards of proof determine outcome. Empirical proof supports that point, as long as the standard applied to the empirical proof itself is not too demanding.¹ In any event, standards are definitely worth worrying about. A firmer understanding would affect the resolution of many legal issues that arise in connections with standards and burdens of proof. Almost incredible, however, is that no one has yet formulated an adequate model of proof-standard application. What does the law call for the factfinder to do?

The standard of proof often finds expression in terms drawn from *traditional probability theory*.² The formulation would be something along the lines that the factfinder must test whether the finding meets or exceeds the required standard on a scale of likelihood, albeit merely a nonnumerical scale with coarse gradations such as: (1) slightest possibility, (2) reasonable possibility, (3) substantial possibility, (4)

-
1. See KEVIN M. CLERMONT, STANDARDS OF DECISION IN LAW: PSYCHOLOGICAL AND LOGICAL BASES FOR THE STANDARD OF PROOF, HERE AND ABROAD 103–16 (2013) (discussing empirical studies). I draw part of my argument from that book, which more fully documents the subject and broadens its reach considerably. In this Article, I try to marshal and synthesize specifically the arguments in favor of utilizing belief functions to produce a single and sound image for the standards and burdens of proof.
 2. See *id.* at 119–21 (explaining the traditional view of probability theories).

equipoise, (5) probability, (6) high probability, and (7) almost certainty. Nonetheless, a yearning for more precision pushes many armchair theorists to use numbers in describing the scale.

Take as a prime example the usual standard of proof in civil cases, which calls for a probability of more likely than not. “As every first-year law student knows, the civil preponderance-of-the-evidence standard requires that a plaintiff establish the probability of her claim to greater than 0.5.”³ A moment’s reflection, however, reveals all sorts of problems with such a formulation of proof. First, there are the routine objections to speaking of proof in numerical terms.⁴ Not only are percentages of likelihood not how people normally think about legal cases, but also use of numbers can mislead the factfinder.⁵ As soon as the theorist thinks more deeply about the nature of proof, those numbers produce all kinds of paradoxes.⁶ Second, the civil standard seems impossibly difficult:

If the plaintiff must prove that some fact, *X*, is more probable than its negation, *not-X*, then the plaintiff should have to show not only the probability that the state of the world is such that *X* is true, but also the probability of every other possible state of the world in which *X* is not true. This would mean that in order to prevail, plaintiffs would have to disprove (or demonstrate the low likelihood of) each of the virtually limitless number of ways the world could have been at the relevant time. This would be a virtually impossible task, and thus, absent conclusive proof, plaintiffs would lose.⁷

-
3. Edward K. Cheng, *Reconceptualizing the Burden of Proof*, 122 YALE L.J. 1254, 1256 (2013) (footnote omitted).
 4. See CLERMONT, *supra* note 1, at 75–78, 113–14 (providing potential problems from quantifying decisionmaking standards).
 5. *Id.*
 6. See Kevin M. Clermont, *Aggregation of Probabilities and Illogic*, 47 GA. L. REV. 165 (2012) (discussing logical problems surrounding aggregation of probabilities); see also Kevin M. Clermont, *Conjunction of Evidence and Multivalent Logic*, in LAW AND THE NEW LOGICS (Lionel Smith ed., forthcoming 2016), available at <http://ssrn.com/abstract=2472383> [<http://perma.cc/A886-47PZ>] (highlighting paradoxes in different standards of proof). The difficulties increase when one tries to account for multiple factfinders, as in a jury. See RICHARD H. FIELD, BENJAMIN KAPLAN, KEVIN M. CLERMONT & CATHERINE T. STRUVE, MATERIALS FOR A BASIC COURSE IN CIVIL PROCEDURE 1387–89 (11th ed. 2014) (resolving problems of which facts the jurors must agree on and of whether different jurors on a nonunanimous jury can provide support for different issues).
 7. Michael S. Pardo, *Second-Order Proof Rules*, 61 FLA. L. REV. 1083, 1093 (2009) (footnote omitted).

But in recognition of inevitably imperfect evidence, the law allows recovery upon much less than a fifty percent showing of probability.⁸ Third, the civil standard simultaneously seems an impossibly easy one. The factfinder supposedly starts in a state of perfect ignorance, wherein the plaintiff's claim has a fifty-fifty chance by the indifference principle. So, introduction of a feather's weight of evidence should suffice for victory over a silent defendant. But we all know that in such a case, the plaintiff would lose by directed verdict. A feather's weight might swing the burden of persuasion, but it does not satisfy the burden of production. The reality is that the law requires much stronger evidence.⁹

Consequently, it is abundantly clear that academics need to "let go of their love for $p > 0.5$."¹⁰ Among the various proffered alternatives,¹¹ the most frequently ballyhooed way to let go of the love is the *relative plausibility theory*.¹² It builds on psychology's story model of holistic

-
8. See Larry Laudan, *Strange Bedfellows: Inference to the Best Explanation and the Criminal Standard of Proof*, 11 INT'L J. EVIDENCE & PROOF 292, 304-05 (2007) ("The trier of fact cannot say, 'Although plaintiff's case is stronger than defendant's, I will reach no verdict since neither party has a frightfully good story to tell'. Under current rules, if the plaintiff has a better story than the defendant, he must win the suit, even when his theory of the case fails to satisfy the strictures required to qualify his theory as the best explanation.").
 9. See 2 MCCORMICK ON EVIDENCE § 338, at 654 (7th ed. 2013) ("A 'scintilla' of evidence will not suffice.").
 10. Cheng, *supra* note 3, at 1258.
 11. E.g., ALEX STEIN, FOUNDATIONS OF EVIDENCE LAW 133-40 (2005) (arguing for viewing standards of proof in terms of allocation of risk); Christoph Engel, *Preponderance of the Evidence Versus Intime Conviction: A Behavioral Perspective on a Conflict Between American and Continental European Law*, 33 VT. L. REV. 435 (2009) (arguing for a standard based on psychological confidence); Leonard R. Jaffee, *Of Probativity and Probability: Statistics, Scientific Evidence, and the Calculus of Chance at Trial*, 46 U. PITT. L. REV. 925 (1985) (rejecting use of probability and arguing that the burdened party must establish truth); Luke Meier, *Probability, Confidence, and the "Reasonable Jury" Standard*, 84 MISS. L.J. 747 (2015) (arguing for a standard based on statistical confidence); Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357 (1985) (arguing that the process of proof aims at generating acceptable statements about past events and thus at projecting behavioral norms to the public).
 12. See, e.g., Ronald J. Allen & Sarah A. Jehl, *Burdens of Persuasion in Civil Cases: Algorithms v. Explanations*, 2003 MICH. ST. L. REV. 893, 929-43 (explaining the origin of the relative plausibility theory). The *weight of the evidence* methodology in science is a similar approach, as is the *differential diagnosis* approach in medicine that diagnoses by successively eliminating plausible causes of a medical condition to reveal the best explanation. See *Milward v. Acuity Specialty Prods. Grp., Inc.*, 639 F.3d 11, 18 (1st Cir. 2011) (admitting expert evidence based on the weight of the evidence approach: "The scientist must (1) identify an association between an exposure and a disease, (2) consider a range of plausible explanations for the association, (3) rank the

evidence-processing.¹³ The relative plausibility theory posits that the factfinder constructs the overall story (or stories, in some variants of the theory) that the plaintiff is spinning and another story (or stories) that the defendant is (or could be) spinning.¹⁴ The factfinder then compares the two stories (or collections of stories) and gives victory to the plaintiff if the plaintiff's version is more plausible than the defendant's.¹⁵

rival explanations according to their plausibility, (4) seek additional evidence to separate the more plausible from the less plausible explanations, (5) consider all of the relevant available evidence, and (6) integrate the evidence using professional judgment to come to a conclusion about the best explanation.”); *Westberry v. Gislaved Gummi AB*, 178 F.3d 257, 262–63 (4th Cir. 1999) (admitting expert evidence based on differential diagnosis). These methods involve consideration and analysis of alternative explanations to get the one that best explains the evidence, a mode of reasoning called *inference to the best explanation*. Professor Allen is drifting in his thinking in this direction. See Ronald J. Allen & Alex Stein, *Evidence, Probability, and the Burden of Proof*, 55 ARIZ. L. REV. 557, 567–70 (2013) (discussing adjudicative factfinding as inference to the best explanation); Michael S. Pardo & Ronald J. Allen, *Juridical Proof and the Best Explanation*, 27 LAW & PHIL. 223 (2008) (discussing how inference to the best explanation explains judicial proof). However, Laudan, *supra* note 8, powerfully demonstrates that inference to the best explanation holds little additional promise of explaining or illuminating standards of proof. *But cf.* Ronald J. Allen & Michael S. Pardo, *Probability, Explanation and Interference: A Reply*, 11 INT’L J. EVIDENCE & PROOF 307, 314–17 (2007) (defending their best explanation approach in a way that pares it back into a form consistent with a relative plausibility approach).

13. See generally JEFFREY T. FREDERICK, *THE PSYCHOLOGY OF THE AMERICAN JURY* 296–99 (1987) (providing a brief overview of the story model of evidence-processing); Reid Hastie & Nancy Pennington, *The Psychology of Juror and Jury Decision Making*, in REID HASTIE, STEVEN D. PENROD & NANCY PENNINGTON, *INSIDE THE JURY* 22–23 (1983) (providing a brief summary of empirical studies supporting the story model); Paula L. Hannaford, Valerie P. Hans, Nicole L. Mott & G. Thomas Munsterman, *The Timing of Opinion Formation by Jurors in Civil Cases: An Empirical Examination*, 67 TENN. L. REV. 627, 629–33 (2000) (discussing three predominant models of jury decisionmaking); Reid Hastie, *What’s the Story? Explanations and Narratives in Civil Jury Decisions*, in CIVIL JURIES AND CIVIL JUSTICE 23, 31–32 (Brian H. Bornstein et al. eds., 2008) (expanding the theory to allow for a party’s multiple stories); Jill E. Huntley & Mark Costanzo, *Sexual Harassment Stories: Testing a Story-Mediated Model of Juror Decision-Making in Civil Litigation*, 27 LAW & HUM. BEHAV. 29, 29 (2003) (presenting research that “extends the story model to civil litigation and tests a story-mediated model against an unmediated model of jury decision-making”); Nancy Pennington & Reid Hastie, *The Story Model for Juror Decision Making*, in *INSIDE THE JUROR: THE PSYCHOLOGY OF JUROR DECISION MAKING* 192 (Reid Hastie ed., 1993) (detailing the story model and summarizing empirical studies testing it); *cf.* Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511, 559–69 (2004) (arguing that factfinders consider evidence holistically rather than atomistically).

14. Allen & Jehl, *supra* note 12, at 937–38.

15. *Id.*

This choice between alternative competing narratives is largely an ordinal process rather than a cardinal one.¹⁶ Relative plausibility has advantages besides drawing on the currently prevailing psychological literature. It shows a nontraditional embrace of relative judgment by the factfinder, in preference to humans' weaker skills at absolute judgment of likelihood.¹⁷ Also, by inventing a test to apply only at the end of a trial, it sidesteps many of the difficulties and paradoxes of using a numerical standard like greater than fifty percent.¹⁸

Yet, even as most of its proponents admit, relative plausibility theory has its own problems.¹⁹ First, an ordinal comparison cannot easily explain standards of proof higher or lower than preponderance of the evidence.²⁰ Standards from a reasonable suspicion up to evidence beyond a reasonable doubt are hard to express as a comparison of stories.²¹ Second, a more obvious difficulty is that it does not track well what the law tells its factfinders about how to proceed.²² The law says

16. *Id.*

17. Relative judgment concerns the considerable capacity of people to distinguish between two or more different stimuli that they can compare directly. Although not entirely distinct, absolute judgment instead involves reference to a remembered scale. See WILLIAM N. DEMBER & JOEL S. WARM, *PSYCHOLOGY OF PERCEPTION* 113, 116–17 (2d ed. 1979) (explaining the difference between absolute and relative judgment).

18. See Ronald J. Allen & Brian Leiter, *Naturalized Epistemology and the Law of Evidence*, 87 VA. L. REV. 1491, 1542–46 (2001) (discussing probabilistic evidence).

19. See Richard D. Friedman, *“E” Is for Eclectic: Multiple Perspectives on Evidence*, 87 VA. L. REV. 2029, 2046–47 (2001) (noting problems with relative plausibility theory).

20. See Ronald J. Allen, *The Nature of Juridical Proof*, 13 CARDOZO L. REV. 373, 413 (1991) (attempting to explain the beyond-a-reasonable-doubt standard as not being satisfied if the factfinder “concludes that there is a plausible scenario consistent with innocence,” while admitting that the clear-and-convincing standard is “troublesome” under his theory because it seems cardinal); Allen & Leiter, *supra* note 18, at 1528 (saying that the prosecution must “show that there is no plausible account of innocence”); Michael S. Pardo, *Group Agency and Legal Proof; or, Why the Jury Is an “It,”* 56 WM. & MARY L. REV. 1793, 1829 (2015) (attempting to explain the clear-and-convincing standard as requiring that “the plaintiff’s explanation must be clearly and convincingly better than the defendant’s explanation”).

21. See Friedman, *supra* note 19, at 2046–47 (discussing relative plausibility theory and standards of decision).

22. See Ronald J. Allen, *Standards of Proof and the Limits of Legal Analysis* 14 (Northwestern University School of Law, Public Law and Legal Theory Research Paper Series, May 3, 2011), <http://ssrn.com/abstract=1830344> [<http://perma.cc/AAR6-5X5W>] (noting the inconsistencies between jury instructions and relative plausibility); Pardo, *supra* note 7, at 1093 (discussing ambiguities in the language of the standard for preponderance of evidence).

to proceed element-by-element and apply the standard of proof to each element, not to create holistic stories and compare them.²³ Third, it diverges from the law by compelling the nonburdened party, or at least imposing a practical obligation, to choose and formulate a competing version of the truth.²⁴ The law allows the defendant to stand mute and still prevail.²⁵ Fourth, comparing the plaintiff's story only to the defendant's favorite story, rather than to all versions of nonliability, will result in recovery by plaintiffs more often than normatively desirable.²⁶ The plaintiff should lose if liability is less likely than nonliability, regardless of which story the defendant prefers.²⁷ Fifth, the theory comes

-
23. See 3 KEVIN F. O'MALLEY, JAY E. GREINIG & WILLIAM C. LEE, FEDERAL JURY PRACTICE AND INSTRUCTIONS: CIVIL § 104.01 (6th ed. 2011) ("Plaintiff has the burden in a civil action, such as this, to prove every essential element of plaintiff's claim by a preponderance of the evidence. If plaintiff should fail to establish any essential element of plaintiff's claim by a preponderance of the evidence, you should find for defendant as to that claim.").
24. Cheng, *supra* note 3, at 1262 n.15.
25. See 2 MCCORMICK ON EVIDENCE, *supra* note 9, § 339, at 660–61 (explaining that because juries bring their own experiences, it is possible for a verdict to find for a defendant who offers nothing in opposition to the plaintiff's evidence).
26. Realization of this difficulty leads some theorists to argue that the aim of the trial system is not truth and minimizing error costs but, say, acceptability of decision. See Nesson, *supra* note 11, at 1390 (describing how the legal system strives to find the most likely of all stories). In fact, those theorists have a better argument for looking only at the defendant's best story than they realize. If they were to extend relative plausibility's "ordered partition" theory into a more sophisticated "ranking theory," then they could show that the looking at defendant's best story allows ignoring all other defendant stories. But accepting ranking theory works out to be the equivalent of accepting belief functions. See Franz Huber, *Belief and Degrees of Belief*, in DEGREES OF BELIEF 1, 16–20 (Franz Huber & Christoph Schmidt-Petri eds., 2009) (comparing belief theory and ranking theory).
27. See David Hamer, *Probabilistic Standards of Proof, Their Complements and the Errors That Are Expected to Flow from Them*, 1 U. NEW ENG. L.J. 71 (2004) (discussing probability theory); D.H. Kaye, *The Error of Equal Error Rates*, 1 LAW, PROBABILITY & RISK 3, 7 (2002) (arguing that the $p > 0.5$ rule is appealing because "it minimizes expected losses"); Mark Schweizer, *Loss Aversion, Omission Bias and the Civil Standard of Proof*, in EUROPEAN PERSPECTIVES ON BEHAVIOURAL LAW AND ECONOMICS 125, 128–32 (Klaus Mathis ed., 2015) (describing the decision theoretic framework); cf. Neil Orloff & Jerry Stedinger, *A Framework for Evaluating the Preponderance-of-the-Evidence Standard*, 131 U. PA. L. REV. 1159, 1168–71 (1983) (considering bias in the distribution of errors).

with baggage.²⁸ It requires, for example, acceptance of some holistic account of factfinding like the story model.²⁹

There must be a better, and perhaps simpler, way to conceive the standard of proof. The mathematical theory of belief functions provides an alternative superior to traditional probability theory and to the newer approach of relative plausibility. Part I will explain belief functions and how they can help understand the idea of a standard of proof applied to a single element of a case. Part II uses the theory to explain both the burden of persuasion and its associated array of standards of proof. Part III then uses the theory to explain both the burden of production and its role in safeguarding certain process and outcome values. Part IV steps back from the theory to see how it will work in the real world, including how belief functions work when applied to the multiple elements of a case.

I. PROOF BY FACTFINDERS' BELIEFS

The first step on the journey is to realize that the key assumption of classical logic makes every proposition absolutely either true or false, an assumption called the principle of bivalence.³⁰ Multivalent logic instead allows propositions to be both true and false to a degree, so they can take on middle values of truth.³¹ Consequently, classical logic has no tools for handling partial truths, propositions that will forever be uncertainly stuck partway between false and true. Traditional probability theory, a mathematical supplement to classical logic, treats only the random odds of a proposition turning out to be either false or true. Contrariwise, multivalent logic developed to handle partial truths. Fuzzy logic is one example of multivalent logic.³² Deciding how to proceed in a world of persisting uncertainty (including how to combine partial truths) logically differs from predicting how uncertainty will resolve itself into certainty (including how to calculate the odds of multiple events occurring together).

28. See Craig R. Callen, Commentary, *Kicking Rocks with Dr. Johnson: A Comment on Professor Allen's Theory*, 13 CARDOZO L. REV. 423, 432–39 (1991) (discussing some problems with the relative plausibility theory).

29. See *id.* at 435–39 (explaining the story model).

30. See THEODORE SIDER, LOGIC FOR PHILOSOPHY 73 (2010) (stating that classic logic is “bivalent,” with exactly two truth values).

31. See generally J.C. BEALL & BAS C. VAN FRAASSEN, POSSIBILITIES AND PARADOX: AN INTRODUCTION TO MODAL AND MANY-VALUED LOGIC (2003) (providing an overview of multivalent logic).

32. *Id.*

Even in the absence of that fundamental realization, a common way for factfinders to express their assessment of evidence³³ is as a gradation of belief in a proposition, where the gradation may take any numerical (or usually nonnumerical) value throughout the whole interval from zero to one. There is nothing controversial about such a formulation, which is compatible with the bivalently based approaches of traditional probability or relative plausibility.

Factfinders could, however, use multivalent “degrees of belief.” These turn out to be an especially useful way to express likelihood because they can capture all the various kinds of uncertainty in the world. The world exhibits several kinds of uncertainty, including the uncertainty characterized as the vagueness of matters of degree and also the indeterminacy resulting from scarce information or conflicting evidence. By following multivalent logic’s rejection of the assumption that all things are either completely true or completely false, degrees of belief can employ the middle values of truth to pick up the extra information about all these uncertainties.³⁴

By contrast, the probability calculus breaks down when the assumption of bivalence no longer holds, as it does in the task of factfinding. The process of proof investigates a world that is not a two-valued world where disputed facts are either true or false. Instead, a good portion, but not all, of the real world is a vague, imprecise, or many-valued world, where fuzzy partial truths exist. Or, the factfinders might never be able to learn whether a disputed fact is certainly true or false, so that any absolute truth remains inaccessible to their minds. Therefore,

33. My interest in this Article is not so much the factfinders’ initial processing of evidence, but rather the subsequent steps that involve their application of a standard of proof to the assessment of the evidence. Application of a standard of proof is a step largely separable from evidential argument. See CLERMONT, *supra* note 1, at 123–29 (detailing the difference between applying a standard of proof and assessing evidence). Although psychologists can tell us something about how humans process evidence, they have contributed almost nothing on how humans would apply standards of proof, leaving the dispute to logicians so far. Fortunately, my discussion of the standard of proof is compatible with any method used initially to process pieces of evidence that reinforce or undermine each other. For example, it would accept factfinders’ using intuitive techniques in a nonquantitative and approximate fashion to evaluate and combine evidence on a factual element, as they pursue the abductive task of seeking truth while they interlace inductive premises and perform deductive testing of interim conclusions. Compare Edmund M. Morgan, *Introduction to Evidence*, in AUSTIN W. SCOTT & SIDNEY P. SIMPSON, CASES AND OTHER MATERIALS ON CIVIL PROCEDURE 941, 943–45 (1950) (discussing the logical methods jurors use to process evidence), with Mark Spottswood, *The Hidden Structure of Fact-Finding*, 64 CASE W. RES. L. REV. 131 (2013) (applying the dual-process psychological framework to legal factfinding).

34. GLENN SHAFER, A MATHEMATICAL THEORY OF EVIDENCE (1976).

humans can accurately represent certain things in the world, such as historical fact, only as partial truths or degrees of belief.

If one embraces this approach, the general problem becomes how to handle a *degree of belief* about the *real world* that will persist unavoidably in *factfinding*. First, by “degree of belief” I do not mean the odds of something being eventually revealed to be an absolute truth, but rather I mean a belief that expresses a fuzzy degree of certainty about the state of the world as represented by the available evidence and that lies somewhere between holding the thing completely false and holding it completely true. Beliefs represent neither firm knowledge nor some squishy personal feeling or strictly internal whim. A belief is an attempt rationally to evaluate the evidence in the pursuit of truth.³⁵ Second, the “real world” is the world as perceived by humans and described by natural language.³⁶ Third, I refer to “factfinding” in its broad sense, as covering anything that a court or other entity subjects to a proof process in order to establish what the entity will treat as truth. It would include application of law to fact, as well as pure fact.

This Part will introduce the theory of belief functions, showing it to work well in representing how imperfect evidence keeps factfinders from committing all of their belief. Then, this Part will introduce the idea of comparing belief and disbelief of a fact, which the factfinder would do only after putting any uncommitted belief aside.

A. Belief Functions

I can pick up the theoretical developments with emergence of the field of imprecise probability.³⁷ This modern field of mathematics provides a useful extension of probability theory whenever information is scarce or conflicting. The basic idea is to use interval specifications of

-
35. See Glenn Shafer, *The Construction of Probability Arguments*, 66 B.U. L. REV. 799, 801–04 (1986) (contrasting three other interpretations of probability); cf. David Enoch, Levi Spectre & Talia Fisher, *Statistical Evidence, Sensitivity, and the Legal Value of Knowledge*, 40 PHIL. & PUB. AFFAIRS 197, 211–15 (2012) (arguing that law’s primary interest is accurate determination of truth, not knowledge); Pardo, *supra* note 20, at 1810–16 (linking “justified true belief” to knowledge). *But cf.* DAVID CHRISTENSEN, PUTTING LOGIC IN ITS PLACE 12–13, 69 (2004) (saying that for certain purposes some theorists use “belief” as an unqualified assertion of an all-or-nothing state of belief); L. Jonathan Cohen, *Should a Jury Say What It Believes or What It Accepts?*, 13 CARDOZO L. REV. 465 (1991) (using “belief,” for his purposes, in the sense of a passive feeling).
36. See D. Michael Risinger, *Searching for Truth in the American Law of Evidence and Proof*, 47 GA. L. REV. 801 (2013) (treating the necessary philosophical assumptions).
37. See generally PETER WALLEY, STATISTICAL REASONING WITH IMPRECISE PROBABILITIES (1991) (detailing imprecise probabilities).

probability, with a lower and an upper probability.³⁸ The rules associated with traditional probability, except those based on assuming an excluded middle, carry over to imprecise probability. Despite its name, imprecise probability theory is more complete and accurate than precise probability in the real world where uncertainty prevails. Imprecise probability can work with multivalent logic as well as with classical logic. In fact, traditional probability built on bivalence appears as a special case in this theory.³⁹

Belief function theory is a further step toward set theory.⁴⁰ To put its multivalent degrees of belief in set theory's terms, the gradation of belief expresses the assessment in terms of the proposition's imprecise and indeterminate degree of membership in the set of true facts. Belief function theory does not constitute a system of logic, unlike multivalent logic. Instead, it remains a branch of mathematics. It indeed rests on a highly rigorous mathematical base, managing to get quite close to achieving a unified theory of uncertainty.⁴¹ Just as traditional probability serves bivalent logic by mathematically handling a kind of uncertainty for which the underlying logic system does not otherwise account,⁴² belief function theory delivers mathematical notions that can extend a logical system. Simple fuzzy logic can pick up the vagueness of the world. Belief function notions can supplement that logic by capturing and expressing in an easy and comprehensible way the indeterminacy resulting from scarce or conflicting evidence concerning fact.⁴³

38. *Id.*

39. *Id.*

40. *See generally* SHAFER, *supra* note 34 (describing belief function theory); Glenn Shafer, *Perspectives on the Theory and Practice of Belief Functions*, 4 INT'L J. APPROXIMATE REASONING 323 (1990) (discussing the implementation of the theory of belief functions). For an accessible introduction, see DAVID A. SCHUM, *THE EVIDENTIAL FOUNDATIONS OF PROBABILISTIC REASONING* 222–43 (1994) (showing the current applications of belief functions). For a fuller historical account, see Rolf Haenni, *Non-Additive Degrees of Belief*, in *DEGREES OF BELIEF* 121, 127–33 (Franz Huber & Christoph Schmidt-Petri eds., 2009) (detailing the history of the theory).

41. *See* Didier Dubois & Henri Prade, *A Unified View of Uncertainty Theories* (Mar. 7, 2012) (unpublished manuscript).

42. *See* IRVING M. COPI, CARL COHEN & KENNETH MCMAHON, *INTRODUCTION TO LOGIC* ch. 14 (14th ed. 2011).

43. I have previously surveyed the various theories on how to handle uncertainty, and explored the ones that best image how the law contemplates uncertainty. For the purposes of expressing imprecise evidential assessments in probability-like terms and conjoining separate findings, fuzzy logic is the optimal theory. *See supra* note 6 (describing and justifying fuzzy logic). For the purposes of understanding how to apply the standards of proof, the compatible theory of belief functions is more expressive. *See* CLERMONT, *supra* note 1, at 201–20 (explaining, and preferring, belief functions as a way to account for imperfect legal evidence and to apply the standards of proof); *cf. id.* at 162–63 (discussing

The key distinction between degrees of belief and probabilities is subtle, as attested by the confusion among people discussing proof over the long years. Both systems numerically quantify uncertainty by using measurements in the unit interval $[0,1]$. But the distinction's consequences are not subtle. Degrees of belief accommodate vagueness better than traditional probability theory, and they better capture the effect of imperfect evidence. The measure of the factfinder's complexly constructed belief in the real world is more relevant than the probability of truth in an imagined world of merely random uncertainty. Finally, belief functions give the tools for translating beliefs about facts into decisions.

1. Basics of Theory

Given that factfinders should not ask how probable is proposition S but rather what is their degree of belief in S as a true proposition (or the degree of S 's truth, which is a degree of membership in the set of true facts), I propose considering a broad version of belief functions as a legal model for human expression of likelihood.⁴⁴ It will give us a handle on how to employ and manipulate such beliefs.⁴⁵

The broad version of the theory of belief functions will also provide us with a good mental image for representing indeterminacy.⁴⁶ On the basis of incomplete, inconclusive, ambiguous, dissonant, or untrustworthy evidence, some of the factfinders' belief should remain indeterminate. In factfinding, we ask how much we believe S to be a real-world truth based on the evidence, as well as how much we believe not S —while remaining conscious of indeterminacy and so recognizing that part of our belief will remain uncommitted.

Beliefs can range anywhere between zero and one. If the belief in S is called $\text{Bel}(S)$, then $0 \leq \text{Bel}(S) \leq 1$. Likewise, belief in not S , which is disbelief of S or belief in S 's contradiction, falls between zero and one.

so-called ultrafuzzy sets as a way to handle imperfect evidence, but acknowledging that the logical operators for ultrafuzzy sets become complicated), 202–03 (describing possibility theory as a way to exploit the compatibility of fuzzy logic and belief functions).

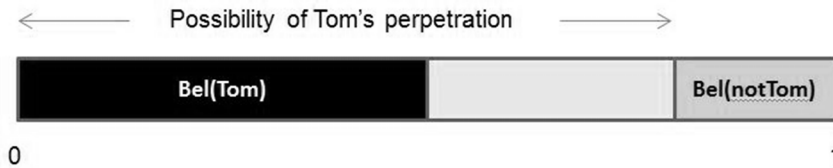
44. Cf. Vilém Novák, *Modeling with Words*, SCHOLARPEdia (2008), http://www.scholarpedia.org/article/Modeling_with_words [<http://perma.cc/2JWT-YLZ8>] (“Mathematical fuzzy logic has two branches: fuzzy logic *in narrow sense* (FLn) and fuzzy logic *in broader sense* (FLb). FLn is a formal fuzzy logic which is a special many-valued logic generalizing classical mathematical logic FLb is an extension of FLn which aims at developing a *formal theory of human reasoning*.”).
45. See SHAFER, *supra* note 34, at 35–37 (discussing belief functions).
46. See Liping Liu & Ronald R. Yager, *Classic Works of the Dempster-Shafer Theory of Belief Functions: An Introduction*, in CLASSIC WORKS OF THE DEMPSTER-SHAFER THEORY OF BELIEF FUNCTIONS 1, 2–19 (Ronald R. Yager & Liping Liu eds., 2008) (formalizing an image of belief functions).

Under the scheme of belief functions, we are squarely in the realm of nonadditive measures.⁴⁷ In other words, a belief and a belief in its contradiction will normally add to less than one. The zone between $\text{Bel}(S)$ and $\text{Bel}(\text{not}S)$ represents the uncommitted belief.

Factfinders form their beliefs and disbeliefs based on the available evidence. Jurors might believe a fact more than they disbelieve it, even if they would not be willing to bet on it as more likely than not if the truth could somehow be revealed. The belief might be quite weak as it rests only on what evidence is available. Contrariwise, a bet must commit total belief to either yes or no, and betting odds always add to one. Thus, a historian of the French Revolution might believe that Robespierre did such-and-such on a given day, but not be willing to bet on the act versus all other possibilities. The historian would also go on to construct a believed narrative of the Reign of Terror without ever treating his or her beliefs as betting odds, for example, by multiplying them to get conjoined facts. This *fundamental difference between non-additive beliefs and betting odds* is subtle but essential, as all else follows from it.

Consider belief function theory's treatment of a single factual hypothesis. Take as an example the issue of whether Tom was the perpetrator of a crime. Although you have no definitive evidence, three witnesses say he was. One seems somewhat credible. But you think that another saw at the scene a different man at a different time, which discounts this evidence of guilt but gives no support to his being innocent. And you think that the third might be lying as part of a cover-up of some other person's guilt, an interpretation that is compatible with both guilt and innocence and so gives some thin support to Tom's not being involved. In sum, this body of evidence supports your .5 belief that Tom was the perpetrator, or $\text{Bel}(\text{Tom})$. The evidence also supports your weaker belief that Tom was not involved, with $\text{Bel}(\text{notTom})$ coming in at .2. That is, $\text{Bel}(\text{notTom})$ is not determined by the value of $\text{Bel}(\text{Tom})$. The remaining .3 is indeterminate, meaning he could be either the perpetrator or not. The evidence is imperfect. The defects in evidence might be probative, affecting $\text{Bel}(\text{Tom})$ or $\text{Bel}(\text{notTom})$; but the defects might be nonprobative, so that they just leave some belief uncommitted. (This example actually involves a so-called power set of four beliefs: Tom, notTom, neither Tom nor notTom, and either Tom or notTom. The belief in the "null" of Tom's being neither perpetrator nor uninvolved is set by definition to be zero. The belief in the "catchall" of Tom's being either perpetrator or uninvolved is one.)

47. See generally Haenni, *supra* note 40, at 121–27 (describing nonadditive degrees of belief); Ron A. Shapira, *Economic Analysis of the Law of Evidence: A Caveat*, 19 CARDOZO L. REV. 1607, 1613–16 (1998) (distinguishing additive from nonadditive).



Again, $\text{Bel}(\text{Tom})$ is the extent to which you believe Tom to be the perpetrator. That belief is sometimes called the lower probability. The upper probability bound represents “possibility.”⁴⁸ It is the extent to which you think his being the perpetrator is possible, that is, the sum of the affirmative belief plus the indeterminate belief. The indeterminate zone between a belief expressed as $\text{Bel}(\text{Tom})$ and the belief in its contradiction expressed as $\text{Bel}(\text{notTom})$ represents uncommitted belief owing to imperfect evidence. The possibility that he is the perpetrator is .8, being $.5 + .3$. The possibility that he was not the perpetrator totals .5, being $.2 + .3$. Possibility equals one minus the belief in the contradiction. (A traditionally expressed probability of his being the perpetrator would fall within the range from the lower to the upper probability.⁴⁹)

The resultant beliefs can be expressed, if expression is ever necessary, as coarsely gradated beliefs. In addition to the benefits of utilizing natural language, these terms convey the uncertainty in determining the belief. Thus, in lieu of expressing beliefs in terms of decimals, one should use the coarse gradations of (1) slightest possibility, (2) reasonable possibility, (3) substantial possibility, (4) equipoise, (5) probability, (6) high probability, and (7) almost certainty. The coarseness of this scale of likelihood also means that the factfinder in comparing beliefs will not have to draw paper-thin distinctions.

In the end, the representation of findings in the form of nonnumerical beliefs best captures the effect of imperfect evidence, which was a

48. See Jeffrey A. Barnett, *Computational Methods for A Mathematical Theory of Evidence*, in CLASSIC WORKS OF THE DEMPSTER-SHAFFER THEORY OF BELIEF FUNCTIONS 197, 200–01 (Ronald R. Yager & Liping Liu eds., 2008) (providing a neat mental image for these bounds); A.P. Dempster, *Upper and Lower Probabilities Induced by a Multivalued Mapping*, 38 ANNALS MATHEMATICAL STAT. 325 (1967) (providing the mathematical proof for upper and lower probabilities); L.A. Zadeh, *Fuzzy Sets as a Basis for a Theory of Possibility*, 1 FUZZY SETS & SYS. 3 (1978) (relating the theory of fuzzy sets to the theory of possibility). In belief function terminology, “possibility” is often phrased as “plausibility.” See SCHUM, *supra* note 40, at 236 (using the phrase “plausibility” in place of “possibility”).

49. See Haenni, *supra* note 40, at 129 (discussing betting probability); cf. Glenn Shafer, *Belief Functions*, in READINGS IN UNCERTAIN REASONING 473, 475–76 (Glenn Shafer & Judea Pearl eds., 1990) (describing this probability in more complicated settings).

rallying cry of Baconian theorists.⁵⁰ The move from probability to belief is also a slight nod to the civil-law emphasis on inner belief as captured by its *intime conviction* standard,⁵¹ and to the frequent *cris de coeur* of theorists who lament any intrusion of probabilistic mathematics into the very human process of proof.⁵²

2. Negation Operator

By traditional probability theory, the probability of a proposition's negation (or contradiction) equals 1 minus the probability of the proposition. If Tom is sixty percent likely the perpetrator, he is forty percent likely uninvolved.

However, $\text{Bel}(\text{Tom})$ and $\text{Bel}(\text{notTom})$ do not necessarily add to 1, because normally some belief remains uncommitted. Thus, for Tom, $\text{Bel}(\text{Tom})=.5$ and $\text{Bel}(\text{notTom})=.2$, so the sum of determinate beliefs adds to .7.

The *complement of Bel(Tom)* equals $(1 - \text{Bel}(\text{Tom}))$, but this gives the possibility of notTom, not the belief in notTom. Indeed, the possibility of notTom equals $(\text{Bel}(\text{notTom}) + \text{uncommitted belief})$. Hence, there is a big difference between the complement and the *belief in the*

-
50. See L. JONATHAN COHEN, *THE PROBABLE AND THE PROVABLE* 49–57, 245–64 (1977) (developing, as an alternative to Pascalian (or mathematicist) probability, a Baconian (or inductive) theory of probability). Baconian theory tries to look not only at the evidence presented, but also at the evidence not available. It makes *evidential completeness* a key criterion, and thereby stresses an important concern. Cf. Hans Rott, *Degrees All the Way Down: Beliefs, Non-Beliefs and Disbeliefs*, in *DEGREES OF BELIEF* 301, 306 (Franz Huber & Christoph Schmidt-Petri eds., 2009) (seeing Baconian probability as a variation on belief functions).
51. See CLERMONT, *supra* note 1, at 221-72 (discussing the divergence between common-law and civil-law countries' standards of proof). With their emphasis on "conviction" in the *intime conviction* standard, the civil-law countries signal their devotion to belief, albeit a belief seemingly built upon a binary worldview (and perhaps a belief compared to an absolute threshold inherited from the criminal model). Such an approach fit better with an inquisitorial model than it did with an adversarial model, allowing it to persist for centuries. But its survival until today may rest instead on the civil-law system's desire to enhance the appearance of legitimacy.
52. See, e.g., Jaffee, *supra* note 11, at 934–51 (1985) (attacking the use of probability in analyzing proof); Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971) (writing the classic version of the lament, in which Professor Tribe stressed not only the risk of misuse of mathematical techniques, including inaccurate meshing of numerical proof with soft or unquantifiable variables, but also the undercutting of society's values, including the dehumanization of the legal process); Adrian A.S. Zuckerman, *Law, Fact or Justice?*, 66 B.U. L. REV. 487, 508 (1986) (arguing that probabilistic assessment diminishes "the hope of seeing justice supervene in individual trials," while seeing factfinding as an individualized but value-laden process).

negation, the difference being the uncommitted belief. Belief function theory thus utilizes the very useful distinction between a disbelief and a lack of belief. After all, disbelief and lack of belief are entirely different states of mind.

In sum, a belief in S normally does not imply anything about the belief in notS, other than that the contradiction cannot be more likely than the complement. Given scarce information or conflicting evidence, one forms a belief in a proposition, while leaving a lot of belief uncommitted and without necessarily forming any belief in the proposition's contradiction. One would need proof or inference of the contradiction before generating any belief in it.

3. Lack of Proof

Traditional probability encounters legendary difficulties with a state of ignorance.⁵³ The reason is that it cannot distinguish between belief, lack of belief, and disbelief. Assume that the factfinder dutifully starts by setting the belief in S at zero. In classical terms, $S=0$ means that S is impossible. And it means that notS is certain. No amount of evidence could alter an impossibility or a certainty into a possibility under Bayes' theorem.⁵⁴ As a way out, probabilists sometimes assert that the ignorant inquirer should start in the middle where the probabilities of true and false under the applicable standard of proof are both fifty percent. But this trick does not accord with the actual probabilities or with the law's instructions. That is, the supposition of fifty percent, on the thought that the fact is either true or false, comports neither with reality nor with where the law tells the factfinder to begin, and it produces inconsistencies when there are more than two hypotheses in play.⁵⁵

Meanwhile, one of the great strengths of belief function theory is that it can well represent a state of ignorance.⁵⁶ An inquirer, if ignorant and well behaved, starts at zero, not at a fifty-percent belief. When $Bel(S)=0$, it does not mean that S is so highly unlikely as to be impossible. It means there is no evidence in support. Accordingly, the

53. See Richard Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, 66 B.U. L. REV. 439, 462–67 (1986) (noting that employing 1:1 as the appropriate odds for someone who is ignorant of the true facts can cause many problems).

54. See Lea Brilmayer, *Second-Order Evidence and Bayesian Logic*, 66 B.U. L. REV. 673, 686–88 (1986) (considering second-order evidence that cannot be accommodated by the Bayesian framework).

55. See *State v. Spann*, 617 A.2d 247, 254 (N.J. 1993) (“.5 assumed prior probability clearly is neither neutral nor objective”); Jaffee, *supra* note 11, at 980–85 (discussing the “incompatibility of proper belief-formation and ‘subjective probability’ statistics”).

56. See SHAFER, *supra* note 34, at 22–24 (referring to belief function as “representation of ignorance”).

inquirer starts out with everything indeterminate, because the lack of evidence makes one withhold all of one's belief. Although $\text{Bel}(S)=0$, $\text{Bel}(\text{not}S)$ equals zero too. The uncommitted belief is the entirety or 1, meaning that S is completely possible, as is $\text{not}S$. In other words, the inquirer does not believe or disbelieve S . Belief function theory thus utilizes the very useful notion of lack of belief.

B. Comparison of Beliefs

My conceptualization has thus far led me to think that the law should not and does not employ the prevailing academic view of the proof process resting on a bivalent logical approach. Factfinders instead determine their beliefs as degrees of real-world truth based on the evidence, just as the law expects of them. Eventually they end up with $\text{Bel}(S)$ and $\text{Bel}(\text{not}S)$, falling between zero and one, but not normally adding to one. What then do they do?

So, finally, I come to the matter of applying a standard of decision. The law dictates that factfinders decide by subjecting their beliefs to a standard of proof in order to come to an unambiguous output. That is, at this point the law forces factfinders back into what looks like a two-valued logic, by making them decide for one party or the other.

The determined theorist could pursue the bivalent image of traditional probability. Then the ultimate task of applying a standard of proof would unavoidably involve placement on a scale of likelihood running from 0% to 100%.⁵⁷ But I contend that speaking in terms of bivalent logic tends to mislead on standards.

A better understanding of standards of proof would result from thinking in terms of multivalent logic and belief functions. Even though decisionmaking requires converting from a multivalent logic to an output that sounds two-valued, the law does not need to require evidence to make the fact more likely than fifty percent or whatever. All the factfinder need do is compare the strengths of belief and disbelief. The path to decision might involve only comparing $\text{Bel}(S)$ and $\text{Bel}(\text{not}S)$ while ignoring the indeterminate belief. By requiring only a comparison, belief functions would never require placement on a scale of likelihood. The workings of this approach is the subject of the next Part.

II. BURDEN OF PERSUASION

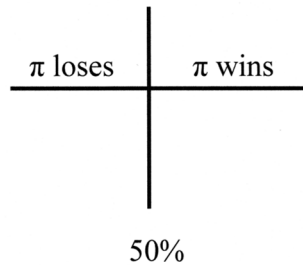
This Part will use belief function theory to explain why the traditional view of the law's burden of persuasion misrepresents the standards of proof. Then, this Part will demonstrate how the law actually conceives of its three standards of proof as different ways of comparing belief and disbelief.

57. See, e.g., *Brown v. Bowen*, 847 F.2d 342, 345 (7th Cir. 1988) (“[T]he trier of fact rules for the plaintiff if it thinks the chance greater than 0.5 that the plaintiff is in the right.”).

A. Traditional View

Let me start with some background on how the law has traditionally viewed the burden of proof, say, in a jury trial. The burden of proof dictates who must produce evidence and ultimately persuade the jury on which elements of the case. Burden of proof thus encompasses two concepts: burden of production and burden of persuasion. The burden of production might require either party at a given time during trial to produce evidence on an element or suffer the judge's adverse determination on that element; one party has the initial burden of production on any particular element, but that burden may shift during the trial if that party produces certain kinds or strengths of evidence. The burden of persuasion requires a certain party ultimately to persuade the jury of the truth of an element or suffer the jury's adverse determination on that element.

To convey the traditional view of the burden of persuasion, a diagram must represent the internal thought process of the factfinder in ultimately weighing the evidence. The following grid measures the factfinder's view in a civil case of the evidential likelihood that the disputed fact exists, with likelihood increasing from 0% on the left to 100% on the right.



The plaintiff in an imagined civil trial starts at the left. By presenting evidence on the issue, he must get beyond the midpoint to win. That is, he must show that it is more likely than not that the disputed fact exists. If after the plaintiff has given his best shot the factfinder thinks that he has not passed the fifty-percent line, then the factfinder should decide for the defendant.

This diagrammatic representation of the burden of persuasion appears to be compatible with traditional probability. Even accepting the unrealism of a probabilistic theory, however, a qualification is necessary that this diagram serves mainly as an impetus to thinking about these matters rather than as a source of definitive statements thereon.

The diagram can confuse. For example, the law handles a finding of equipoise in a civil case by means of the burden of persuasion. If the evidence ends up as evenly matched, the burden-bearer loses. That observation usually generates the reaction that the burden of persuasion does not matter much. After all, in theory, it should work only as a

tiebreaker in the highly unusual case of a precise tie. Yet, in practice, lawyers and judges fight and suffer over the burden of persuasion. Why?

First, the diagram does not mean that a fifty-percent line exists in reality. The psychological truth is that equipoise is more of a zone, or range of probabilities, than a line. As already suggested, a useful way to envisage the whole scale of likelihood is as a set of fuzzy categories, or intervals, of likelihood.⁵⁸ Each category, such as more likely than not, embodies some range of approximate likelihood. Equipoise is no different from the other categories. A range of evidential states may strike the factfinder as evenly balanced. Equipoise being a zone means that the burden of persuasion will affect many more cases than those few in which the conflicting evidence results precisely in a dead heat.⁵⁹

Second, given the selection effect, close cases are common.⁶⁰ Uneven cases falling far from the standard of proof tend to settle, while the cases where the parties can disagree on the predictions of outcome tend disproportionately to go to trial.

Third, still other reasons for caring about the persuasion-burden lie in psychology. How the law frames a question—whether the plaintiff or the defendant bears the risk of nonpersuasion of a fact, that is, whether the plaintiff or the defendant appears to start from “zero”—matters.⁶¹ An anchoring heuristic lowers the willingness of the factfinder to determine that the burdened party has prevailed, because people fail to adjust fully from a given starting point, even if the starting point was arbitrarily set.⁶² Another reason is that loss aversion, status quo bias, and omission bias make it difficult for the burdened party to carry the burden. The idea is that the disutility generated by a defendant’s loss

58. *Supra* Part I.A.1.

59. *See* United States *ex rel.* Bilyew v. Franzen, 686 F.2d 1238, 1248 (7th Cir. 1982) (stressing importance of the persuasion burden and observing that “a judge or a jury can experience only a small, finite number of degrees of certainty Thus cases when the evidence . . . seem[s] in balance are not unique among some infinite variety of evidentiary balances, but instead are among a much smaller number of [ranges of] possibilities that may be perceived by the fact-finder.”).

60. *See* George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1 (1984) (presenting a nonrandom model of the relationship between disputes settled and disputes litigated, focusing on solely economic determinants and the parties’ rational estimates of the outcome).

61. *See* Eyal Zamir & Ilana Ritov, *Loss Aversion, Omission Bias, and the Burden of Proof in Civil Litigation*, 41 J. LEGAL STUD. 165, 197 n.23 (2012) (describing the challenge that placing the burden of proof on the plaintiff adds even with no heightened standard of persuasion).

62. *See* Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCIENCE (n.s.) 1124, 1128 (1974) (describing the anchoring phenomenon that happens when people make estimates starting from an initial value).

of something by verdict exceeds the utility reaped from a plaintiff's equal gain; that the burdened party seems undesirably intent on disrupting the status quo; and that the factfinder's harmful commission by making an award would inflict more cost psychologically than would its harmful omission. Accordingly, people perceive the loss to the defendant by a judgment as larger than the gain to the plaintiff, while such a judgment requires the system to elevate action over inaction and so risk incurring regret costs.⁶³

Thus, the factfinder will rely on the burden of persuasion more often than one might imagine. But having to draw a fat fifty-percent line encourages a more general reconsideration of the proof standards. The conclusion will be that this diagram for the burden of persuasion is fundamentally misleading, in need of redrawing rather than mere refinement. The redrawing will entail a reformulation of those standards into a diagram of belief functions. The law does not and should not conform to the traditional academic view.

B. Reformulated View

The traditional approach, reeking of probability theory, does not do a terribly good job of accounting for the usual state of evidence. It poses odd questions to the factfinder: given imperfect evidence, what is the chance the plaintiff is right in an absolute sense, and how does that chance compare to the applicable standard of proof?

The law has settled on three standards of proof that apply in different circumstances: (1) The standard of preponderance of the evidence translates into a more-likely-than-not standard. It is the usual standard in civil litigation, but it appears throughout law.⁶⁴ (2) Next comes the intermediate standard or standards, often grouped under the banner of clear and convincing evidence and roughly translated as a much-more-likely-than-not standard. These variously phrased but equivalently applied standards govern on certain issues in special situations, such as when terminating parental rights.⁶⁵ (3) The standard of proof beyond a reasonable doubt means a virtual-certainty standard. It very rarely prevails outside criminal law.⁶⁶

In applying the standard, belief functions better reflect the factfinder's actual frame of mind: some belief will remain uncommitted in the absence of perfect evidence. Instead of betting that Tom's identity is or is not probable, the factfinder should think and speak in terms of

63. See Schweizer, *supra* note 27, at 134–38 (explaining how cognitive psychological factors, like loss aversion, omission bias, and status quo bias, heighten the standard of proof).

64. See CLERMONT, *supra* note 1, at 16–18 (outlining the debate over the meaning of the preponderance-of-the-evidence standard).

65. See *id.* at 23–25 (describing the clear-and-convincing-evidence standard).

66. See *id.* at 26–31 (describing the proof-beyond-a-reasonable-doubt standard).

degrees of belief. That is, on a fact to which the standard of proof applies, the belief function route is the one to take, rather than invoking a simplistic scale of likelihood.

The factfinder would ask itself the natural question that the law seems to pose to it: do you believe the burdened party's allegations more than you disbelieve them? The factfinder need not compose betting odds. (Nor does the factfinder have to follow the invented approach of relative plausibility. The other party does not have to choose and formulate a story. It can, even by silence, take advantage of disjunction if the factfinder disbelieves any element as much as or more than it believes the element.)

Finally, belief functions could invoke the factfinder's considerable powers of relative judgment rather than its absolute judgment of likelihood. What could the standards of proof mean in a comparative sense? To begin, what could preponderance of the evidence, or its translation as more likely than not, mean in a comparative sense?

1. Preponderance of the Evidence

One could measure the proof against some absolute threshold and require, say, that the evidence have some specified content. But for ages such a formulation has not accorded with the import of real cases. The law should be more comparative. It does not simply inquire which side has the stronger evidence, however. It looks instead to belief in the burdened party's position.⁶⁷

The law could examine the belief for some absolute strength, say, requiring that $\text{Bel}(S)$ exceed a fifty-percent likelihood. But it does not require the completeness of proof that would be necessary to get a belief above fifty percent. It is willing to rest decisions on the evidence presented. Accordingly, any talk of requiring elements to exceed fifty percent is misleading. The better approach is more directly to invoke the more powerful human ability of relative judgment by comparing beliefs.

One could compare $\text{Bel}(S)$ relative to $\text{Bel}(\text{not}S)$.⁶⁸ In comparing them, $\text{Bel}(\text{not}S)$ is the belief in the negation of S , not the complement of $\text{Bel}(S)$. It represents how much the factfinder actively disbelieves S , the fact in dispute. The factfinder should ignore uncommitted belief and then compare the affirmative belief to any belief in its contradiction.

The comparison thus should look at actual belief in S and actual disbelief of S . If the factfinder were to work with only those two beliefs,

67. See J.P. McBaine, *Burden of Proof: Degrees of Belief*, 32 CALIF. L. REV. 242, 248–49 (1944) (examining what level of proof a factfinder needs to believe in the burdened party's position).

68. See COHEN, *supra* note 50, at 255 (“The cardinal question to be settled by the trier of fact may always be construed as this: on the facts before the court, is the conclusion to be proved by the plaintiff more inductively probable than its negation?”).

and discard the indeterminate belief, the most obvious course in civil cases would be to say that the burdened party should win if and only if $\text{Bel}(S) > \text{Bel}(\text{not}S)$. The factfinder would decide for the plaintiff if $\text{Bel}(S)$ exceeds $\text{Bel}(\text{not}S)$, but decide for the defendant if $\text{Bel}(S)$ does not exceed $\text{Bel}(\text{not}S)$. This standard not only is readily comprehensible but also avoids any need to quantify the beliefs.

Although belief functions do not require placement on a scale, the factfinder in effect might end in believing the burdened plaintiff's position on a disputed fact to be only "substantially possible" on the coarse scale of seven gradations. That situation does not mean that the plaintiff should lose, however. The factfinder might, if forced to express likelihood, believe the falsity of the plaintiff's position merely to a "reasonable possibility." The plaintiff should win, by use of belief functions. All the factfinder must do is to compare belief and disbelief: all that preponderance of the evidence requires is that the strength of the factfinder's belief that the plaintiff is right must exceed its belief that the plaintiff is wrong. Belief functions so add the idea that the factfinder in such a case must have a belief in the element's truth stronger than his belief in its falsity; while some of the factfinder's belief remains uncommitted, it did find the plaintiff's position to be a good one in the sense of more likely true than false. Sometimes, then, we would be talking of whether the smallish belief exceeds its smallish contradiction.

The comparative approach to the civil standard of proof means that the nonburdened party does not need to develop a competing version of the truth, but can rely on negation of any element. A belief in the falsity of the burdened party's version of the truth may develop naturally in the course of trial. It could arise even upon hearing only the burdened party's evidence. Any evidence from the nonburdened party should contribute to raising $\text{Bel}(\text{not}S)$.

Relatedly, the nonburdened party need not fight imaginary fights. Some scholars worry that looking at negation puts the burdened party in the impossible situation of disproving every alternative possibility.⁶⁹ But that worry comes from confusing lack of belief with disbelief. Disbelieving S entails the degree to which the factfinder thinks S is false. The mere possibility of other states of the world in which S is not true go into the uncommitted belief, not usually into $\text{Bel}(\text{not}S)$; recall that the possibility of $\text{not}S$ equals $\text{Bel}(\text{not}S)$ plus the uncommitted belief; again, the degree of believing that Tom is not the perpetrator is quite different from envisaging the chance that he is possibly not the perpetrator. The proposed comparison involves the belief in $\text{not}S$, and does not involve the possibility of $\text{not}S$.

Not only does this comparative approach comport with the natural cognitive method that follows from telling the factfinders they must

69. See, e.g., Pardo, *supra* note 7, at 1093–94 (observing this paradox created by a probabilistic approach to the more-likely-than-not standard, and so proposing a relative plausibility approach).

look to their beliefs and then decide for one side or the other, but also it does nothing to interfere with the current procedural and substantive functioning of the standard of proof. Picture a normalization process of disregarding the indeterminate beliefs and scaling $\text{Bel}(S)$ and $\text{Bel}(\text{not}S)$ up proportionately so that they add to 1. Call the recalculations $b(S)$ and $b(\text{not}S)$. If $\text{Bel}(S)=.50$ and $\text{Bel}(\text{not}S)=.20$, then $b(S)=.71$ and $b(\text{not}S)=.29$. These new numbers represent much less mental distance from the traditional view of standards of proof, because $b(S) > b(\text{not}S)$ if and only if $b(S) > .50$. Thus, preponderance could retain a connotation of likelihood exceeding fifty percent. This normalization renders the reformulation much less jarring, and it also demonstrates that I did not pull the reformulation out of thin air.

Furthermore, something about the traditional view of the preponderance standard as a showing of a probability greater than fifty percent just seems appropriate for civil cases: among competing fixed standards, it minimizes the expected number of erroneous decisions and also the expected sum of wrongful amounts of damages, which is a goal that the law has chosen to pursue by its civil standard. The reformulated standard has the same error-cost minimizing properties, but achieves them in the real world where the assumption of bivalence does not hold and where considerable indeterminacy prevails. For an idea of a proof adapted from the probabilists' proof,⁷⁰ let $b(S)=p$ be the apparent probability that the defendant is liable (for D dollars). If $\text{Bel}(S) > \text{Bel}(\text{not}S)$, then $p > \frac{1}{2}$; call p by the name p_1 in that case. If $\text{Bel}(S) \leq \text{Bel}(\text{not}S)$, call it p_2 . On the one hand, under the preponderance standard, the expected sum of false positives and false negatives over the run of cases is $[(1 - p_1)D + p_2D]$. On the other hand, under a very high standard that eliminates false positives, the analogous sum is $[p_1D + p_2D]$. Therefore, given that $(1 - p_1)$ is less than p_1 , the reformulated preponderance standard lowers the system's expected error costs.

Finally, one of civil procedure's most fundamental principles is that, where possible, the parties should ordinarily receive equal treatment and bear equal risk. The common law has long championed the traditional view of the preponderance rule as a manifestation of that principle. By comparing belief to disbelief, the reformulated standard preserves the level playing field.

2. Clear and Convincing Evidence

Now, as to the other standards of proof, clear and convincing evidence should mean $\text{Bel}(S) \gg \text{Bel}(\text{not}S)$.⁷¹ This standard would not be

70. See *supra* note 27 (discussing the preponderance standard as a $p > 0.5$ approach).

71. See McBaine, *supra* note 67, at 263 (proposing an instruction to the effect that "the probability that they are true or exist is substantially greater than the probability that they are false or do not exist"); Edmund M.

that difficult to apply. We are quite used to such a standard of being clearly convinced, in life and in law. Judges apply it on a motion for a new trial based on the verdict's being against the weight of the evidence. Appellate courts use it in reviewing judge-found facts. Those standards of decision mean that it is not enough to disagree with the jury or the judge, and instead the reviewer must think there was a serious error.

However, I admit that the cases do not make perfectly evident what clear and convincing means. Alternatively, or perhaps additionally, it imposes a requirement about the completeness of evidence. It may require admission of enough evidence to reduce uncommitted belief to the point that $\text{Bel}(S)$ exceeds the possibility of notS. I am open to those viewpoints, but unconvinced so far. In the meantime, one could partially capture the standard by explicating clear and convincing to the factfinder as the standard that lies between preponderance and reasonable doubt.

3. Beyond a Reasonable Doubt

As to proof beyond a reasonable doubt, it is demanding of course. It must require more than $\text{Bel}(S) \gg \text{Bel}(\text{not}S)$. Indeed, proof beyond a reasonable doubt seems to differ in kind, suggesting that it is not simply $\text{Bel}(S) \gg \gg \text{Bel}(\text{not}S)$. Instead, by placing separate demands on $\text{Bel}(S)$ and $\text{Bel}(\text{not}S)$, the criminal standard should mean that no great uncommitted belief remains and that no reasonable doubt persists.⁷²

“No great uncommitted belief” reflects the idea that $\text{Bel}(S)$ cannot be weak, measured in an absolute sense. We do not want to convict when, although there is some evidence of guilt, we really do not know what happened. The belief in guilt must outweigh all alternative possibilities, including fanciful ones. The *belief* in guilt must at least exceed the *possibility* of innocence, so that $\text{Bel}(S) > .50$. Given the usual limits on available evidence, achieving such a high degree of absolute belief represents a demanding standard.⁷³

Morgan, *Instructing the Jury upon Presumptions and Burden of Proof*, 47 HARV. L. REV. 59, 67 (1933) (“its truth is much more probable than its falsity”); cf. Laudan, *supra* note 8, at 299–300 (discussing attempts to append such notions to the approach of inference to the best explanation).

72. See Pardo, *supra* note 20, at 1829 & n.142 (seeing the beyond-a-reasonable-doubt standard as imposing a double requirement).

73. On justifying what some may still consider a low threshold, $\text{Bel}(S) > .50$, see Larry Laudan & Harry D. Saunders, *Re-thinking the Criminal Standard of Proof: Seeking Consensus About the Utilities of Trial Outcomes*, 7 INT’L COMMENT. ON EVIDENCE iss. 2, art. 1 (2009), at 3, 14–17, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1369996 [<http://perma.cc/N4MN-J8UR>] (analyzing burdens of proof by comparing the costs and benefits of convictions and acquittals that may or may not be accurate).

“No reasonable doubt” means that no reasonable person could hold $\text{Bel}(\text{not}S) > 0$. On the view that anything is possible, zero as a coarsely gradated degree of belief equates to a *slightest possibility*. $\text{Bel}(\text{not}S) > 0$ thus refers to a step up from the slightest possibility of innocence. Consequently, that no reasonable person could hold $\text{Bel}(\text{not}S) > 0$ actually means that no reasonable factfinder should see a *reasonable possibility* of innocence. In other words, for a conviction the prosecutor must show that no reasonable possibility of innocence exists.⁷⁴

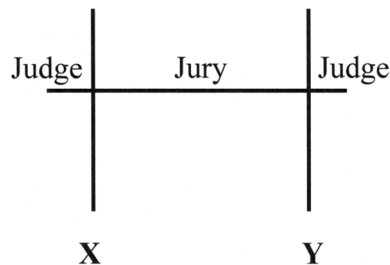
III. BURDEN OF PRODUCTION

This Part will use belief function theory to explain why the traditional view of the law’s burden of production tends slightly to mislead. Then, this Part will demonstrate why the law’s initial burden of production starts the factfinders at point zero.

A. Traditional View

In going from discussing the burden of persuasion to explaining the academic view of the burden of production, I need to use a different diagram. This one represents the role of the judge in patrolling the extreme outer limits of rationality on the jury’s task of applying the standard of proof.

Imagine a single disputed issue of typical fact, S, on which the plaintiff bears the initial burden of production and the burden of persuasion. Then imagine a grid representing the judge’s disagreement with a potential verdict for the plaintiff, or equivalently the judge’s view of likelihood of error in such a verdict, with disagreement or likelihood decreasing from one on the left to zero on the right.⁷⁵ It is important to



74. See Allen & Leiter, *supra* note 20, at 1528 (saying that the prosecution must “show that there is no plausible account of innocence”); McBaine, *supra* note 67, at 266 (proposing an instruction to the effect that a reasonable doubt exists when “you cannot honestly say that it is almost certain that the defendant did the acts which he is charged to have done”); *cf.* Laudan, *supra* note 8, at 300–02 (discussing attempts to append such a notion to the approach of inference to the best explanation).

75. See 9 JOHN H. WIGMORE, EVIDENCE § 2487 (James H. Chadbourne ed. 1981) (providing a diagram representing the interactions between judicial rulings

realize that this diagram represents the likelihood of jury error in finding that the disputed fact exists, not the judge's view of the evidential likelihood that the disputed fact exists. In other words, this diagram represents the judge's thought process in externally overseeing the jury that acts as factfinder, not the judge's thought process as if the judge were finding facts. Alternatively stated, this diagram represents the burden of production, not the burden of persuasion.

The plaintiff in the imagined case starts at the left of the diagram. If he presents no evidence, the judge would ordinarily grant a motion for judgment as a matter of law against him. He is consequently bound to go forward with his evidence until he satisfies the judge that a reasonable jury would be warranted in finding for him. That is, he must get to line X in order to make a jury question of the imagined single issue of fact, doing so by presenting evidence. The plaintiff's getting to or beyond line X means that although the judge might still disagree with a verdict for the plaintiff, the judge thinks a reasonable jury could find that the plaintiff sustained his persuasion-burden, and therefore the judge will hold that the plaintiff sustained his production-burden. If the plaintiff does not get to line X, the judge would so vehemently disagree with a verdict for the plaintiff as to consider the jury irrational, and so the judge can grant the motion for judgment as a matter of law. Line X, again, represents the judge's view on the limit of rationality in the jury's finding for the plaintiff, rather than the judge's view of the evidential likelihood that the disputed fact exists. For example, if the judge disbelieved all of the plaintiff's abundant evidence, but still acknowledged that a reasonable jury could believe it, then the judge should rule that the plaintiff has carried his production-burden, because a reasonable jury could conclude that the plaintiff sustained his persuasion-burden.

If the plaintiff produces enough evidence to get beyond point Y, he is entitled to judgment as a matter of law in his favor unless the defendant comes forward with enough evidence to push the case back to point Y. If the defendant so succeeds in carrying her burden of production, it is again a case for the jury. She may, however, be so successful that her evidence carries the case beyond point X. If so, the defendant becomes entitled to judgment as a matter of law unless the plaintiff in his turn comes forward with more evidence. If, at the close of all the evidence, the case lies between points X and Y, it goes to the jury and the plaintiff has the persuasion-burden. He will lose if the jury is not persuaded.

As a theoretical matter, the production-burden may thus shift several times with the pull and haul of the evidence. As a practical matter, however, such multiple shifting on a single issue of fact is very unlikely.

and evidentiary burdens); cf. John T. McNaughton, *Burden of Production of Evidence: A Function of a Burden of Persuasion*, 68 HARV. L. REV. 1382 (1955) (offering alternative diagrams).

The reason is that conflicting evidence on a single issue would, in most realistic settings, remain in the realm where decision is properly for the jury: a reasonable jury could find either way, and so the judge should not grant judgment as a matter of law. Thus the pull and haul of the evidence will result only in oscillation within the jury's realm.

This diagrammatic scheme works pretty well to represent the law's approach. Moreover, the diagram helps in understanding other concepts and special rules. A *permissive inference* (and *res ipsa loquitur* is one in the view of most courts⁷⁶) describes an inference that a jury is authorized but not required to draw from certain evidence; in other words, the inference satisfies the plaintiff's production-burden by getting the case to line X, although not beyond line Y. A true *presumption* (such as the presumption against suicide as the cause of death) shifts the burden of production to the opponent after the introduction of the evidential premise; in other words, the presumption puts the case to the right of line Y and so requires the jury to find the presumed fact, unless the opponent introduces enough evidence to carry her production-burden and push the case at least back into the jury zone between Y and X.⁷⁷

Most significant among special rules, certain kinds of evidence will not satisfy a burden of production. To satisfy that burden, the burdened party cannot rely on the opponent's failure to testify,⁷⁸ on mere disbelief of the opposing testimony,⁷⁹ or on demeanor evidence drawn from the opponent's testimony.⁸⁰ Similarly, naked statistical evidence normally will not satisfy the burden of production.⁸¹ However, any of these kinds

76. See John Farley Thorne III, Comment, *Mathematics, Fuzzy Negligence, and the Logic of Res Ipsa Loquitur*, 75 Nw. U. L. REV. 147 (1980) (justifying the doctrine by use of fuzzy logic).

77. See FED. R. EVID. 301 ("In a civil case, unless a federal statute or these rules provide otherwise, the party against whom a presumption is directed has the burden of producing evidence to rebut the presumption. But this rule does not shift the burden of persuasion, which remains on the party who had it originally.").

78. See *Stimpson v. Hunter*, 125 N.E. 155, 157 (Mass. 1919) ("The failure of the defendant and of his son to testify although present in court was not equivalent to affirmative proof of facts necessary to maintain the action.").

79. See *Cruzan v. N.Y. Cent. & Hudson River R.R. Co.*, 116 N.E. 879, 880 (Mass. 1917) ("Mere disbelief of denials of facts which must be proved is not the equivalent of affirmative evidence in support of those facts.").

80. See *Dyer v. MacDougall*, 201 F.2d 265, 268-69 (2d Cir. 1952) (holding that although demeanor evidence is probative, it does not suffice to escape a summary judgment).

81. See *Guenther v. Armstrong Rubber Co.*, 406 F.2d 1315, 1318 (3d Cir. 1969) (dictum) (saying, in a case where the plaintiff had been injured by an exploding tire, that a seventy-five to eighty percent chance it came from the defendant manufacturer was not enough for the case to go to the jury).

of evidence is perfectly proper to introduce as a supplement to positive evidence that satisfies the burden of production.⁸² The idea behind these special rules is that they are necessary to protect the notion of a burden of production.

Why protect the burden of production? The reason is that it serves important functions. It facilitates early termination of weak claims or defenses, safeguards against irrational error, and effectuates other process and outcome values.⁸³ In the absence of these special rules, any burdened party could produce enough evidence to reach the factfinder, this evidence possibly being merely in the form of silence, disbelief, demeanor, or general statistics (such as that the defendant manufactured sixty percent of the supply of the injury-causing device of unknown provenance). Perhaps we harbor a special fear of the factfinder's mishandling of such weak evidence when undiluted by other admitted evidence and consequently rendering an unreasoned decision for the proponent based either on prejudice without regard to the evidence or on undue deference to such bewildering evidence. To avoid such an outcome, and to ensure that the burden of production means something, the judge should require sufficient evidence of other kinds. Once the proponent clears that hurdle, the tribunal should allow the feared evidence its probative effect.

B. Reformulated View

At least at first glance, the accepted scheme seems fairly compatible with traditional probability. The only diagrammatic qualification would be that representing the judge's view of jury error as a series of fuzzy intervals rather than lines would better capture reality. The leftmost judge zone would correspond to slightest possibility and the rightmost to almost certainty. The jury zone would run from reasonable possibility all the way to high probability, albeit subject to a motion for new trial.

A difficulty for traditional probability in this area is fixing the starting point for factfinding. The probabilist might assume that when you know nothing, the rational starting point is fifty percent.⁸⁴ Indeed, some experimental evidence indicates that lay people do tend to start at fifty

For a more complete consideration of statistical evidence and its ultimately nonparadoxical nature, see FIELD ET AL., *supra* note 6, at 1314–19 (explaining how a factfinder converts statistical evidence into a belief).

82. See, e.g., *Baxter v. Palmigiano*, 425 U.S. 308, 316–20 (1976) (treating failure to testify as supplemental evidence).

83. See Robert S. Summers, *Evaluating and Improving Legal Processes—A Plea for “Process Values,”* 60 CORNELL L. REV. 1 (1974) (discussing generally the importance of “process values”).

84. *Supra* Part I.A.3.

percent.⁸⁵ Then, if the plaintiff offers a feather's weight of evidence, he in theory would thereby carry not only his burden of production but also his burden of persuasion.

The real-life judge, however, hands only defeat to the plaintiff with nothing more than a feather's weight of evidence, and does so by summary means. Why is that? The reason is that the factfinder should start not at fifty percent but at the far left, and to get to X requires more than a feather's weight. The proper representation of lack of proof is zero belief in the plaintiff's position—but also zero belief in the defendant's position. The full range of belief is properly uncommitted.⁸⁶ That insight makes sense of the notion of the burden of production. It also suggests that, in starting at zero belief, the law is proceeding by belief function theory.

Thus, the paradoxical difficulties in applying the burden of production to weak proof dissipate. For an example, imagine a directed verdict motion by a civil defendant in a single-issue case. This example meshes the burden of production with the new view of the preponderance standard. The motion requires the judge to ask if no reasonable jury could find for the burdened plaintiff by viewing $\text{Bel}(S) > \text{Bel}(\text{not}S)$.⁸⁷ At the end of the plaintiff's case, if a maximally reasonable $\text{Bel}(\text{not}S)$ is zero (effectively a slightest possibility), then the inequality requires a minimally reasonable $\text{Bel}(S)$ to exceed zero (effectively a reasonable possibility). That the plaintiff must have established a reasonable possibility is the embodiment of the burden of production, and it is what keeps the plaintiff from surviving with a mere feather's weight

85. See Anne W. Martin & David A. Schum, *Quantifying Burdens of Proof: A Likelihood Ratio Approach*, 27 JURIMETRICS J. 383, 390–93 (1987) (surveying a small sample of students for their odds of guilt used as the prior probability, which turned out to be 1:1 or fifty percent).

86. Cf. LARRY LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW 104–06 (2006) (parsing the presumption of innocence to mean an epistemic blank slate).

87. The reference to a “reasonable” jury reflects the fact that on such a motion the judge is reviewing the jury's hypothesized application of the standard of proof. The judge's standard of decision turns on whether a jury could not reasonably, or rationally, find for the nonmovant. That is, the defendant must show that a verdict for the plaintiff, given the standard of proof, is not reasonably possible. See, e.g., Kevin M. Clermont, *Procedure's Magical Number Three: Psychological Bases for Standards of Decision*, 72 CORNELL L. REV. 1115, 1126–27 (1987) (discussing when a judge should grant a motion for judgment as a matter of law). We can state this standard of review simply in terms of the law's coarsely gradated scale of possibilities and probabilities, without the complications that belief functions impose on the standard of proof. The reason is that we do not expect the judge to retain uncommitted belief in applying a standard of review. The “evidence” for applying the standard is complete. We want from the judge the likelihood of jury error in finding for the plaintiff, with the complement being the likelihood that the jury has authority to find for the plaintiff.

of evidence. An illustrative situation would be where the plaintiff has produced only a little evidence, but it is “pure” evidence that gives the defendant no support.⁸⁸ If a reasonable jury could find for the plaintiff on such proof, the judge should deny the directed verdict motion. If the defendant then produces no effective evidence during the rest of the trial, but moves again for a directed verdict at the end of all the evidence, the judge should deny the motion and the case should go to the jury. The jury, if it were to take the same view of the evidence as the judge hypothesized, could find for the plaintiff—even on such thin evidence.

In sum, once one recognizes that the burden of persuasion is a comparative operation on sometimes thin evidence, the notion of a burden of production becomes a necessary one. The judge must patrol the sufficiency of the evidence to ensure that there is rationally enough to warrant the factfinder’s applying the burden of persuasion. Otherwise the factfinder might irrationally begin at fifty percent or rely on very slim evidence.

IV. OVERVIEW OF STANDARDS OF PROOF

My views, then, are not at all subversive. Overall I merely contend, in accordance with belief functions’ teaching, that the law charges factfinders to form a coarsely gradated degree of belief in the burdened party’s position, while leaving some belief uncommitted in the face of imperfect evidence, and then apply the standard of proof by comparing that belief to their coarsely gradated belief in its negation. Many observers of the legal system would find that contention, putting its slightly new vocabulary to the side, unobjectionable.

A. Compatibility of Reformulated and Current Standards

A reader always entertains the temptation, upon seeing what looks like a plea for reconceptualization, to dismiss it as a pie-in-the-sky academic musing. When the reconceptualization involves the standards of proof, the specialists have the added temptation of dismissing it as another of the common anti-probabilist rants or pro-probabilist paeans. After all, if my view were a sound one, someone would have come up with it before. So I hasten to undercut my contribution by stressing that my ideas are not that new. I am trying little more than to explain what the law has been doing all along.

The easiest way to convey the lack of newness is to refer back to the normalization process that converts beliefs into probabilistic outputs.⁸⁹ That process allows expression consistent with traditional probability theory. Yet, I resist taking that normalization route for ordinary

88. See Liu & Yager, *supra* note 46, at 18–19 (discussing Leibniz’s notions of pure and mixed evidence).

89. See *supra* Part II.B.1.

use. First, converting to additive beliefs loses information and would reintroduce the probabilistic imaging that originally led to the problems and paradoxes of the traditional view. Second, normalization requires quantification of $\text{Bel}(S)$ and $\text{Bel}(\text{not}S)$, a complicated step otherwise unnecessary, and a step that is much more difficult for humans to perform than simply comparing beliefs. Third, I contend that directly comparing $\text{Bel}(S)$ and $\text{Bel}(\text{not}S)$ actually conforms better to the actual law's instructions than normalization does.

Now take a look at practice. Courts sometimes express divergent views of the standard of proof. Some writers have concluded that courts interpret preponderance in one of three ways: (1) "more convincing," which requires the burdened party to tell a better tale than the opponent tells; (2) "more likely than not," which requires a showing of the fact's existence stronger than the showing of its nonexistence; or (3) "really happened," which requires a showing by evidence of what probably transpired outside in the real world.⁹⁰ The reformulated standard would conform to the middle option, rather than either (1) the comparison of relative plausibility theory or (3) the absolute measure of probability theory. The evidence at trial will support S to an extent while supporting $\text{not}S$ to another extent, and the reformulated standard says that the factfinder need only compare these two beliefs.

Consider a couple of classic cases on how option (2) gets applied. In *Livanovitch v. Livanovitch*,⁹¹ the trial court gave the following charge: "If . . . you are more inclined to believe from the evidence that he did so deliver the bonds to the defendant, even though your belief is only *the slightest degree* greater than that he did not, your verdict should be for the plaintiff."⁹² The appellate court said:

The instruction was not erroneous. It was but another way of saying that the slightest preponderance of the evidence in his favor entitled the plaintiff to a verdict All that is required in a civil case of one who has the burden of proof is that he establish his claim by a preponderance of the evidence When the equilibrium of proof is destroyed, and the beam inclines toward him who has the burden, however slightly, he has satisfied the requirement of the law, and is entitled to the verdict. "A bare preponderance is sufficient, though the scales drop but a feather's weight." This rule accords with the practice in this state as

90. See, e.g., J.S. COVINGTON, JR., *THE STRUCTURE OF LEGAL ARGUMENT AND PROOF* 99–100 (2d ed. 2006) (describing the three explanations courts have given of the term "preponderance of evidence").

91. 131 A. 799 (Vt. 1926).

92. *Id.* at 800.

remembered by the justices of this court, and is well supported by the authorities.⁹³

In *Lampe v. Franklin American Trust Co.*,⁹⁴ one of the defendant's contentions was that the note in suit had been altered after it had been signed by the defendant's decedent. The trial court refused the defendant's request for an instruction that the jury should find that the instrument was not the decedent's note "if you find and believe that it is *more probable* that such changes or alterations have been made in the instrument after it was signed by the deceased and without his knowledge and consent, than it is that such alterations and changes were made at or about the time that the deceased signed the instrument and under his direction and with his knowledge and consent."⁹⁵ Holding the denial of that instruction to have been proper, the appellate court said:

The trouble with this statement is that a verdict must be based upon what the jury finds to be facts rather than what they find to be "more probable." . . . This means merely that the party, who has the burden of proof, must produce evidence, tending to show the truth of those facts, "which is more convincing to them as worthy of belief than that which is offered in opposition thereto."⁹⁶

These two cases' formulations sound contradictory. But if one interprets the quotations as speaking in terms of the coarsely gradated belief in the fact compared with the coarsely gradated belief in the fact's negation, based on the evidence presented, the apparent contradiction evaporates. They both seem to be saying that the burdened party should win if and only if $\text{Bel}(S) > \text{Bel}(\text{not}S)$.

In the end, I submit that comparison of coarsely gradated beliefs is the most accurate representation of what the law tells a factfinder to do with a standard of proof. In civil cases, when the judge explains preponderance of the evidence, the explanation should convey the idea that the factfinder has to find that $\text{Bel}(S)$ is more likely than *not*, which means $\text{Bel}(S) > \text{Bel}(\text{not}S)$.

To say it differently: if you were to put the evidence favorable to plaintiff and the evidence favorable to defendant on opposite sides of the scales, plaintiff would have to make the scales tip somewhat

93. *Id.*

94. 96 S.W.2d 710 (Mo. 1936).

95. *Id.* at 723.

96. *Id.* (quoting *Rouchene v. Gamble Constr. Co.*, 89 S.W.2d 58, 63 (Mo. 1935)).

on his side. If plaintiff fails to meet this burden, the verdict must be for defendant.⁹⁷

Or, preponderance means that the evidence “produces in your minds belief that what is sought to be proved is more likely true than not true”⁹⁸ or “more probably true than false.”⁹⁹ By literally instructing factfinders to decide between S and notS, already the law effectively urges them to focus on belief and disbelief, and then compare them.

B. Application of Reformulated Standards to Multiple Elements

I have written extensively on how to conjoin findings on multiple elements.¹⁰⁰ Using fuzzy logic, I showed how using the product rule to multiply probabilities is improper for factfinding, and thus resolved the so-called conjunction paradox (which posits that if the factfinder proceeds element-by-element and the cause of action entails more than one element, no assurance exists that the product of the elements’ likelihoods will meet the standard of proof).¹⁰¹ The conjunction and disjunction functions work this way for sets in fuzzy logic, when x and y can take any truth value from zero to one and where these two rules are called the MIN and MAX rules:

$$\text{truth}(x \text{ AND } y) = \text{minimum}(\text{truth}(x), \text{truth}(y))$$

$$\text{truth}(x \text{ OR } y) = \text{maximum}(\text{truth}(x), \text{truth}(y))$$

One big reason not to apply the product rule in conjoining elements is the mathematical fact that the product rule works only in an additive regime where the convincingness of a proposition implies the complementary likelihood of its contradiction. Given scarce information or conflicting evidence, however, a factfinder’s belief in x does not imply anything about a belief in not x , other than that the contradiction cannot be more likely than x ’s complement, or one minus the belief in x . Even when factfinding entails a yes-or-no issue, the complement represents not contradiction but only what is not known. One would need

97. MODEL CIVIL JURY INSTRUCTIONS FOR THE DISTRICT COURTS OF THE THIRD CIRCUIT ¶ 1.10 (2015).

98. 3 O’MALLEY ET AL., *supra* note 23, § 104.01; see 4 LEONARD B. SAND ET AL., MODERN FEDERAL JURY INSTRUCTIONS: CIVIL ¶ 73-2 (2015) (“To establish a fact by a preponderance of evidence means to prove that the fact is more likely true than not true.”).

99. Nissho-Iwai Co. v. M/T Stolt Lion, 719 F.2d 34, 38 (2d Cir. 1983) (“The term ‘preponderance’ means that ‘upon all the evidence . . . the facts asserted by the plaintiff are more probably true than false.’” (quoting Porter v. Am. Exp. Lines, Inc., 387 F.2d 409, 411 (3d Cir. 1968))).

100. See *supra* note 6 (citing two articles that the author wrote on the topic of how to conjoin findings on multiple elements).

101. *Id.*

proof or inference of the contradiction before generating any belief in it. One forms a belief in a proposition as a partial truth, while leaving a lot of belief uncommitted. Because the partial truth does not measure the truth of its contradiction, one should not account for the contradiction in combining partial truths. Instead, one should combine beliefs by stringing them together into a chain, with the conjunction of these elements being as true as the weakest link.

Under the MIN rule, if and only if each element passes the standard of proof, the conjoined elements meet the standard of proof. The conjoined story of liability will not only be the most believable story, but will be more believable than all the stories of nonliability combined. To minimize error costs in these circumstances, the law should decide in conformity with the stronger belief. If the plaintiff so satisfies the standard, giving the plaintiff a recovery and the defendant a loss would be economically efficient. Refusing to accept the MIN rule's version of the overall truth will always involve choosing a lesser truth at some step in telling the combined story of a series of two, or more, elements.

Suppose that someone has seriously injured Katie, in circumstances suggesting negligence. She sues Tom, which means that she must prove his identity as the tortfeasor—as well as fault, causation, and injury. She introduces a fair amount of evidence. First, the factfinder would assess that evidence and might conclude as follows: (1) The evidence points to Tom being the perpetrator. If the factfinder were to speak in terms comprehensible to a bettor, he would put the odds at 3:2, or sixty percent. Using words, he would say that Tom was probably the perpetrator. (2) The question of fault was a tough one. There are uncertainties as to what was done, but there is also a vagueness concerning how wrongful the supposed acts really were. The factfinder needs commensurable measures, so that he can evaluate a mix of random and nonrandom uncertainty. If forced to assess all the evidence on this issue and put it on a scale of truth running from zero to one, he would say .7. He might feel more comfortable saying fault was probable or more. (3) The acts, whatever they were, apparently caused the injury. Proximate cause is about as vague and multivalent as a legal concept can get. The factfinder is pretty convinced nevertheless. He would put causation at .8, or highly probable. (4) Katie's injuries are not really very vague or uncertain. He would put this element of the tort at a .9 probability, or almost beyond a reasonable doubt.

The factfinder may want to combine these findings. They are a mixture of probabilities and degrees of truth. But viewing them all as degrees of truth invokes the MIN operator, so that he can say that Katie's story comes in at .6, or probable. Katie should win, by use of fuzzy logic.

However, I am contending here that the better approach would be to look at the four findings through the lens of belief functions. The four findings should be thought of and stated as degrees of belief, which would be markedly lower than the probabilities that do not leave

uncommitted belief. Recall that factfinders sense their beliefs and disbeliefs based on the available evidence; they might believe a fact more than they disbelieve it, even if they would not be willing to bet on it as more likely than not if the truth could somehow be revealed; and the belief might be quite weak as it rests only on what evidence is available, while the bet must commit total belief to either yes or no. Note well that belief functions do not require quantification of beliefs and disbeliefs. Nonetheless, for the purposes of discussion, let us say that the hypothetical's belief functions work out this way: (1) The belief that Tom was the perpetrator is .35, while .25 is the disbelief. (2) As to fault, Bel(fault) might be .50, while .20 is Bel(nofault). (3) Causation is clearer, with Bel(cause) being .70, while .15 is Bel(nocause). (4) Injury is clearer still, with Bel(injury) being .85, while .05 is Bel(noinjury). Katie should win by an element-by-element application of the standard of proof to the beliefs and disbeliefs.

What if the factfinder were to apply the standard of proof to Katie's combined story? How does belief function theory combine the beliefs? It would be disquieting if switching from the theory of fuzzy logic to belief functions produced a different resolution to the conjunction paradox. Happily, the resolution remains much the same. Fuzzy logic and belief functions are fundamentally compatible, being two similar ways to account for uncertainty.¹⁰²

One way to show the similarity of their approaches to conjunction would be to convert the beliefs into fuzzy findings at the decisional stage. The law forces a binary decision. Theorists separate out the credal stage, where the factfinder works with beliefs, from the pignistic stage, where the factfinder must make a decision.¹⁰³ I have argued that the decision comes by comparing Bel(S) to Bel(notS) for each element. Instead, the beliefs could be normalized, and the decision would turn

102. See Didier Dubois & Henri Prade, *A Set-Theoretic View of Belief Functions: Logical Operations and Approximations by Fuzzy Sets*, in CLASSIC WORKS OF THE DEMPSTER-SHAFFER THEORY OF BELIEF FUNCTIONS 375 (Ronald R. Yager & Liping Liu eds., 2008) (arguing for the basic compatibility of fuzzy logic and belief functions); see also SCHUM, *supra* note 40, at 266–69 (observing that one can fuzzify belief functions); John Yen, *Generalizing the Dempster-Shafer Theory to Fuzzy Sets*, in CLASSIC WORKS OF THE DEMPSTER-SHAFFER THEORY OF BELIEF FUNCTIONS 529 (Ronald R. Yager & Liping Liu eds., 2008) (showing how to form beliefs about membership in a fuzzy set).

103. See Philippe Smets & Robert Kennes, *The Transferable Belief Model*, in CLASSIC WORKS OF THE DEMPSTER-SHAFFER THEORY OF BELIEF FUNCTIONS 693, 703–11 (Ronald R. Yager & Liping Liu eds., 2008); cf. Michael S. Pardo, *The Nature and Purpose of Evidence Theory*, 66 VAND. L. REV. 547 (2013) (calling these two stages the micro-level and the macro-level of proof); *supra* note 33 (separating evidence-processing from application of the standard of proof).

on whether the normalized $b(S)$ exceeds $b(\text{not}S)$.¹⁰⁴ To get the conjoined $b(\text{liable})$, the $b(S)$ for each element would combine by the MIN operator. Here, $b(\text{liable})$ is .58. The conjoined $b(\text{liable})$ will meet the standard of proof if and only if the $b(S)$ for each element met the standard of proof.

But how does one more technically conjoin belief functions? The heavy theoretical work on belief functions consists mainly of developing tools for combining pieces of evidence to determine a combined belief. For the most part, the prominent Dempster-Shafer rule governs the task.¹⁰⁵ That rule is very complicated, because it abstractly addresses the problem in the most general terms possible (Bayes' theorem turns out to be a special case of that approach).¹⁰⁶

Conjoining findings and disjoining findings on elements are simpler, however, than the more general problem of updating beliefs as more evidence comes in. Further, shifting the image from mathematical formulas to set theory makes the solution still easier to picture.¹⁰⁷ The situation is that the plaintiff must prove the conjunction of elements; and the defendant gets to rely on the disjunction, winning if the fact-finder disbelieves any element as much as or more than it believes the element. Combined beliefs and disbeliefs on multiple elements appear as a new belief function. At a rough and ready level, the lower probability will be the minimum of the conjoined affirmative beliefs and the upper probability will be the maximum of the disjoined disbeliefs.¹⁰⁸ On our hypothetical, $\text{Bel}(\text{liable})=.35$ and $\text{Bel}(\text{notliable})=.25$.

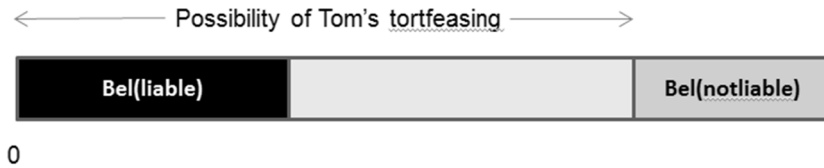
104. See Nicholas J.J. Smith, *Degree of Belief Is Expected Truth Value*, in CUTS AND CLOUDS: VAGUENESS, ITS NATURE, AND ITS LOGIC 491, 503 (Richard Dietz & Sebastiano Moruzzi eds., 2010) (describing a betting scheme based on such comparison).

105. See SHAFER, *supra* note 34, at 6, 57–67 (using orthogonal sums); Barnett, *supra* note 38, at 198–204. By the Dempster-Shafer rule, “we construct a belief function to represent the new evidence and combine it with our ‘prior’ belief function—i.e., with the belief function that represents our prior opinions. This method deals symmetrically with the new evidence and the old evidence on which our prior opinions are based: both bodies of evidence are represented by belief functions, and the result of the combination does not depend on which evidence is the old and which is the new.” SHAFER, *supra* note 34, at 25.

106. See Glenn Shafer & Amos Tversky, *Languages and Designs for Probability Judgment*, in CLASSIC WORKS OF THE DEMPSTER-SHAFER THEORY OF BELIEF FUNCTIONS 345 (Ronald R. Yager & Liping Liu eds., 2008) (comparing Bayesian probability judgments and belief functions).

107. See Huber, *supra* note 26, at 10–15 (discussing the use of possibility theory for this purpose).

108. See *id.* at 14. He suggests that the belief measure, or the necessity N , for conjunction of beliefs is $N(A \cap B) = \min\{N(A), N(B)\}$, and for disjunction of disbeliefs is $N(\text{not}A \cup \text{not}B) = \max\{N(\text{not}A), N(\text{not}B)\}$. But to make the beliefs and disbeliefs fully comparable is more complicated. See Rott,



Normally, if and only if the belief in each element meets the standard of proof, then the belief in the conjoined elements will too. That is, Bel(liable) would meet the standard of proof as compared to Bel(notliable)—but not necessarily. A fairly strong but insufficient disbelief on one element might be bigger than the sufficient belief on another element¹⁰⁹ This complication results from belief functions' taking into account the uncertainty produced by imperfect evidence.

Nevertheless, the law would not want to, and does not, charge its factfinders to perform this difficult mental task of comparing conjunction and disjunction. Even with the rough and ready formula, it is challenging to gauge the overall disbelief, which is a belief in a disjunction. We can ask the factfinder about its disbelief of a single fact, but disjunctive disbelief in a series of facts is difficult even to verbalize. Moreover, the comparison on the basis of the whole case might involve comparing a belief in one element to the disbelief of a different element, which is apt to stymie any factfinder. The law's element-by-element method is more comprehensible (and corraling) than any holistic method, and it works out to be largely equivalent to the proper but difficult holistic method.

The wisdom of the law in proceeding element-by-element appears even more obviously when one considers the other consequences. First, because applying the standard of proof element-by-element comfortably works out to be largely equivalent to applying the standard to the whole story properly conjoined, we do not have to worry much about renegade factfinders who construct an overall story. If the factfinder in actual practice approaches the case holistically, that practice would not directly endanger the standard of proof. Second, given that equivalence, general and special verdicts will work the same way. A nonequivalent holistic standard will run aground upon encountering a nongeneral verdict (or a systematic judge in a bench trial). Third, the equivalence causes the apparent criticality of the number and scope of elements to

supra note 50, at 310 (describing the “tension between degrees for beliefs and degrees for disbeliefs”).

109. For example, if Bel(A)=.50 and Bel(notA)=.40, and if Bel(B)=.30 and Bel(notB)=.20, then Bel(A AND B)=.30 and Bel(notA OR notB)=.40. Thus, the element-by-element approach would produce a result different from the holistic approach. But this situation would not be common. Under a set of coherent beliefs and disbeliefs, if B is less likely than A, then notA should normally be less likely than notB. *Cf.* Rott, *supra* note 50, at 311 (setting this relationship as an axiom).

melt away. “Element” would therefore best be seen as a synonym for a finding necessary to a cause of action or defense under the substantive law, one that the factfinder (or factfinders) must find to meet the standard of proof. In sum, the law is wise to tell its factfinders to proceed element-by-element.

At the least, it is clear that the product rule does not apply to conjoining beliefs. The comparison of stories, as relative plausibility calls for, is not so clearly inappropriate. But if one views the stories through the lens of belief functions, the factfinder ought to believe the plaintiff’s story as much as the minimum belief among the elements, and the defendant’s best story as much as the minimum belief among that story’s elements. That is, the factfinder ought to gauge the convincingness of any story by its weakest link. The defendant’s best story might have a weak link comparable to the plaintiff’s weakest link, even though disbelief in some other element is stronger than the plaintiff’s corresponding link. Thus, comparing plaintiff’s story to the defendant’s best story while focusing on their weakest links seriously stacks the deck against the defendant, effectively denying the benefits of disjunction to the defendant. This analysis explains why belief function theory says that we should instead look to the maximum disbelief among in the defendant’s links, and also why the law says to proceed element-by-element.

All these insights about conjunction establish the single biggest advantage of belief functions over relative plausibility. Belief functions prove that the conjunction paradox does not exist. Relative plausibility theorists, having built their theory on the faulty assumption of bivalence, are terribly troubled by conjunction. In fact, they created their whole theory to sidestep that paradox.¹¹⁰ If there is no conjunction paradox, we need not invent a biased holistic standard and so overturn established law in order to suppress the paradox.

The superiority of belief functions, thanks to their mathematical sophistication, also resolves the aforementioned five big problems of relative plausibility.¹¹¹ First, belief functions allow formulation of standards other than preponderance of the evidence. Second, they track well the instructions currently given to factfinders by the law. Third, they conform to the law by not requiring the nonburdened party to forward a competing story. Fourth, they come closest to minimizing expected

110. See, e.g., Allen & Jehl, *supra* note 12, at 896 (describing two scholars’ attempts to “explain away the proof paradoxes”). Indeed, the motivating force of much of the latest theorizing about proof is the modern concern with the conjunction paradox, theorizing that tends to collapse along with that paradox. See, e.g., Jason Iuliano, Essay, *Jury Voting Paradoxes*, 113 MICH. L. REV. 405 (2014) (arguing for a method of avoiding the conjunction paradox in the jury decision-making process).

111. See *supra* text accompanying notes 19–29 (listing five major problems with relative plausibility).

error costs. Fifth, they can accommodate any psychological theory on the evidence-processing that precedes applying the standard of proof.

CONCLUSION

This Article deployed belief functions to conceptualize the standards of proof. It was not a heavily prescriptive endeavor, which would have tried to argue normatively for the best way to apply standards. Instead, it was mainly a descriptive and explanatory endeavor, trying to unearth how standards of proof actually work in the law world. Compared to the traditionally probabilistic account and the newer conjecture of relative plausibility, this conceptualization conforms more closely to what we know of people's cognition, captures better what the law says its standards are and how it manipulates them, and improves our mental image of the factfinders' task.

One virtue of the conceptualization is that it is not radically new, as it confirms the law's ancient message that factfinders should simply compare their nonquantified views of the fact's truth and falsity. It leaves the law's standards essentially intact to accomplish their current purposes. Another virtue is that the conceptualization nevertheless manages to resolve some stubborn problems of proof: the theory implies that the factfinders should start the case, being in a state of ignorance with lack of proof, at a zero belief; and it also implies that the factfinders at the end of a case should apply the standard only to each separate element. Thus, for understanding the standards of proof, degrees of belief work well.