



## Alabama Law Scholarly Commons

---

Working Papers

Faculty Scholarship

---

12-11-2018

### The Paradoxes of Legal Proof: A Critical Guide

Michael S. Pardo

University of Alabama - School of Law, mpardo@law.ua.edu

Follow this and additional works at: [https://scholarship.law.ua.edu/fac\\_working\\_papers](https://scholarship.law.ua.edu/fac_working_papers)

---

#### Recommended Citation

Michael S. Pardo, *The Paradoxes of Legal Proof: A Critical Guide*, (2018).

Available at: [https://scholarship.law.ua.edu/fac\\_working\\_papers/307](https://scholarship.law.ua.edu/fac_working_papers/307)

This Working Paper is brought to you for free and open access by the Faculty Scholarship at Alabama Law Scholarly Commons. It has been accepted for inclusion in Working Papers by an authorized administrator of Alabama Law Scholarly Commons.

---

## THE PARADOXES OF LEGAL PROOF: A CRITICAL GUIDE

MICHAEL S. PARDO\*

### ABSTRACT

*For several decades, a variety of paradoxes have fascinated and frustrated legal scholars and courts discussing evidence, procedure, and legal proof. These paradoxes concern issues such as statistical evidence, burdens and standards of proof, and rules for jury verdicts. As with other types of paradoxes, the paradoxes of legal proof raise fundamental issues and assumptions. In the legal context, the issues and assumptions are not merely of theoretical interest; they also have significant practical implications at trial and, indeed, throughout the processes of civil and criminal litigation. At the same time, there remains very little agreement about any of the paradoxes, and the issues they raise remain highly contested.*

*This Article explores the different types of legal-proof paradoxes and the connections between them. In analyzing the structure of the paradoxes, the Article makes three contributions. First, it explains the practical significance of the paradoxes and why they raise fundamental issues for evidence law and civil and criminal procedure. Second, it reveals the necessary connections between the different types of paradoxes and the ways in which one type of paradox has implications for the others. These unexplored connections help to explain why individual paradoxes have resisted consensus and continue to prompt discussion and disagreement. The paradoxes are too often treated as isolated problems to be “solved” without appreciating how the underlying issues relate to those raised by the other paradoxes. Third, the analysis clarifies the primary source of confusion for each type of paradox: namely, the popular but mistaken assumption that standards of proof are probabilistic thresholds. Abandoning this spurious assumption provides a number of salutary theoretical and practical consequences, including greater clarity of the legal issues underlying the paradoxes, a better explanation of legal doctrine, and an improved understanding of the manifold litigation issues dependent on the legal-proof process.*

---

\* Henry Upton Sims Professor of Law, Culverhouse School of Law, The University of Alabama. For helpful comments on previous drafts, my thanks to Ron Allen, Kevin Clermont, Shahar Dillbary, Meredith Render, Fred Vars, and participants at an Alabama faculty workshop. My thanks to Dean Mark Brandon and the Alabama Law School Foundation for generous research support.

## CONTENTS

INTRODUCTION .....	235
I. THE PROCESS OF PROOF.....	241
A. <i>Central Features</i> .....	241
B. <i>Two Theoretical Accounts</i> .....	245
1. The Probabilistic Account .....	245
2. The Explanatory Account.....	249
II. THE STATISTICAL-EVIDENCE PARADOX .....	253
A. <i>The Examples and Their Assumptions</i> .....	253
B. <i>Proposed Solutions</i> .....	256
C. <i>Statistical vs. Individualized Evidence</i> .....	262
III. THE CONJUNCTION PARADOX.....	266
A. <i>What Is the Conjunction Paradox?</i> .....	267
B. <i>Proposed Solutions</i> .....	270
1. Denying Premise 1 (Elements).....	271
2. Rejecting Premise 1 (Elements) .....	273
3. Alternative Probabilistic Accounts.....	275
4. The Explanatory Account.....	279
C. <i>The Significance of the Conjunction Paradox</i> .....	280
IV. THE VERDICT PARADOX .....	282
CONCLUSION.....	288

“[P]aradoxes raise serious problems. . . . To grapple with them is not merely to engage in an intellectual game, but is to come to grips with key issues.”<sup>1</sup>

#### INTRODUCTION

Paradoxes have played important roles in evidence scholarship.<sup>2</sup> Since their introduction into the legal literature, paradoxes have provided a continuing source of both fascination and frustration—and, if recent publications are any indication, this trend shows no signs of waning.<sup>3</sup> On one hand, the paradoxes

<sup>1</sup> R.M. SAINSBURY, PARADOXES 1 (3d ed. 2009).

<sup>2</sup> See Roger C. Park & Michael J. Saks, *Evidence Scholarship Reconsidered: Results of the Interdisciplinary Turn*, 47 B.C. L. REV. 949, 986-88 (2006) (discussing role of paradoxes in evidence scholarship). There have been numerous influential discussions over the past five decades. See generally L. JONATHAN COHEN, *THE PROBABLE AND THE PROVABLE* (1977) (discussing issues with using probability in judicial process); FREDERICK SCHAUER, *PROFILES, PROBABILITIES, AND STEREOTYPES* (2003) (discussing “Blue Bus” paradox and sufficiency of statistical evidence in judicial process); Ronald J. Allen, *A Reconceptualization of Civil Trials*, 66 B.U. L. REV. 401 (1986) (using paradoxes to explore limits on compatibility between conventional probability theory and conventional views of trials); David Kaye, *The Paradox of the Gatecrasher and Other Stories*, 1979 ARIZ. ST. L.J. 101 (discussing paradoxes to expose misconceptions about burden of proof and value of probability estimates); Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357 (1985) (discussing relationship between probability in factfinding process and public acceptance of jury verdicts); Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in Legal Process*, 84 HARV. L. REV. 1329 (1971) (considering suitability of utilizing mathematical methods in legal process and possible dangers).

<sup>3</sup> See generally DALE A. NANCE, *THE BURDENS OF PROOF: DISCRIMINATORY POWER, WEIGHT OF EVIDENCE, AND TENACITY OF BELIEF* (2016) (exploring nature of burdens of proof through discussion of balancing evidence against opposing evidence and total “weight” of evidence); Edward K. Cheng, *Reconceptualizing the Burden of Proof*, 122 YALE L.J. 1254 (2013) (arguing that conceptualization of preponderance standard as probability ratio eliminates difficulties posed by classical evidentiary paradoxes); Kevin M. Clermont, *Common Sense on Standards of Proof*, 48 SETON HALL L. REV. 1057 (2018) [hereinafter Clermont, *Common Sense*] (examining use of belief functions to eliminate conjunction paradox problem for factfinders); Kevin M. Clermont, *Death of Paradox: The Killer Logic Beneath the Standards of Proof*, 88 NOTRE DAME L. REV. 1061 (2013) [hereinafter Clermont, *Death of Paradox*] (arguing that modern systems of logic, such as fuzzy logic and belief functions, better reflect realities of legal factfinding as opposed to classical logic and probability theory); David Enoch & Talia Fisher, *Sense and “Sensitivity”: Epistemic and Instrumental Approaches to Statistical Evidence*, 67 STAN. L. REV. 557 (2015) (providing potential justification for inferiority of statistical evidence as compared to individualized evidence and implications stemming from this conclusion); Jason Iuliano, *Jury Voting Paradoxes*, 113 MICH. L. REV. 405 (2014) (applying two philosophical paradoxes to jury decision-making and exploring potential problems); David S. Schwartz & Elliott R. Sober, *The Conjunction Problem and the Logic of Jury Findings*, 59 WM. & MARY L. REV. 619 (2017) (arguing conjunction paradox is of little consequence and problems it raises are insufficient to abandon probability theory in adjudicative fact-finding); Mark Spottswood, *Unraveling the Conjunction Paradox*, 15 LAW PROBABILITY & RISK 259 (2016) (critiquing arguments about conjunction paradox and suggesting problems posed by paradox can be

provide stimuli for new research, theories, and analyses of legal issues; on the other hand, when ignored, they provide potential stumbling blocks for discussions of evidence and proof.

This Article provides a comprehensive account of the legal-proof paradoxes with two goals in mind. The first goal is explanatory: to analyze the different types of paradoxes, the issues they raise, and their significance for law. In addition to this groundwork, the second goal is to defend specific positions with regard to the paradoxes and the underlying legal issues.

Before introducing the proof paradoxes, it will be helpful to briefly outline features they share with paradoxes in other fields as well as the context in which they arise in the legal literature. The paradoxes share three features with paradoxes in other fields. First, they are relatively “easy to state and immediately provoke one into trying to ‘solve’ them.”<sup>4</sup> This feature no doubt explains some of their continuing fascination.<sup>5</sup> Second, they arise because there appears to be an inconsistency between accepted assumptions or because accepted assumptions appear to lead to unacceptable conclusions.<sup>6</sup> Third, they “raise serious problems” and force one “to come to grips with key issues.”<sup>7</sup> In the legal context, these issues cut to the foundations of the evidentiary-proof process—and, indeed, to the systems of civil and criminal litigation more generally. These

---

resolved through jury instruction). The evidence paradoxes are also a topic of recent interest in the philosophical literature. *See, e.g.*, SARAH MOSS, *PROBABILISTIC KNOWLEDGE* 216-20 (2018) (examining legal-proof paradoxes to describe difficulties in applying standards of proof); Michael Blome-Tillmann, *Sensitivity, Causality, and Statistical Evidence in Courts of Law*, 4 *THOUGHT* 102, 102-10 (2015) (offering counterexamples for proposed resolution of Gatecrasher paradox); Lara Buchak, *Belief, Credence, and Norms*, 169 *PHIL. STUD.* 285, 289-92 (2014) (using legal-proof paradoxes to discuss difference between belief and credence); Marcello Di Bello, *Trial by Statistics: Is High Probability of Guilt Enough to Convict?*, *MIND* (forthcoming Jan. 2019) (manuscript at 1) (exploring explanations for reluctance to basing verdicts on pure statistical evidence and providing potential solutions); Martin Smith, *When Does Evidence Suffice for Conviction?*, 127 *MIND* 1193, 1193 (2018) (same).

<sup>4</sup> SAINSBURY, *supra* note 1, at 1.

<sup>5</sup> Despite the considerable attention devoted to the legal-proof paradoxes, no consensus has emerged with regard to any of them—as the ample scholarship suggests. *See* sources cited *supra* notes 2-3.

<sup>6</sup> *See* SAINSBURY, *supra* note 1, at 1 (“This is what I understand by a paradox: an apparently unacceptable conclusion derived by apparently acceptable reasoning from apparently acceptable premises.”); ROY SORENSEN, *A BRIEF HISTORY OF THE PARADOX: PHILOSOPHY AND THE LABYRINTHS OF THE MIND*, at xii (2003) (“Paradoxes mark fault lines in our common-sense world.”); Stephen Schiffer, *Skepticism and the Vagaries of Justified Belief*, 119 *PHIL. STUD.* 161, 165 (2004) (“[E]ach premise seems to some degree creditable when considered on its own, but the two together entail a conclusion we’re apt to feel has got to be false.”).

<sup>7</sup> SAINSBURY, *supra* note 1, at 1; *see* George P. Fletcher, *Paradoxes in Legal Thought*, 85 *COLUM. L. REV.* 1263, 1292 (1985) (“One would expect that unless we properly address the questions that lie at the foundation of our legal system, we will generate paradoxes and antinomies.”). On this point, Professor Sainsbury distinguishes paradoxes from related phenomena such as puzzles, riddles, and games. SAINSBURY, *supra* note 1, at 1.

issues include, among others: which cases proceed to trial and which will be terminated pre-trial; burdens and standards of proof at trial; the relevance, probative value, and admissibility of evidence; the nature of verdicts; and which verdicts will be upheld and which will be overturned.<sup>8</sup> Moreover, the deep issues these paradoxes raise affect every application of the law in these areas, not merely cases that resemble the paradoxical examples.

For several decades, evidence scholarship has been broadly interdisciplinary, drawing on insights from the sciences and humanities.<sup>9</sup> Within this broader context, however, one prominent approach has involved the application of formal tools from statistics and probability theory to aspects of evidence, legal rules, or the process of proof.<sup>10</sup> The use of these tools is the context in which the paradoxes arise.

What, then, are the proof paradoxes? There are three general types. The first and most familiar type concerns the hypothetical use of statistical evidence to prove a contested fact. The second type concerns burdens of proof and arises because crimes, civil causes of action, and affirmative defenses are typically defined in terms of discrete elements. The third type concerns juries and the process by which individual votes are aggregated into verdicts. A brief example of each follows.

The first type of paradox concerns the use of statistical evidence. The most familiar of these is “Blue Bus”:

Plaintiff is injured by a bus. Plaintiff, however, cannot identify the color of the bus that hit her. Plaintiff has undisputed evidence that the Blue Bus Co. owns seventy-five percent of the buses in town. Plaintiff sues the Blue Bus Co.<sup>11</sup>

Assuming no other evidence is introduced, is the evidence sufficient to support a finding for Plaintiff? On one hand, there is a plausible argument that the evidence is sufficient given the applicable standard of proof. Plaintiff must prove by a “preponderance of the evidence” (or “more likely than not”) that a bus owned by Blue Bus was responsible. If the preponderance standard is interpreted to mean “a probability of greater than 0.5” (as it often is)<sup>12</sup> then

---

<sup>8</sup> See *infra* Section I.A (explaining legal-proof paradoxes concern three aspects of law: burdens and standards of proof, probative value of evidence, and requirements for verdicts).

<sup>9</sup> See Park & Saks, *supra* note 2, at 949 (asserting that evidence scholarship has become “decidedly interdisciplinary”).

<sup>10</sup> See Richard Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, 66 B.U. L. REV. 439, 441-42 (1986) (tracing origins of use of statistical theory in evidence); Michael S. Pardo, *The Nature and Purpose of Evidence Theory*, 66 VAND. L. REV. 547, 574-96 (2013) (outlining probabilistic conception of evidence and proof).

<sup>11</sup> This famous example is based on dicta in *Smith v. Rapid Transit, Inc.*, 58 N.E.2d 754, 755 (Mass. 1945). See *infra* note 95.

<sup>12</sup> See Cheng, *supra* note 3, at 1254 (“The preponderance standard is conventionally described as an absolute probability threshold of 0.5.”).

evidence that seventy-five percent of the buses are Blue Buses appears to surpass the 0.5 threshold. On the other hand, the overwhelming consensus is that the evidence is *not* sufficient to establish liability and perhaps not even sufficient to get to trial.<sup>13</sup> The tension between these two lines of reasoning (the evidence is sufficient and not sufficient) creates an apparent “paradox,” and there is no shortage of attempts to explain where or why the first line of reasoning goes awry.<sup>14</sup> The Blue Bus example is one of several similar examples in the literature. Underlying these hypothetical (and highly artificial) cases are deep issues, with important practical implications, concerning legal evidence and standards of proof.<sup>15</sup>

The second type of paradox concerns burdens of proof. Burdens of proof apply to *individual elements* of civil causes of action, crimes, and affirmative defenses. In other words, the proof standard (e.g., preponderance, clear and convincing evidence, and beyond a reasonable doubt) applies to each legal element, and factfinders are instructed to apply them in this manner.<sup>16</sup> For example, suppose that trespass is defined with two elements: (1) entering the land in possession of another, and (2) doing so intentionally or knowingly. In a civil case, a plaintiff would have to prove each of these elements by a preponderance of the evidence. If we assume that the preponderance standard means “a probability greater than 0.5,” then the plaintiff will win in a case in which she proves each element to 0.6. This result, however, gives rise to a “conjunction paradox.”<sup>17</sup> Even though the plaintiff wins, the probability of two (independent) events being true is the conjunction or multiplication of their individual probabilities.<sup>18</sup> In other words, under these circumstances, the

---

<sup>13</sup> See Gary L. Wells, *Naked Statistical Evidence of Liability: Is Subjective Probability Enough?*, 62 J. PERSONALITY & SOC. PSYCHOL. 739, 739 (1992) (describing multiple studies strongly suggesting individuals are reluctant to find for plaintiffs relying solely on some types of statistical evidence).

<sup>14</sup> See *infra* Section II.B (exploring potential solutions to statistical-evidence paradox).

<sup>15</sup> See *infra* Part II (describing inconsistency between requirements of applicable legal rules and actual results).

<sup>16</sup> See, e.g., *United States v. Delgado*, 672 F.3d 320, 331 (5th Cir. 2012) (“[D]ue process requires the government to present evidence sufficient to prove each element of a criminal offense beyond a reasonable doubt . . . .”); *Spieler v. Evans Inc.*, No. 94-50339, 1995 WL 29305, at \*1 (5th Cir. Jan. 13, 1995) (“Following a four-day trial, the district court included the following in its jury instructions: The plaintiff must prove by a preponderance of the evidence each of the following [elements] . . . .”).

<sup>17</sup> COHEN, *supra* note 2, at 2 (introducing conjunction paradox into evidence scholarship). An early discussion of the conjunction effect appears in *The Nature of Judicial Proof*. See JEROME MICHAEL & MORTIMER J. ADLER, *THE NATURE OF JUDICIAL PROOF* 140-44 (1931) (discussing how conjunction paradox impacts determinations which must be made beyond a reasonable doubt).

<sup>18</sup> This theorem—also referred to as the “product rule”—states that, if *A* and *B* are independent of each other, then the probability of *A* and *B* both being true is *A* multiplied by *B*. Independence is a simplifying assumption that will often not be the case, though similar

plaintiff's claim as a whole would only be 0.36 probable. And thirty-six percent is much lower than the preponderance threshold (if the latter is 0.5).<sup>19</sup> The apparent tension between legal practice and probabilistic reasoning creates the paradox and raises important issues about burdens and standards of proof. How these issues are resolved has important practical implications for both civil and criminal litigation.<sup>20</sup>

The third type of paradox concerns the requirements for jury verdicts.<sup>21</sup> A variety of legal rules specify what is required for juror votes to constitute a verdict.<sup>22</sup> These "aggregation" rules sometimes generate paradoxical results. For example, suppose a jurisdiction has a non-unanimous voting rule that requires nine out of twelve votes.<sup>23</sup> When a jury is considering a four-element claim or crime, a verdict may arise that every juror rejects. This could arise because the jurors disagree about which elements are not proven. Jurors 1-3 may conclude that the first element is not proven, jurors 4-6 may conclude the second element is not proven, jurors 7-9 the third element, and jurors 10-12 the fourth element. In this circumstance, a defendant will lose even though every single juror thinks the defendant should win. The inconsistency between conclusions on the elements and conclusions about the case as a whole produces a paradox and

---

problems arise when elements are probabilistically dependent. *See infra* text accompanying notes 162-71.

<sup>19</sup> Other variations on the problem are discussed in Part III, *infra*.

<sup>20</sup> *See infra* Section III.C (explaining perception of factfinders as applying probabilistic thresholds to individual elements of claims creates conceptual problem, and presenting solution).

<sup>21</sup> *See* JON ELSTER, SECURITIES AGAINST MISRULE: JURIES, ASSEMBLIES, ELECTIONS 62-64 (2013) (discussing "Poisson paradox" of belief-aggregation and how it impacts jury decisions); Iuliano, *supra* note 3, at 423-26 (explaining that special jury verdicts are rendered unreliable due to problems presented by discursive dilemma paradox and lottery paradox); Michael S. Pardo, *Group Agency and Legal Proof; or, Why the Jury Is an "It,"* 56 WM. & MARY L. REV. 1793, 1793-94 (2015) (conceptualizing juries as group agents and exploring resulting conceptual problems for probabilistic theory of evidence). Scholars have explored similar problems with legal issues decided by multi-member courts. *See* Lewis A. Kornhauser & Lawrence G. Sager, *The Many as One: Integrity and Group Choice in Paradoxical Cases*, 32 PHIL. & PUB. AFF. 249, 250-51 (2004) (explaining how tallying votes of judges on panels can produce different outcomes than tallying their votes as a whole).

<sup>22</sup> There is considerable doctrinal confusion concerning when juror disagreements are acceptable and when they undermine verdicts. *See* Peter Westen & Eric Ow, *Reaching Agreement on When Jurors Must Agree*, 10 NEW CRIM. L. REV. 153, 156 (2007) (detailing division of opinion between Supreme Court justices as to when, if ever, jurors must agree upon which means defendant used to commit offense).

<sup>23</sup> *See infra* Section I.A (describing features of the evidentiary proof process including jury voting requirements). There is wide variation among jurisdictions in terms of jury size and voting rules (unanimity, supermajority, etc.). *See* DAVID B. ROTTMAN & SHAUNA M. STRICKLAND, U.S. DOJ, OFFICE OF JUSTICE PROGRAMS, BUREAU OF JUSTICE STATISTICS, STATE COURT ORGANIZATION 2004, at 233-37 tbl.42 (2006) (detailing different jury sizes depending on state and trial type).



raises important issues about the doctrinal requirements for verdicts. As with the other types of paradoxes, how these issues are resolved has significant practical implications for litigation.<sup>24</sup>

The paradoxes of legal proof share a general structure. What unites them—and, indeed, what makes them “paradoxical”—is that they each reveal tensions or inconsistencies in conventional ways of understanding the process of legal proof.<sup>25</sup> These tensions or inconsistencies call out for an explanation, one that will “solve” the paradox. As a general matter, we can distinguish two different types of solutions: “happy-face” and “unhappy-face.”<sup>26</sup> Happy-face solutions reveal how a paradox can be resolved or explained away.<sup>27</sup> Such a solution will also reveal where and how a line of reasoning goes wrong. It may lead us to abandon an assumed premise or change current practices, or it may allow us to harmonize propositions that appeared to be contradictory. Unhappy-face solutions, by contrast, explain why no happy-face solutions are possible. They reveal that because of some sort of glitch or incoherence in our concepts, rules, or practices, there is an inherent tension that cannot be resolved.<sup>28</sup>

This Article argues that there are, in fact, happy-face solutions to the proof paradoxes (although some will be happier than others). The general lesson is that the most pernicious consequences that appear to follow from the paradoxes arise not from the legal rules and practices themselves—rather, they arise from a particular theoretical account of the proof process in which legal standards of proof are thought of as probabilistic thresholds. This Article demonstrates that, in addition to better explaining the legal rules and practices, an alternative theoretical account of the proof process in terms of competing explanations also explains why the legal system avoids or ameliorates the most pernicious consequences thought to follow from the paradoxes. Moreover, understanding the reasons why will illuminate fundamental aspects of legal evidence and proof, with implications far beyond the particular paradoxical examples.

---

<sup>24</sup> See *infra* Part IV (explaining how juror disagreement on aspects of cases seemingly produce paradoxical outcomes).

<sup>25</sup> See SAINSBURY, *supra* note 1, at 1 (explaining paradoxes as result of unacceptable conclusions arrived at based on acceptable reasoning); SORENSEN, *supra* note 6, at xii (describing paradoxes as tension underlying reasoning).

<sup>26</sup> Schiffer, *supra* note 6, at 178-81 (introducing this distinction); see also DUNCAN PRITCHARD, EPISTEMIC ANGST 192 n.13 (2016) (employing distinction).

<sup>27</sup> See Schiffer, *supra* note 6, at 178-79 (“A happy face solution to a paradox does two things, assuming that the propositions comprising the set really are mutually incompatible: first, it identifies the odd-guy-out, the member of the set that’s not true; and second, it shows us why this spurious proposition deceived us, strips from it its patina of truth, so that we’re not taken in by it again.”).

<sup>28</sup> *Id.* at 179 (“When a paradox lacks a happy-face solution it’s because there’s a certain kind of glitch in the concept, or concepts, generating the paradox. Aspects . . . are in tension, pull us in different directions, and there is nothing else in the concept or elsewhere to resolve that tension . . .”). Schiffer gives “free will” as an example. *Id.*

The Article proceeds in five parts. Part I provides background for the proof paradoxes. This Part first describes the legal rules and practices that comprise the proof process, and then discusses the theoretical accounts that purport to explain or conceptualize these rules and practices. With this background in place, Part II discusses the statistical-evidence paradox, Part III discusses the conjunction paradox, and Part IV discusses the verdict paradox. The Conclusion explains some general lessons to be drawn from the paradoxes and their practical implications.

## I. THE PROCESS OF PROOF

Before turning to the proof paradoxes, it is important to first understand the features of the evidentiary proof process that give rise to the paradoxes. The paradoxes concern three aspects of the law: (1) burdens and standards of proof, (2) the probative value of evidence, and (3) the requirements for verdicts. This Part first outlines these features, and then discusses the theoretical accounts that purport to explain them. Separating the features and the accounts of these features is a critical first step toward understanding the paradoxes. This is because, as the next parts will show, the paradoxes do not necessarily arise from the features themselves but rather from particular ways of *conceptualizing* or *interpreting* the features.

### A. *Central Features*

The process of proof at trial is structured around the burden of proof, on one hand, and the requirements for admissible evidence, on the other. The burden of proof specifies the issues that parties are responsible for proving at trial.<sup>29</sup> For each such issue, an applicable standard instructs factfinders on when to conclude that the issue has been proven. Admissibility rules specify the evidence that factfinders may rely upon in making this determination. In jury trials, additional rules specify when individual juror conclusions constitute a verdict. The discussion below elaborates on the legal rules that underlie these three features: burdens of proof, evidence, and verdicts.

The burden of proof provides an overarching structure to the evidentiary process at trial.<sup>30</sup> For the legal elements of crimes, civil causes of action, or affirmative defenses, one party has the burden of proving each element.<sup>31</sup> Typically, the prosecution and civil plaintiffs bear the burden of proving the elements of crimes and civil claims, respectively.<sup>32</sup> Defendants bear the burden

---

<sup>29</sup> For an overview of burdens of proof, see RONALD J. ALLEN ET AL., *AN ANALYTICAL APPROACH TO EVIDENCE* 804-13 (6th ed. 2016).

<sup>30</sup> *Id.* (explaining roles, rationales, and allocations of burden of production and persuasion at trial).

<sup>31</sup> *Id.* at 808-13 (describing which party typically bears which burden in certain cases).

<sup>32</sup> *Id.*

of proof, in whole or in part, for most affirmative defenses.<sup>33</sup> The burden of proof at trial consists of two components: a burden of *production* and a burden of *persuasion*. The burden of production, as its name suggests, requires the party with that burden to produce evidence. This raises the question: how much evidence is needed to meet the production burden? The answer: enough to meet the persuasion burden.<sup>34</sup> The burden of persuasion is set by the applicable standard of proof, which is both a key structural feature in legal proof and at the heart of many of the proof paradoxes.

Standards of proof specify when the party with the burden of proof has met its burden. In essence, these standards dictate when a disputed fact has been “proven” for legal purposes. In civil cases, the typical standard is “preponderance of the evidence.”<sup>35</sup> In criminal cases, the higher standard of “beyond a reasonable doubt” applies.<sup>36</sup> A third, intermediate standard of proof, “clear and convincing evidence,” occasionally applies to the elements of civil claims and affirmative defenses.<sup>37</sup> The standards serve a number of important functions. Most importantly, they instruct factfinders on when to conclude that a disputed fact has been proven.<sup>38</sup> In addition, they guide a number of judicial

---

<sup>33</sup> *Id.* For example, in criminal cases, defendants may have a burden of production (but not the burden of persuasion) for an affirmative defense, or they may have both burdens. *See, e.g.*, *Clark v. Arizona*, 548 U.S. 735, 769-73 (2006) (discussing burdens of proof for insanity defense).

<sup>34</sup> *See* John T. McNaughton, *Burden of Production of Evidence: A Function of a Burden of Persuasion*, 68 HARV. L. REV. 1382, 1390 (1955) (concluding burden of production is “function of [the] burden of persuasion”).

<sup>35</sup> *Grogan v. Garner*, 498 U.S. 279, 286 (1991) (“Because the preponderance-of-the-evidence standard results in a roughly equal allocation of the risk of error between litigants, we presume that this standard is applicable in civil actions between private litigants . . .”).

<sup>36</sup> *In re Winship*, 397 U.S. 358, 362 (1970) (explaining that “beyond a reasonable doubt” standard is constitutionally required in criminal cases).

<sup>37</sup> *Clark*, 548 U.S. at 796 (finding no issue with Arizona’s requirement to prove insanity by clear and convincing evidence); *Addington v. Texas*, 441 U.S. 418, 424 (1979) (“[T]his Court has used the ‘clear, unequivocal and convincing’ standard of proof to protect particularly important individual interests in various civil cases.”).

<sup>38</sup> Standards of proof instruct legal decision-making on how to resolve factual disputes in the face of *uncertainty*. The standards each express policy choices regarding accuracy and allocating the risk of error among the parties. The “preponderance of the evidence” standard attempts to divide the risk of error roughly evenly between the parties, in part because the two types of errors (i.e., false positives and false negatives) are thought to be similar in terms of social costs. *Grogan*, 498 U.S. at 286 (“[P]reponderance-of-the-evidence standard results in a roughly equal allocation of the risk of error between litigants . . .”). The standard also expresses a notion of equality among litigants. *See* Mike Redmayne, *Standards of Proof in Civil Litigation*, 62 MOD. L. REV. 167, 171-74 (1999); Lawrence B. Solum, *Procedural Justice*, 78 S. CALIF. L. REV. 181, 287-88 (2004). Under this standard, the law favors whichever side the evidence appears to support, with “ties” going against the party with the proof burden. Accordingly, each side bears a risk that the evidence may appear to support the opposing party, even though they ought to win. Moreover, given certain assumptions, the

assessments of the sufficiency of evidence—including whether a party has sufficient evidence to survive a motion for summary judgment, whether a case should go to the jury during a trial, and whether the evidence is sufficient to support a judgment.<sup>39</sup> These determinations require an assessment of whether an outcome is “reasonable” given the evidence and the standard of proof.<sup>40</sup>

Burdens of proof provide only half of the proof picture. The other half concerns the evidence itself. Whether a standard of proof is met in a given case depends on whether a factfinder concludes the standard is satisfied *based on the admissible evidence*.<sup>41</sup> Evidentiary rules structure the admissibility process and the inferences that may be drawn from evidence. The two foundational concepts

---

preponderance standard is thought to maximize overall accuracy (or minimize the total number of errors). See Richard S. Bell, *Decision Theory and Due Process: A Critique of the Supreme Court's Lawmaking for Burdens of Proof*, 78 J. CRIM. L. & CRIMINOLOGY 557, 572 (1987) (stating that policy underlying standard in civil trials is equal distribution of risk of error); Edward K. Cheng & Michael S. Pardo, *Accuracy, Optimality, and the Preponderance Standard*, 14 LAW, PROBABILITY & RISK 193, 194 (2015) (concluding optimal standard of proof for ensuring accuracy of civil litigation verdicts is preponderance of the evidence); 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, 75-81 (2004) (stating preponderance of evidence standard is “highly defensible” and generally minimizes risk of error). Higher standards of proof (“beyond a reasonable doubt” and “clear and convincing evidence”) attempt to shift the risk of error away from one side (typically, defendants) in order to minimize one type of error, under the assumption that one type of error is more socially costly. See *Addington*, 441 U.S. at 424 (explaining “clear and convincing” standard applies when there are asymmetric risks in civil cases); *In re Winship*, 397 U.S. at 364 (discussing asymmetric risks of error in criminal cases); Bell, *supra*, at 580 (discussing impact of “beyond a reasonable doubt” on shifting allocation of risk of errors); Hamer, *supra*, at 86 (discussing burden-shifting effect of “beyond a reasonable doubt” standard).

<sup>39</sup> See FED. R. CIV. P. 50; FED R. CIV. P. 56; FED. R. CRIM. P. 29. The preponderance standard also applies to disputed facts relevant for admissibility determinations. *Bourjaily v. United States*, 483 U.S. 171, 175 (1987). A number of other issues throughout the litigation process also depend on evidentiary considerations (including jurisdiction, venue, pleadings, discovery, class certification, joinder, preclusion, and settlement, among others). See Michael S. Pardo, *Some Remarks on the Importance of Evidence Outside of Trials*, 36 REV. LITIG. 443, 447 (2018) (“Evidence still matters to litigation even in a world with relatively few trials because evidence influences every stage of the litigation process.”).

<sup>40</sup> See *Reeves v. Sanderson Plumbing Prods., Inc.*, 530 U.S. 133, 150 (2000) (discussing standard for judgment as matter of law); *Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242, 252 (1986) (discussing standard for motion for summary judgment); *Jackson v. Virginia*, 443 U.S. 307, 318 (1979) (discussing sufficiency of evidence in criminal cases).

<sup>41</sup> In addition, courts review whether that determination is reasonable in light of the evidence and the standard. See *Anderson*, 477 U.S. at 248, 252 (explaining that summary judgment standard depends on whether “reasonable jury could return a verdict for the nonmoving party” and that this depends on “evidentiary standard of proof”); see also *Reeves*, 530 U.S. at 150 (explaining that standard for judgment as matter of law “mirrors” standard for summary judgment); *Jackson*, 443 U.S. at 319 (explaining that sufficiency standard for criminal cases depends on whether “any rational trier of fact could have found the essential elements of the crime beyond a reasonable doubt”).

underlying these rules are *relevance* and *probative value*. Relevance is a necessary condition for admissibility.<sup>42</sup> Evidence is relevant if it has any tendency to make a material fact more or less probable than without the evidence.<sup>43</sup> Probative value refers to the strength of relevant evidence in proving a disputed fact.<sup>44</sup> Courts have discretion to exclude an item of evidence when its probative value is “substantially outweighed” by its potential for unfair prejudice, confusing the issues, being misleading, or for efficiency considerations.<sup>45</sup> These rules aim to improve the accuracy of fact-finding and reduce the costs of litigation by minimizing inferential errors and incentivizing parties to introduce epistemically better evidence.<sup>46</sup> Within this basic structure, a variety of other rules impose additional requirements on the admissibility and use of evidence.<sup>47</sup>

A third feature of the proof process concerns verdicts. Burdens of proof and admissible evidence together provide a basic picture; for each disputed element, each factfinder must conclude, based on the admissible evidence, whether the applicable standard of proof has been satisfied.<sup>48</sup> Additional legal rules regulate how to aggregate the decisions of individual jurors for purposes of a verdict. Constitutional and procedural rules specify the size of juries and the number of votes needed.<sup>49</sup> In federal criminal cases, juries typically have twelve members

---

<sup>42</sup> See FED. R. EVID. 402.

<sup>43</sup> *Id.* 401.

<sup>44</sup> “Probative value” may refer to the value of an item of evidence by itself or to multiple items in combination. A single item of evidence may have low probative value by itself but significantly more probative value when combined with other evidence, and vice versa. Probative value also depends on the proponent’s need for the evidence. See *id.* 403; *Old Chief v. United States*, 519 U.S. 172, 184 (1997) (“[W]hat counts as the Rule 403 ‘probative value’ of an item of evidence . . . may be calculated by comparing evidentiary alternatives.”).

<sup>45</sup> FED. R. EVID. 403.

<sup>46</sup> See *id.* 102 (“These rules should be construed so as to administer every proceeding fairly, eliminate unjustifiable expense and delay, and promote the development of evidence law, to the end of ascertaining the truth and securing a just determination.”).

<sup>47</sup> These rules serve a variety of different goals. Some rules are related to accurate fact-finding. See, e.g., *id.* 301 (governing allocation of burden of proof for presumptions); *id.* 404 (governing admissibility of character evidence); *id.* 702 (governing testimony of experts); *id.* 801-07 (governing admissibility of hearsay). Other rules are external to the proof process. See, e.g., *id.* 407-11 (describing categories of evidence not admissible to prove liability).

<sup>48</sup> In other words, each factfinder must conclude whether the probative value of the evidence as a whole exceeds the threshold set by the standard of proof. See cases cited *supra* note 41.

<sup>49</sup> See ROTTMAN & STRICKLAND, *supra* note 23, at 233-37 tbl.42 (showing state-specific rules derived from state constitutions). Six-member juries are constitutional, but five-member juries are not. See *Ballew v. Georgia*, 435 U.S. 223, 239 (1978) (finding juries with less than six members unconstitutional); *Williams v. Florida*, 399 U.S. 78, 86 (1970) (holding that six-member juries are constitutional). Twelve-member juries need not be unanimous, but six-member juries must be. *Burch v. Louisiana*, 441 U.S. 130, 134 (1979) (holding that six-member juries must be unanimous); *Apodaca v. Oregon*, 406 U.S. 404, 410-11 (1972)

and the verdicts must be unanimous.<sup>50</sup> State criminal cases typically require between six and twelve members and typically require unanimous verdicts.<sup>51</sup> Oregon, however, allows for non-unanimous verdicts in criminal cases and Louisiana, until recently, did as well.<sup>52</sup> In federal civil cases, juries require at least six members and the verdict must be unanimous (unless the parties stipulate otherwise).<sup>53</sup> There is considerable variation among the states in civil cases.<sup>54</sup> Most juries range from six to twelve members, although some states allow parties to stipulate fewer than six.<sup>55</sup> Some states require unanimous verdicts, some require a supermajority, and others reduce from unanimous to supermajority rule after a period of deliberation has passed.<sup>56</sup>

### B. *Two Theoretical Accounts*

Scholars have provided two general theoretical accounts of legal proof: probabilistic and explanatory.<sup>57</sup> The two approaches provide different ways of explaining the three central features discussed above: standards of proof, the probative value of evidence, and verdicts.

#### 1. The Probabilistic Account

Probability theory has been a highly influential tool among legal scholars attempting to explain the central features of legal proof.<sup>58</sup> The probabilistic

---

(holding that twelve-member juries need not be unanimous).

<sup>50</sup> FED. R. CRIM. P. 23(b) (“A jury consists of twelve persons unless this rule provides otherwise.”); *id.* 31(a) (stating that “verdict must be unanimous”).

<sup>51</sup> See ROTTMAN & STRICKLAND, *supra* note 23, at 233-37 tbl.42 (showing jury size and verdict-unanimity rules by state).

<sup>52</sup> See LA. CODE CRIM. PROC. ANN. art. 782 (2014) (requiring just ten of twelve jurors to concur in cases “in which punishment is necessarily confinement at hard labor”); OR. REV. STAT. § 136.450 (2015) (stating general rule that “verdict of a trial jury in a criminal action shall be by concurrence of at least 10 of 12 jurors”). In November 2018, voters in Louisiana approved an amendment to the state constitution that will require unanimous verdicts in criminal cases (beginning in January 2019). See Official Election Results from Nov. 6, 2018, LOUISIANA SECRETARY OF STATE R. KYLE ARDOIN, <https://voterportal.sos.la.gov/Graphical/#/TabFull> [<https://perma.cc/9EHQ-P3Q6>] (select “Tues Nov 6 2018” in drop down menu; then follow the “Statewide” hyperlink).

<sup>53</sup> FED. R. CIV. P. 48(b).

<sup>54</sup> See ROTTMAN & STRICKLAND, *supra* note 23, at 233-37 tbl.42.

<sup>55</sup> *Id.*

<sup>56</sup> *Id.*

<sup>57</sup> See Ronald J. Allen & Michael S. Pardo, *Relative Plausibility and Its Critics*, INT’L J. EVID. & PROOF (forthcoming 2019) (manuscript at 8-19) (contrasting two approaches in detail).

<sup>58</sup> See, e.g., RICHARD EGGLESTON, EVIDENCE, PROOF & PROBABILITY 6-9 (1978) (outlining chapters discussing role of probability in fact-finding, and decision-making based on evidence and applicable standard of proof); Michael O. Finkelstein & William B. Fairley, *A Bayesian Approach to Identification Evidence*, 83 HARV. L. REV. 489, 498-509 (1970) (proposing

account interprets standards of proof as points or thresholds on a scale between zero and one. Under this scheme, one represents certain truth and zero represents certain falsity,<sup>59</sup> and the standards are commonly defined as follows: “preponderance of the evidence” as a probability of greater than 0.5, “beyond a reasonable doubt” as a probability of 0.9 or greater, and “clear and convincing evidence” as a probability around 0.75.<sup>60</sup> Under this account, factfinders assess the probability of each legal element (along the spectrum between zero and one). An element is considered proven if its probability exceeds the threshold for the applicable standard.<sup>61</sup> The element is not proven when its probability is at or below the threshold. These interpretations are intended to reflect the normative

---

probability theory as useful tool for evaluating probative significance of statistical evidence); John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065, 1066-67, 1071-77, 1083-90 (1968) (describing probability as fundamental concept in decision theory, proposing formula for determining probability necessary to return verdict, and discussing determination of probabilities by factfinders); Richard O. Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021, 1022-27, 1032-37 (1977) (describing utility of probability theory for determining relevance of evidence). There are differences among various probabilistic accounts; the discussion below highlights the common features, focusing on the details that are important for understanding the proof paradoxes. The theory relied upon is the standard account of probability based on the Kolmogorov axioms. A.N. KOLMOGOROV, FOUNDATIONS OF THE THEORY OF PROBABILITY 2-3, 14-16 (Nathan Morrison trans., 2d ed. 1956) (setting forth famous axioms of probability theory); see IAN HACKING, AN INTRODUCTION TO PROBABILITY AND INDUCTIVE LOGIC 140-49 (2001) (providing clear overview of various theories about probability); discussion *infra* Section III.B.3 (discussing alternative probabilistic accounts).

<sup>59</sup> See Buchak, *supra* note 3, at 285 (discussing complex relationship between beliefs and “credences,” or “degrees of belief”).

<sup>60</sup> See *Brown v. Bowen*, 847 F.2d 342, 345-46 (7th Cir. 1988) (“The preponderance standard is a more-likely-than-not rule, under which the trier of fact rules for the plaintiff if it thinks the chance greater than 0.5 that the plaintiff is in the right. The reasonable doubt standard is much higher, perhaps 0.9 or better. The clear-and-convincing standard is somewhere in between.”); Bell, *supra* note 38, at 558 (noting assumption that triers of fact conclude “that the truth of a party’s assertion is probable at some value . . . between 0 and 1”); Hamer, *supra* note 38, at 87 (describing civil standard as “50 per cent” standard and discussing views that criminal standard is 0.9 or higher). Although these numbers reflect common assumptions, there is disagreement about what numbers to attach to each of the standards. See, e.g., *United States v. Fatico*, 458 F. Supp. 388, 409-10 (E.D.N.Y. 1978) (citing studies and survey of judges indicating disagreement regarding which numbers to attach to different standards); Rita James Simon & Linda Mahan, *Quantifying Burdens of Proof: A View from the Bench, the Jury, and the Classroom*, 5 LAW & SOC’Y REV. 319, 319 (1971) (investigating differences in numbers judges, juries, and students attach to civil and criminal standards); Eyal Zamir & Ilana Ritov, *Loss Aversion, Omission Bias, and the Burden of Proof in Civil Litigation*, 41 J. LEGAL STUD. 165, 186-91 (2012) (discussing test results indicating “standard of persuasion implemented by legal fact-finders is considerably higher” than is generally accepted).

<sup>61</sup> See Pardo, *supra* note 10, at 595 (“[W]hether the evidence in a case is sufficient to satisfy a probabilistic proof standard will depend simply on whether the particular fact-finders think that the evidence surpasses the threshold.”).

policy choices underlying the standards.<sup>62</sup> For the preponderance standard, for example, proving an element to a probability of greater than 0.5 may appear to allocate the risk of error roughly evenly between the parties and minimize the total number of errors.<sup>63</sup> By contrast, requiring proof to a higher probabilistic threshold—for example, 0.95 for “beyond a reasonable doubt” or 0.75 for “clear and convincing evidence”—shifts more of the risk of error to the party with the burden of proof.<sup>64</sup>

Under the probabilistic account, the evidence must be quantified in some manner in order to compare it with the probabilistic threshold that represents the standard of proof.<sup>65</sup> In addition to quantifying the evidence as a whole (to measure against the standard) scholars have also sought to define the relevance and probative value of individual items of evidence based on how the evidence affects the probability of the fact for which it is offered.<sup>66</sup> As within probability

<sup>62</sup> See, e.g., Hamer, *supra* note 38, at 73 (“[T]he ordinary civil standard of 50 per cent minimises the subjective expected rate of error while the criminal standard is set at a higher level to reduce the risk of the more costly error—the wrongful conviction.”).

<sup>63</sup> According to this interpretation, the probability space between 0 and 1 is divided roughly evenly, with the plaintiff bearing the risk when the result is between 0 and 0.5 and the defendant bearing the risk when the probability is greater than 0.5. Moreover, accuracy is thought to be advanced by siding with the plaintiff when the claim is more likely true and with the defendant when it is not. *But see* Allen & Pardo, *supra* note 57 (manuscript at 12-13) (explaining why these appearances about accuracy and risk of error are false).

<sup>64</sup> For example, civil plaintiffs, rather than defendants, now bear the risk of error for issues proven greater than a preponderance but less than “clear and convincing” evidence. Moreover, in criminal cases, the prosecution bears the risk of error for issues that might be proven greater than a preponderance, but not “beyond a reasonable doubt.”

<sup>65</sup> The quantification need not require attaching a precise number to the evidence, only a determination whether or not it surpasses the probabilistic threshold.

<sup>66</sup> See FED. R. EVID. 401-03 (governing admissibility and exclusion of evidence based on relevance and probative value). Probabilistic accounts rely on the concept of “likelihood ratios,” which compare two values: (1) the probability of receiving evidence, given that a proposition is true; and (2) the probability of receiving the same evidence, given that the same proposition is false. See D.H. Kaye, *The Relevance of “Matching” DNA: Is the Window Half Open or Half Shut?*, 85 J. CRIM. L. & CRIMINOLOGY 676, 683-84 (1995) [hereinafter Kaye, *The Relevance of “Matching” DNA*] (“The best developed and most plausible theory of probative value articulated by legal scholars builds on a statistical concept known as the likelihood ratio . . .”); Lempert, *supra* note 58, at 1025 (describing likelihood ratios as “mathematical equivalent” of Federal Rule of Evidence 401 definition of “relevant evidence”). Relevance is defined as any evidence with a likelihood ratio other than 1:1. Lempert, *supra* note 58, at 1025 (“Where the likelihood ratio for an item of evidence differs from one, that evidence is *logically relevant*.”). Probative value is defined based on the size of the ratio or the extent to which the evidence changes the prior probability. Compare Richard D. Friedman, *A Close Look at Probative Value*, 66 B.U. L. REV. 733, 733-35 (1986) (presenting mathematical definition of probative value reflecting extent to which new evidence affects prior probability), with D.H. Kaye, *Quantifying Probative Value*, 66 B.U. L. REV. 761, 761-64 (1986) [hereinafter Kaye, *Quantifying Probative Value*] (noting plausible alternatives to Friedman’s mathematical methods for measuring probative value). Bayes’s



theory generally, the quantification of evidence (as a whole or as individual items) may be based on either *objective* or *subjective* considerations.<sup>67</sup> “Objective” bases for probabilistic conclusions include relative frequencies or known statistical distributions—for example, the percentage of people with a particular blood type or the percentage of blue buses in the town. “Subjective” probabilistic conclusions are based on *degrees of belief* or how strongly a person believes a proposition to be true—for example, a juror may conclude that a confession is twenty times more likely if the defendant is guilty than if he is innocent, even if no such data is presented or available.<sup>68</sup> Probabilistic approaches to legal evidence have relied on both objective and subjective conceptions.<sup>69</sup>

Finally, the probabilistic account of verdicts requires the aggregation of jurors’ probabilistic conclusions. Individual jurors decide whether the probabilistic standard for each element has been satisfied.<sup>70</sup> If the requisite number of jurors concludes that each element has been proven, then a verdict

---

Theorem provides a method for updating the prior probability in light of new items of evidence (the posterior probability equals the likelihood ratio multiplied by the prior probability).

<sup>67</sup> See, e.g., Alan Hájek, *Interpretations of Probability*, STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Dec. 19, 2011), <http://plato.stanford.edu/entries/probability-interpret> [https://perma.cc/56JU-BBNS] (providing clear overview of methods probability interpretation); see also, e.g., DONALD GILLIES, *PHILOSOPHICAL THEORIES OF PROBABILITY 1* (2000) (surveying methods of probability interpretation, including objective and subjective methods, used to identify probability based on available evidence); D.H. MELLOR, *PROBABILITY: A PHILOSOPHICAL INTRODUCTION* 66 (2005) (demonstrating objective and subjective considerations in quantifying evidence in context of betting on horses, where available evidence may objectively indicate low chance of horse winning but bettor decides based on subjective hunch horse will win).

<sup>68</sup> See, e.g., LEONARD J. SAVAGE, *THE FOUNDATIONS OF STATISTICS* 27-30 (1954) (setting forth theory of “personal probability” based on subjective degrees of belief).

<sup>69</sup> See, e.g., Friedman, *supra* note 66, at 733 (advocating probabilistic account that takes subjective approach to probability); Alvin I. Goldman, *Quasi-Objective Bayesianism and Legal Evidence*, 42 *JURIMETRICS* 237, 237, 239-41 (2002) (proposing augmented subjective Bayesianism approach called “quasi-objective Bayesianism” which can “identify conditions for objectively expected increases in truth possession”); Kaye, *The Relevance of “Matching” DNA*, *supra* note 66, at 676-77 (discussing probabilistic approach to DNA evidence, where probative value of DNA match is based on objective likelihood calculations). Both possibilities face serious problems: a disconnect from reality and thus accuracy in one case (subjective) and the lack of data for most items of evidence in the other (objective). See Allen & Pardo, *supra* note 57 (manuscript at 9-11) (discussing these problems in detail).

<sup>70</sup> See Iuliano, *supra* note 3, at 413, 422 (explaining that jurors decide whether burden of proof has been satisfied with regard to each element, but discussing “lottery paradox” that occurs when jurors find burden satisfied for individual elements but not satisfied for conjunction of elements); Saul Levmore, *Conjunction and Aggregation*, 99 *MICH. L. REV.* 723, 723-34 (2001) (noting that jurors assess probability of each element and that law ignores problems inherent in conjunction and aggregation for purposes of reaching verdicts).

has been reached.<sup>71</sup> And if the requisite number concludes that the probability of an element is *not* greater than the threshold, then an opposite verdict has been reached. If neither of these conditions obtains, then no verdict has yet been reached.

## 2. The Explanatory Account

A second theoretical account of legal proof focuses on explanatory considerations to explain the proof process—including standards of proof, the relevance and probative value of evidence, and verdicts.<sup>72</sup> Because the probabilistic and explanatory accounts overlap to some degree, it will be helpful to highlight key similarities and differences between the two approaches before delving into the details of the explanatory account.

First, the probabilistic and explanatory accounts of legal proof are similar in that both attempt to explain the central features of the proof process: burdens and standards of proof, the probative value of evidence, and the requirements for verdicts. In doing so, both accounts provide a way to conceptualize or give content to these features. Moreover, both accounts focus on the same type of general decision-making task—that is, drawing inferences from evidence under conditions of uncertainty.<sup>73</sup>

The primary differences between the two theories concern: (1) how each characterizes the inferential process, and (2) the criteria each employs. As discussed above, probabilistic approaches characterize the inferential process as one that involves probabilistic judgments about each legal element. Within this

---

<sup>71</sup> See *supra* Section I.A (discussing jury majority and unanimity requirements for verdicts in civil and criminal cases); see also Iuliano, *supra* note 3, at 413, 422 (describing jurors' finding burden satisfied with regard to each element as necessary for reaching general verdict).

<sup>72</sup> See Michael S. Pardo & Ronald J. Allen, *Juridical Proof and the Best Explanation*, 27 *LAW & PHIL.* 223, 227-42 (2008) (describing how "inference to the best explanation" explains both macro-structure of proof at trial and micro-level issues regarding relevance and probative value of particular items of evidence); see also Ronald J. Allen, *Factual Ambiguity and a Theory of Evidence*, 88 *NW. U. L. REV.* 604, 606 (1994) (describing ability of relative-plausibility account to explain structure of proof of liability); Ronald J. Allen, *The Nature of Juridical Proof*, 13 *CARDOZO L. REV.* 373, 373 (1991) (describing how relative-plausibility account addresses epistemological problems inherent in juridical proof); Pardo, *supra* note 10, at 596-99 (describing explanatory conception of evidence and proof).

<sup>73</sup> See, e.g., GILBERT HARMAN, *CHANGE IN VIEW: PRINCIPLES OF REASONING* 65-75 (1986) (describing "explanatory coherence" of evidence as enabling inference-drawing under conditions of uncertainty); PETER LIPTON, *INFERENCE TO THE BEST EXPLANATION* 1-3, 107-20 (2d ed. 2004) (providing explanatory account of how individuals weigh evidence and make inferences, and arguing this account is compatible with probabilistic account, which likewise focuses on drawing inferences from evidence but is based on probabilistic calculations rather than explanatory heuristics); Tania Lombrozo, *Explanation and Abductive Inference*, in *THE OXFORD HANDBOOK OF THINKING AND REASONING* 260, 260 (Keith J. Holyoak & Robert G. Morrison eds., 2012) (examining cognitive process of "explanation" and its role in inference-making).

framework, both standards of proof and the value of evidence are given probabilistic interpretations. Accordingly, the central decision-making task is to match the probability of each element (based on the evidence) against the probabilistic threshold associated with the standard of proof.<sup>74</sup> By contrast, explanatory approaches characterize the inferential process as an evaluation of possible explanations of the evidence and events.<sup>75</sup> Within this explanatory framework, both standards of proof and the value of evidence depend on explanatory considerations. Accordingly, the central decision-making task is not to attach probabilities to the elements—it is to determine whether an explanation of the evidence (that includes all of the legal elements) satisfies the explanatory threshold associated with the applicable standard of proof.<sup>76</sup> The discussion below describes the explanatory account of standards of proof, the value of evidence, and verdicts.

Under the explanatory account, standards of proof specify thresholds for explanations that support the party with the burden of proof. The thresholds vary depending on the applicable standard, with higher standards requiring a higher threshold.<sup>77</sup> Under the “preponderance of the evidence” standard, factfinders determine whether the best available explanation of the evidence (and events) favors the plaintiff or the defendant.<sup>78</sup> An explanation favors the plaintiff if it includes all of the legal elements of the plaintiff’s claim; an explanation favors

---

<sup>74</sup> Pardo, *supra* note 10, at 595 (“At the macro-level, whether the evidence in a case is sufficient to satisfy a probabilistic proof standard will depend simply on whether the particular fact-finders think that the evidence surpasses the threshold.”).

<sup>75</sup> Pardo & Allen, *supra* note 72, at 223-34 (describing inferential process behind explanatory accounts of legal proof as one of “abduction” or “inference to the best explanation”).

<sup>76</sup> In general, the quality of an explanation serves as a proxy for likelihood: the better the explanation, the more likely, when compared with the available alternatives. *See* LIPTON, *supra* note 73, at 59 (describing “inference to the best explanation” as “the explanation that is most warranted: the likeliest”); Timothy Williamson, *Abductive Philosophy*, 2016 PHIL. F. 263, 267 (“Inference to the best explanation does not directly rank potential explanations according to their probability. This does not automatically make it inconsistent with a probabilistic epistemology . . . . [It] may be a good heuristic to use when—as often happens—probabilities are hard to estimate . . . .”); *see also* *Anderson v. Griffin*, 397 F.3d 515, 521 (7th Cir. 2005) (“[I]f in a particular case all the alternatives are ruled out, we can be confident that the case presents one of those instances in which [a] rare event did occur.”); *Bammerlin v. Navistar Int’l Transp. Corp.*, 30 F.3d 898, 902 (7th Cir. 1994) (“[Plaintiff] proceeded by eliminating the alternatives . . . . [Plaintiff] produced evidence that could lead a rational jury to eliminate the hypotheses inconsistent with his favored theory, which in turn permits an inference that his hypothesis is true.”); JOHN D. NORTON, *THE MATERIAL THEORY OF INDUCTION* chs. 8 & 9 (June 26, 2018) (unpublished manuscript), [https://www.pitt.edu/~jdnorton/homepage/cv.html#material\\_theory](https://www.pitt.edu/~jdnorton/homepage/cv.html#material_theory) [<https://perma.cc/8S9T-G8UK>] (providing illuminating discussion of explanatory inferences in science).

<sup>77</sup> In this respect, the probability and explanatory accounts are similar.

<sup>78</sup> *See* Pardo & Allen, *supra* note 72, at 234-35.

the defendant when it fails to include one or more elements.<sup>79</sup> A number of general criteria affect the strength or quality of an explanation. These criteria include considerations such as consistency, coherence, fit with background knowledge, simplicity, and the number of unlikely assumptions that need to be made.<sup>80</sup> For example, suppose a civil trespass case involving two disputed elements: whether the defendant entered the plaintiff's land and whether the defendant did so intentionally. The factfinder will compare whether the plaintiff's explanation (e.g., "the defendant intentionally entered land owned by the plaintiff") or the explanation advanced by the defendant (e.g., "the plaintiff is mistaken about the identity of the trespasser" or "defendant entered the land accidentally" or both) better fits with the evidence presented at trial.

Higher standards of proof require more. In criminal cases, under the "beyond a reasonable doubt" standard, the prosecution must do more than present a better explanation than the defense; factfinders should convict only when the prosecution's explanation (which includes all of the legal elements) is plausible, given the evidence, and there is *no* plausible defense explanation.<sup>81</sup> The "clear and convincing evidence" standard requires a threshold in between a preponderance and beyond a reasonable doubt. In explanatory terms, this means that the plaintiff's explanation must be not only better than the defendant's but also clearly more plausible in the decision maker's eyes.<sup>82</sup>

The explanatory account comports with underlying goals of the standards.<sup>83</sup> These goals include policy choices regarding accuracy and the risk of error.<sup>84</sup> Under the preponderance standard, the risk of error is divided roughly evenly

---

<sup>79</sup> Parties are generally allowed to provide alternative or disjunctive explanations, and it will sometimes be to their advantage to do so. *See id.* at 249 ("[I]f the jury believes that two mutually incompatible stories favor a party, the party gets the benefit of the disjunction of their probabilities.").

<sup>80</sup> *See id.* at 230 (explaining that "explanation is . . . better to the extent that it is consistent, simpler, explains more and different types of facts . . . better accords with background beliefs . . . is less ad hoc, and so on; and is worse to extent it betrays these criteria").

<sup>81</sup> *See, e.g.,* O'Laughlin v. O'Brien, 568 F.3d 287, 304-08 (1st Cir. 2009) (reversing conviction where prosecution's theory of case was implausible based on circumstantial evidence); United States v. Beard, 354 F.3d 691, 692-93 (7th Cir. 2004) (sustaining conviction for possession of firearm by passenger in borrowed automobile who conducted drug transaction from automobile because it was unlikely that gun belonged to car owner and defense proffered no alternative explanation); United States v. Newell, 239 F.3d 917, 920 (7th Cir. 2001) (affirming conviction where government presented strong case and defendant presented implausible alternative explanation, noting that government's "burden, even in a criminal case, is not to disprove every possibility that might exonerate the defendant"). The process is comparative in that it takes into account possible alternatives on both sides, but it requires more of the prosecution than merely having the better of the available explanations.

<sup>82</sup> Pardo, *supra* note 10, at 604 ("For the clear-and-convincing standard, the explanation must be substantially better than the alternatives.").

<sup>83</sup> *See id.* at 603-610.

<sup>84</sup> *See supra* note 38.

among the parties (each side bears the risk that the jury may mistakenly adopt the opposing side's explanation).<sup>85</sup> Moreover, under certain assumptions, accuracy improves to the extent that better explanations are more likely to be true than worse explanations.<sup>86</sup> Similarly, under higher standards of proof, the higher thresholds shift the risk of error away from the party without the burden of proof.<sup>87</sup> Accordingly, this reflects the preference to minimize one type of error over another.

The explanatory theory also provides an account of the relevance and probative value of evidence.<sup>88</sup> Evidence is relevant if it is part of either side's explanation or helps to distinguish between competing explanations.<sup>89</sup> Probative value depends on the significance of evidence for the explanations the parties advance—i.e., how well the evidence supports or falsifies one of the competing explanations.<sup>90</sup> This account of these concepts does not have the precision that follows from more formal definitions, but it has other advantages. First, it better fits actual assessments by judges and juries.<sup>91</sup> Second, it does not require numbers to quantify items of evidence (for most of which data will not be available). Third, because it concerns the relationships between evidence and explanations, it provides a more objective basis for decisions than subjective probability assessments.<sup>92</sup>

Finally, the explanatory account also explains verdict requirements. Individual jurors decide whether the explanation that includes all of the legal elements—typically, the one advanced by the prosecution or plaintiff—meets

---

<sup>85</sup> On this point, the explanatory account is a better fit than the probabilistic account. See Pardo, *supra* note 10, at 604 (“Each side bears the risk that the jury will select an explanation favoring the other side, with ties going to the party without the burden of proof.”).

<sup>86</sup> See *id.* at 609 (“The explanatory conception depends on the assumption that, other things being equal, a better explanation is more likely to be true than a worse explanation.”).

<sup>87</sup> As the standard increases, the party with the burden of proof must meet a higher explanatory threshold.

<sup>88</sup> See Pardo, *supra* note 10, at 600-03 (distinguishing explanatory theory's account of relevance and probative value from probabilistic theory).

<sup>89</sup> *Id.* (“Relevance . . . depend[s] on whether evidence supports a party's explanation or is a challenge to the other side's explanation.”); see also Maggie Wittlin, *Hindsight Evidence*, 116 COLUM. L. REV. 1323, 1342-44 (2016) (discussing value of hindsight evidence in terms of competing explanations).

<sup>90</sup> Pardo, *supra* note 10, at 610-11 (“Evidence that renders one side's explanation plausible or the other side's explanation implausible will be highly probative.”).

<sup>91</sup> Cf. Shari Seidman Diamond, Beth Murphy & Mary R. Rose, *The “Kettleful of Law” in Real Jury Deliberations: Successes, Failures, and Next Steps*, 106 NW. U. L. REV. 1537, 1563-64, 1575 (2012) (discussing juror use of burden-of-proof instructions in deciding cases and that jurors attempt to construct plausible account of what happened). See generally Allen & Pardo, *supra* note 57 (summarizing empirical literature).

<sup>92</sup> Therefore, it provides a more robust basis for “sufficiency of the evidence” determinations. See *supra* text accompanying notes 39-41.

the explanatory threshold of the applicable standard of proof.<sup>93</sup> If the requisite number of jurors (to constitute a verdict) concludes that the explanation meets the threshold, then there is a valid verdict for the party with the burden of proof.<sup>94</sup> By contrast, if the requisite number of jurors concludes the explanation does not meet that threshold, then a verdict for the other side (typically, defendants) has been reached. If neither of these conditions obtains, then no verdict has been reached.

\*\*\*\*\*

To sum up: this Part has provided background context necessary to understand the proof paradoxes. This context includes three features that are central to the proof process: standards of proof, the relevance and probative value of evidence, and verdicts. Two different theoretical accounts have been advanced to explain these legal features: probabilistic and explanatory. The Article now turns to the paradoxes themselves. One central theme that will emerge is that assumptions that drive the paradoxes follow from the probabilistic account of proof and not from the law itself.

## II. THE STATISTICAL-EVIDENCE PARADOX

The first paradox is the best known and includes a series of hypothetical cases involving *statistical* evidence. In each of these cases, there is one disputed fact on which liability or guilt depends (i.e., the identity of the defendant). The evidence on this issue consists of a single item of statistical evidence. What makes the examples “paradoxical” is that the evidence appears on its face to surpass the applicable standard of proof, and yet the judgment of most people is that the evidence is insufficient to prove liability or guilt. This apparent inconsistency between what the applicable legal rules appear to require, on one hand, and judgments about what the correct result ought to be, on the other, creates a tension that calls out for explanation. Scholars have offered a variety of suggestions to explain away the tension.

### A. *The Examples and Their Assumptions*

The most famous example is the “Blue Bus” hypothetical.<sup>95</sup> Here is one version:

---

<sup>93</sup> See Allen & Pardo, *supra* note 57 (manuscript at 13) (“Under the explanatory account, the central fact-finding task is not to attach probabilities to the individual elements; it is to determine whether potential explanations of the evidence and events satisfy the applicable standard of proof.”).

<sup>94</sup> See Pardo, *supra* note 21, at 1828-29 (describing how juries reach verdicts under explanatory theory).

<sup>95</sup> The example—which Professor Tribe referred to as a “famous chestnut,” Tribe, *supra* note 2, at 1341 n.37—is based on *Smith v. Rapid Transit*, 58 N.E.2d 754 (Mass. 1945). The

---

Plaintiff is negligently run down by a bus on Main Street. The Blue Bus Company owns seventy-five percent of the buses in town. The only witness is Plaintiff, who is color-blind and cannot testify to the color of the bus involved. Plaintiff sues the Blue Bus Company.

If we assume that neither party introduces other evidence, two questions arise. First, is a finding for the plaintiff *permissible*? This means that the evidence is sufficient to get to a jury, and, if the jury finds for the plaintiff, the evidence is sufficient as a matter of law to support that finding?<sup>96</sup> Second, is a finding for the plaintiff *required*? This means the plaintiff is entitled to judgment as a matter of law.<sup>97</sup>

According to one line of reasoning, the answer to both questions is *yes*: the plaintiff has met the proof requirements (question one) and indeed has done so as a matter of law (question two). The plaintiff must prove “by a preponderance of the evidence” that the bus involved was owned by the defendant. This line of thought formalizes the preponderance standard as a probabilistic threshold of greater than 0.5, and the evidence of ownership makes it 0.75 that it was a Blue Bus. Assuming no other evidence is introduced, the plaintiff has established her case to a 0.75 probability, and 0.75 is well beyond the preponderance threshold of 0.5. Therefore, not only is the plaintiff’s evidence sufficient to support a finding—a finding for the plaintiff is required as a matter of law because it is the only reasonable conclusion from the evidence.

Despite this line of reasoning, the “overwhelming intuition” is that the plaintiff has not met her proof requirements.<sup>98</sup> The general consensus is to reject

---

plaintiff in *Smith* did not attempt to introduce statistical evidence. Rather, the plaintiff introduced evidence that the defendant was the only bus company licensed to operate a public bus line on the street in question. *Id.* at 755. In upholding the trial court’s directed verdict for the defendant, Justice Spalding stated that the evidence was insufficient because of the lack of information about private or chartered buses also operating on the street. *Id.* The opinion then asserted the sentences that would inspire the paradox: “The most that can be said of the evidence in the instant case is that perhaps the mathematical chances somewhat favor the proposition that a bus of the defendant caused the accident. This was not enough.” *Id.* For a discussion of the case and an analysis of the paradox created by the Blue Bus problem, see SCHAUER, *supra* note 2, at 79-107.

<sup>96</sup> In other words, is the evidence sufficient to survive a motion for summary judgment or judgment as a matter of law? This possibility is consistent with a jury *not* finding for the plaintiff at trial. It only rules out the idea that any such finding would necessarily be unreasonable. *See supra* notes 40-41.

<sup>97</sup> Under this stronger claim, the evidence is not only sufficient to survive summary judgment or judgment as a matter of law for the defendant; the evidence mandates a judgment for the plaintiff. *See* Nesson, *supra* note 2, at 1379-80 (distinguishing these questions and noting “logic of the standard decision-theory model holds that the plaintiff is entitled to win” in the Blue Bus paradox).

<sup>98</sup> *See* Mike Redmayne, *Exploring the Proof Paradoxes*, 14 LEGAL THEORY 281, 281-82 (2008). As will be discussed in Section III.C, this intuition follows from a mistaken view about the standard of proof rather than from the nature of evidence itself.

the conclusions for *both* questions. In other words, the plaintiff is not entitled to judgment as a matter of law, and moreover, the evidence is not even sufficient to support a jury finding for the plaintiff (or to even get to trial).<sup>99</sup> According to this counterview, the *defendant* is entitled to judgment as a matter of law. These conflicting lines of reasoning indicate that something is amiss.

The tension arises from the conjunction of the following three premises:

**Premise 1 (Evidence):** the statistic in the evidence expresses the probative value of the evidence in proving the fact at issue. In other words, 0.75 reflects the probability that the bus involved was a Blue Bus.

**Premise 2 (Standard):** the preponderance standard requires proof beyond a 0.5 probability. In other words, the plaintiff must prove that it is more than fifty percent probable that the bus involved was a Blue Bus.

**Premise 3 (Outcome):** the plaintiff should lose.

The tension arises from the fact that Premise 1 and Premise 2 jointly imply that the burden of proof has been satisfied and therefore that Premise 3 is false. The tension thus arises from a contradiction between the first two premises (i.e., plaintiff should win) and Premise 3.

Before discussing possible solutions, it will be helpful to introduce an additional, well-known example that displays the same structure and tension among its underlying premises (involving evidence, standard, and outcome, respectively).<sup>100</sup> The “Gatecrashers” example was introduced into the legal literature by L. Jonathan Cohen:

Consider, for example, a case in which it is common ground that 499 people paid for admission to a rodeo, and that 1,000 are counted in the seats, of whom *A* is one. Suppose no tickets were issued and there can be no testimony as to whether *A* paid for admission or climbed over the fence. So

<sup>99</sup> *Howard v. Wal-Mart Stores, Inc.*, 160 F.3d 358, 360 (7th Cir. 1998) (“[S]hould a jury be allowed to award judgment to the plaintiff [in the Blue Bus hypothetical]? The law’s answer is ‘no.’”). Experimental data likewise support this conclusion. See Wells, *supra* note 13, at 740 (reporting that most subjects refused to find for plaintiff based on statistical evidence).

<sup>100</sup> The examples above involve proof of identity, but the paradox is not limited to disputes about identity. Similar issues arise for other disputed facts, such as causation, state of mind, and damages. See *United States v. Veysey*, 334 F.3d 600, 605 (7th Cir. 2003) (admitting statistical evidence to prove fires were not accidental); Charles Nesson, *Agent Orange Meets the Blue Bus: Factfinding at the Frontier of Knowledge*, 66 B.U. L. REV. 521, 530-32 (1986) (examining relationship of clinical and statistical proof of causation in context of Agent Orange); Tribe, *supra* note 2, at 1342-43 (noting examples involving disputes about intention); Richard W. Wright, *Causation, Responsibility, Risk, Probability, Naked Statistics, and Proof: Pruning the Bramble Bush by Clarifying the Concepts*, 73 IOWA L. REV. 1001, 1049-67 (1988) (describing different types of evidence used to determine whether specific causal law was fully instantiated).



by any plausible criterion of mathematical probability there is a 0.501 probability, on the admitted facts, that he did not pay.<sup>101</sup>

Suppose that the owner of the rodeo sues A. Is the evidence sufficient to support a finding for the owner? Is a finding for the owner required as a matter of law? The example raises the same issue as in Blue Bus—an item of statistical information appears to make the probability of the plaintiff's claim exceed the 0.5 threshold. At the same time, observers are reluctant to assign liability or to conclude that the evidence is sufficient to get to a jury in the first place. Moreover, Gatecrashers has the added issue that the same item of evidence that appears to make A liable could also be used in a lawsuit by the owner against any of the other 999 attendees.<sup>102</sup>

### B. *Proposed Solutions*

The paradox raises questions about whether the evidence satisfies the standard of proof. Possible solutions focus on each of the three premises. Let's start with the possible "Premise 3 (Outcome)" solutions. One possible strategy is simply to abandon Premise 3. According to this strategy, the evidence on its face surpasses the proof standard and therefore the plaintiff/prosecution is entitled to win. Accordingly, any anti-liability intuitions produced by the examples must be mistaken—perhaps the result of cognitive bias—and should be rejected.<sup>103</sup>

<sup>101</sup> COHEN, *supra* note 2, at 75; *see also* Blome-Tillmann, *supra* note 3, at 102-04.

<sup>102</sup> Is the plaintiff entitled to a victory in all 1,000 lawsuits, even though it is known that 499 bought tickets? *See* G. Alexander Nunn, *The Incompatibility of Due Process and Naked Statistical Evidence*, 68 VAND. L. REV. 1407, 1423-28 (2015) (discussing this feature of the Gatecrasher paradox). A third classic example, "Prisoners," introduced by Charles Nesson, raises similar issues in the criminal context:

In an enclosed yard are twenty-five identically dressed prisoners and a prison guard. The sole witness is too far away to distinguish individual features. He sees the guard, recognizable by his uniform, trip and fall, apparently knocking himself out. The prisoners huddle and argue. One breaks away from the others and goes to a shed in the corner of the yard to hide. The other twenty-four set upon the fallen guard and kill him. After the killing, the hidden prisoner emerges from the shed and mixes with the other prisoners. When the authorities later enter the yard, they find the dead guard and the twenty-five prisoners. Given these facts, twenty-four of the twenty-five are guilty of murder. Suppose that a murder indictment is brought against one of the prisoners—call him Prisoner I. . . . Nothing distinguishes Prisoner I from the other twenty-four prisoners.

Charles R. Nesson, *Reasonable Doubt and Permissive Inferences: The Value of Complexity*, 92 HARV. L. REV. 1187, 1192-93 (1979); *see also* Barbara Davidson & Robert Pargetter, *Guilt Beyond Reasonable Doubt*, 65 AUSTRALASIAN J. PHIL. 182, 183 (1987). The statistic surpasses the 0.9 threshold commonly associated with "beyond a reasonable doubt," but the common response is that the evidence is insufficient. Moreover, as with Gatecrashers, the incriminating evidence applies to every prisoner in the yard.

<sup>103</sup> *See* Amos Tversky & Daniel Kahneman, *Evidential Impact of Base Rates, in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 153, 157 (Daniel Kahneman, Paul Slovic &

Also focused on Premise 3, another possibility is to concede that liability follows from the first two premises (i.e., the evidence satisfies the standard of proof), but argue that liability should not be imposed for reasons external to Premises 1 and 2.<sup>104</sup> According to this strategy, other legal policies should preclude a liability finding *even though the evidence otherwise satisfies the burden of proof*.

Both types of Premise 3 solutions (liability and no liability) assume the truth of Premises 1 and 2. Solutions that depend on these additional considerations thus become necessary only if Premises 1 and 2 jointly imply liability.<sup>105</sup> For this reason, the discussion puts these solutions to the side and examines solutions focused on the first two premises. These premises are at the heart of the paradoxes—they are the ones that force one “to come to grips with key issues” for the law of evidence and proof.<sup>106</sup>

Premise 1 (Evidence) assumes that the statistic expresses the strength or the *probative value* of the evidence. In other words, the evidence proves that the probability a Blue Bus struck the plaintiff is 0.75.<sup>107</sup> The probability the bus was blue compared with the probability that it was not blue is 0.75/0.25, or three to

---

Amos Tversky eds., 1982) (“In spite of the witness’s report, therefore, the hit-and-run cab is more likely to be Green than Blue, because the base rate is more extreme than the witness is credible.”).

<sup>104</sup> These policy reasons tend to fall into two main categories: (1) instrumental reasons, and (2) additional epistemic reasons for not imposing liability. Examples in the first category include incentivizing ex ante non-litigation behavior, public acceptability, incentivizing the search for more evidence, and ensuring autonomy. *See, e.g.*, Enoch & Fisher, *supra* note 3, at 583 (incentivizing ex ante non-litigation behavior); Nesson, *supra* note 2, at 1366-68 (discussing public acceptance of jury verdicts); Richard A. Posner, *An Economic Approach to the Law of Evidence*, 51 STAN. L. REV. 1477, 1486-87 (1999) (discussing incentives to search for more evidence); David T. Wasserman, *The Morality of Statistical Proof and the Risk of Mistaken Liability*, 13 CARDOZO L. REV. 935, 942-43 (1991) (discussing how reliance solely on statistical proof demeans defendant’s individuality and autonomy). Additional epistemic requirements include causal relationships between evidence and verdicts, or between evidence and disputed facts, and judgments about the “weight” of evidence. *See, e.g.*, NANCE, *supra* note 3, at 111-12 (discussing weight of evidence); Judith Jarvis Thomson, *Liability and Individualized Evidence*, 49 LAW & CONTEMP. PROBS. 199, 201-02 (1986) (speculating that factfinders seek evidence indicating fault in addition to statistical evidence when determining causation); Wright, *supra* note 100, at 1063-65 (noting that jurors must find actual causation, not merely statistical likelihood of causation).

<sup>105</sup> This is not to diminish the significance of these issues. The point, rather, is that for purposes of the paradoxes they assume, rather than answer, fundamental issues raised by the paradoxes. *See* Sean P. Sullivan, *A Likelihood Story: The Theory of Legal Fact-Finding*, 90 U. COLO. L. REV. (forthcoming 2019) (manuscript at 44) (“Other arguments suggest that auxiliary policy objectives . . . may require the plaintiff to show more than technically needed to win . . . [T]his is again dodging the core difficulty of each puzzle.”).

<sup>106</sup> SAINSBURY, *supra* note 1, at 1.

<sup>107</sup> In the other examples, *see supra* notes 101-02 and accompanying text, the probability that the attendee crashed the gate would be 0.501 and the probability that the prisoner attacked the guard would be 0.95.

one.<sup>108</sup> Attempting to quantify the value of evidence in this manner, however, faces serious limitations. According to this conception of evidentiary value, data about relative distributions (e.g., blue and non-blue buses) provide the basis for quantification.<sup>109</sup> What could be wrong with that? There is, of course, nothing wrong with relying on objective data to draw inferences. But there is a problem with quantifying the *probative value* of evidence in this manner: the “reference class” problem.<sup>110</sup>

To illustrate this problem, notice that the distribution of buses in the town, for where the statistic arises is merely one reference class about bus distribution. There are countless other possible classes that include the event at issue (i.e., a bus accident). For example, the street where the accident occurred (suppose only five percent of the buses are Blue Buses), the time of day (suppose ten percent of the buses are Blue), and the percentage of bus accidents (suppose only one percent are Blue). Each one of these alternative classes provides a different “objective” value for the probability that the bus involved was Blue. The point of these hypothetical possibilities is to reveal that the probative value of evidence depends on much more than the likelihood ratio that follows from one particular reference class.<sup>111</sup> Probative value thus depends on a host of assumptions beyond the data; for example, assumptions about the appropriateness of the class for which data are available (why distribution in the town and not on the street?) and knowledge or assumptions about other classes (what if Blue Bus drivers have spotless driving records, and Red Bus drivers do not?). As a theoretical matter, the reference-class problem means that probative value cannot be *quantified* based on statistical distributions in the manner presupposed by the paradoxes.<sup>112</sup>

Once we recognize that probative value depends on more than the statistical distribution (or likelihood ratio), liability in the examples no longer necessarily follows. In other words, it is no longer the case that the evidence necessarily

---

<sup>108</sup> See *supra* note 66 and accompanying text.

<sup>109</sup> See *supra* note 66 and accompanying text.

<sup>110</sup> Ronald J. Allen & Michael S. Pardo, *The Problematic Value of Mathematical Models of Evidence*, 36 J. LEGAL STUD. 107, 109 (2007) (“Each of the reference classes leads to a different inference about which company is more likely liable, and nothing determines the correct class, save one: the very event under discussion, which has a likelihood of one and which we are trying to discover.”).

<sup>111</sup> Different ratios will follow from the same item of evidence simply by changing the reference class. *A fortiori*, no particular class by itself captures the probative value of the evidence.

<sup>112</sup> Of course, this does not mean that the evidence may not be particularly probative or that qualitative differences may not be made between different reference classes. See Edward K. Cheng, *A Practical Solution to the Reference Class Problem*, 109 COLUM. L. REV. 2081, 2095-97 (2009) (arguing that model selection methods provide tool for determining which proffered reference class is most appropriate because legal proceedings involve finite number of possible reference classes proposed by parties).

meets the standard of proof, *even if* the standards are thought of as probabilistic thresholds. Scholars otherwise attracted to probabilistic conceptions of proof have recognized this point and have offered alternative readings of the evidence in the paradoxes based on *subjective* probabilities.<sup>113</sup>

According to this subjective conception, the relative distribution of buses does not mean that the probability a Blue Bus caused the accident is 0.75. Rather, factfinders assign a probability to the latter based on how probable they *believe* this fact to be.<sup>114</sup> Conceptualizing evidence in this manner provides one way of responding to the paradoxes. In particular, it rejects Premise 1 and thereby *permits* (but does not require) factfinders to conclude that the evidence does not satisfy the standard (and thus accept Premise 3 without contradiction).

This conception, however, is a poor interpretation of legal evidence.<sup>115</sup> Relying on subjective probabilities to define the value of evidence is inconsistent with the fundamental goals and the basic features of the evidentiary-proof process. Factfinders' subjective beliefs could be anything at all, regardless of the evidence. For example, jurors in Blue Bus could think it is 0.01 or 0.99 probable that the bus was Blue, and nothing in this conception of evidence says whether either is incorrect or unreasonable. Defining the value of evidence in this manner is unlikely to advance accurate fact-finding at trial.<sup>116</sup> A fundamental assumption is that the quality of the evidence is distinct from the subjective beliefs of factfinders and that the former should guide the latter (not vice versa).<sup>117</sup> Defining the quality of evidence based solely on the subjective beliefs of decision makers ignores the truth-conducive aspects of evidence. Better evidence, other things being equal, should lead to more accurate outcomes. But this is not necessarily so under a subjective conception; the relationship between subjective beliefs about evidence and truth could be anything at all.

This conception is also inconsistent with basic aspects of legal proof. In particular, it cannot account for "sufficiency of evidence" requirements at trial

---

<sup>113</sup> See, e.g., Kaye, *supra* note 2, at 106 ("[I]t may be appropriate to treat the subjective probability as less than one-half, and therefore insufficient to support a verdict for plaintiff . . ."); see also SAVAGE, *supra* note 68, at 56 ("It is my tentative view that the concept of personal probability . . . is . . . the only probability concept essential to science and other activities that call upon probability.").

<sup>114</sup> This belief may be based on any idiosyncratic opinion whatsoever, no matter how ridiculous, or based on nothing at all.

<sup>115</sup> See Pardo, *supra* note 10, at 591 ("Subjective assessments could be anything at all, and there is simply no reason to think they will be truth conducive.").

<sup>116</sup> Goldman, *supra* note 69, at 239 ("[S]ubjective Bayesianism does not commend itself as a basis for truth acquisition. It is not at all clear how purely subjective Bayesian methods, applied to the legal context, hold any promise of leading a trier of fact to truth.").

<sup>117</sup> See SUSAN HAACK, *Legal Probabilism: An Epistemological Dissent*, in EVIDENCE MATTERS: SCIENCE, PROOF, AND TRUTH IN THE LAW 47, 56-64 (William Twining, Christopher McCrudden & Bronwen Morgan eds., 2014) (distinguishing warranted inferences based on evidence from subjective beliefs of factfinders); LARRY LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW: AN ESSAY IN LEGAL EPISTEMOLOGY 51-54 (2006) (same).

and throughout the processes of civil and criminal litigation. Legal doctrine requires courts, at several litigation stages, to assess whether the evidence is sufficient to support a “reasonable” or “rational” jury finding.<sup>118</sup> Subjective conceptions of evidence provide no criteria for making this determination for the simple reason that *every* finding would count as reasonable or rational.<sup>119</sup> In addition to being inconsistent with several aspects of legal doctrine and practice, this point also connects with the paradoxes—even if the subjective view *permits* jurors to find no liability, it also permits them to find liability in each of the examples. This conception also implies that in each example, the case should go to the jury because they might, based on their subjective beliefs, find liability. Indeed, according to this conception, *every* case should go to the jury—there would be no basis, for example, for a court deciding a motion for summary judgment or judgment as a matter of law to declare that a particular finding is unreasonable.<sup>120</sup> These implications are even more problematic than those raised by the liability conclusions in the paradoxical examples. When it comes to this possible solution, the “cure” is worse than the ailment.

In sum, Premise 1 raises deep questions about the probative value of evidence. Solutions that define probative value in terms of either objective (i.e., relative frequencies) or subjective (i.e., “degrees of belief” or “credences”) probabilities, however, run into devastating problems. As discussed in Part I, the *explanatory* account of proof provides an alternative conception of probative value. According to this account, probative value depends on the role the evidence plays in supporting and challenging one or more of the explanations proffered by the parties.<sup>121</sup> This account obviously lacks the precision that comes from the quantification of probative value. Moreover, it is not clear how to integrate this conception with the standards of proof if the law defines the latter as probabilistic thresholds (Premise 2). Nor is it clear, at this point, how the explanatory account characterizes the evidence in the paradoxes. Without an account of standards of proof (Premise 2), we cannot draw clear conclusions about the examples. The explanatory account, however, avoids several of the pitfalls facing objective and subjective probabilistic conceptions of evidence: (1) it is more feasible to implement than objective probability approaches, (2) it does not require that decision makers quantify the evidence at all, and (3) it provides criteria other than subjective beliefs by which to assess particular judgments.<sup>122</sup>

---

<sup>118</sup> See *supra* notes 39-41 and accompanying text.

<sup>119</sup> Or, alternatively, the only unreasonable conclusions would be those that failed to comply with minimal consistency requirements. See SAVAGE, *supra* note 68, at 57 (“According to the personalistic view, the role of the mathematical theory of probability is to enable the person using it to detect inconsistencies in his own real or envisaged behavior. It is also understood that, having detected an inconsistency, he will remove it.”).

<sup>120</sup> See *supra* notes 39-41 and accompanying text.

<sup>121</sup> See *supra* text accompanying notes 88-92.

<sup>122</sup> See Pardo, *supra* note 10, at 600-03 (discussing these features).

In short, the existence of an alternative (i.e., non-probabilistic) theoretical account of probative value should give us pause before drawing conclusions about Premise 1. We must look more closely at Premise 2.

Premise 2 (Standard) plays a key role in the paradox. Liability in each case depends on what the standard of proof means and requires. The reasoning that leads to liability based on the statistical evidence (Premise 1) presupposes that the standard is a probability threshold—namely, that “preponderance of the evidence” means proof beyond a 0.5 probability. This interpretation, however, is not inherent in the standards themselves, and this way of conceptualizing the standard is just one theoretical possibility.<sup>123</sup> If standards of proof are *not* best conceived of as probability thresholds, then liability in the statistical-evidence examples no longer necessarily follows (regardless of what one concludes about Premise 1). Importantly, the examples tell us nothing about the standards of proof and what they require—they merely assume a probabilistic threshold of 0.5.

Unlike the statistical-evidence paradox, the other types of paradoxes (discussed in Parts III and IV) focus on fundamental issues regarding standards of proof. This is one way in which the paradoxes are interrelated—it is simply not possible to determine a correct answer to the statistical-evidence examples unless and until one has also grappled with the fundamental issues about standards of proof the other paradoxes raise.<sup>124</sup> Therefore, we are not yet in a position to evaluate possible Premise 2 solutions to the statistical-evidence paradoxes.

We are in a position, however, to diagnose some of the reasons why the paradoxes continue to fuel fascination and frustration among scholars. Liability appears to follow from particular assumptions, typically left implicit, about probative value and standards of proof. These assumptions, however, depend upon a particular theoretical conception that is itself contested and problematic.<sup>125</sup> Faced with the implications that follow directly from this conception, scholars naturally look outside of Premises 1 and 2 for a solution. This category of solutions, however—while unearthing potentially important legal policies and principles—essentially sidesteps the fundamental proof issues.<sup>126</sup>

---

<sup>123</sup> See *supra* Section I.B.

<sup>124</sup> The discussion returns to these issues in Parts III and IV, *infra*. For a discussion within the philosophical literature, see Smith, *supra* note 3, at 18-19 (suggesting possible reforms to standards of proof as potential solution to statistical-evidence paradox).

<sup>125</sup> Moreover, no alternative conception implies a conclusion as clearly and cleanly as the probabilist one supports liability.

<sup>126</sup> See *supra* note 104 and accompanying text.

C. *Statistical vs. Individualized Evidence*

We are also now in a position to clarify another important confusion generated by the statistical-evidence paradox, one deserving of its own subsection. It is widely assumed that the paradoxes reveal that there is something deficient with *statistical* evidence as opposed to evidence that is not explicitly in statistical form.<sup>127</sup> For reasons discussed below, however, this assumption is false. Although the paradoxes each involve statistical evidence, nothing in the examples themselves supports a general distinction between statistical and non-statistical evidence.<sup>128</sup> Rather, the skeptical conclusion about statistical evidence is thought to follow from two additional sources. The first source is evidence—both experimental research and the intuitions of commentators—that individual items of non-statistical evidence *would be* sufficient to prove liability in the examples.<sup>129</sup> For example, an eyewitness in Blue Bus testifies that the bus was blue. Such evidence is typically characterized as “individualized” or “case specific.”<sup>130</sup> The second source is actual legal cases in which courts have expressed skepticism about statistical evidence or have found such evidence to be insufficient, inadmissible, or irrelevant.<sup>131</sup> Neither source, however, supports a general distinction between statistical and non-statistical evidence.

---

<sup>127</sup> See Wells, *supra* note 13, at 748 (concluding that “people are reluctant to accept naked probabilities as evidence of liability”); see also Enoch & Fisher, *supra* note 3, at 558 (“Despite the voluminous body of literature dedicated to the issue of statistical evidence, it continues to generate great controversy in evidence law scholarship.”).

<sup>128</sup> Another distraction in these debates is the idea that the examples concern so-called “naked” statistical evidence, which is supposed to somehow be distinct from other types of evidence. Exactly what this means, however, is not entirely clear, and there does not appear to be a non-question-begging way of drawing a distinction along these lines. At best, the term simply means base rates; at worst, it simply means evidence people think is insufficient. In any event, it is not clear that the term is doing analytical work; therefore, it is best to abandon this term and talk directly about the features of the evidence at issue. Other scholars have also noted problems with this term. See Craig R. Callen, *Spotting a Preponderance of the Evidence in the Wild: Inference to the Best Explanation and Sufficiency of the Evidence*, 48 SETON HALL L. REV. 1517, 1531 (2018) (“Debates over blue bus hypotheticals often touched on the status of ‘naked’ or ‘solely’ statistical evidence . . . . Discussions of the question are confusing, because the distinction between naked statistics and other sorts of evidence is ill-founded. A ‘naked’ or ‘sole’ statistic would simply be a number, with no information about what it quantified, or the context in which one might make inferences based on it. There would be no reason to regard it as probative of anything.”); Wells, *supra* note 13, at 739 (“Naked statistical evidence is ill defined in the legal literature but typically refers to probabilities that are not case specific in the sense that the evidence was not created by the event in question but rather existed prior to or independently of the particular case being tried.”).

<sup>129</sup> See Redmayne, *supra* note 98, at 281-82; Wells, *supra* note 13, at 746.

<sup>130</sup> See Enoch & Fisher, *supra* note 3, at 559-61 (explaining prevailing preference in legal system for individualized evidence); Wells, *supra* note 13, at 746 (discussing insufficiency of subjective probability in driving verdicts).

<sup>131</sup> See Jonathan J. Koehler, *When Do Courts Think Base Rate Statistics Are Relevant?*, 42 JURIMETRICS 373, 377-80 (2002) (discussing examples).

Legal scholars have distinguished the evidence in Blue Bus from more “individualized” evidence such as an eyewitness.<sup>132</sup> Experimental findings provide support for the significance of such a distinction, even when it is acknowledged that the eyewitness could be mistaken.<sup>133</sup> Moreover, and perhaps surprisingly, this is so even if one stipulates that the eyewitness is reliable seventy-five percent of the time (i.e., the same number as in Blue Bus). Thus, it is claimed, the paradoxes reveal the law’s skepticism toward statistical evidence.<sup>134</sup> This skepticism is then taken to reveal either the epistemic inferiority of statistical evidence or a cognitive bias in the law.<sup>135</sup>

These lines of reasoning, however, are based on the mistaken assumption that there *is* a meaningful general distinction between statistical and non-statistical evidence that follows from the examples.<sup>136</sup> Four points, in combination, will explain why the assumption does not follow from the examples. We can use the Blue Bus and eyewitness cases to illustrate these points. First, probative value includes more than the statistical information. Therefore, the fact that two items of evidence possess the same statistical distribution does not mean that they necessarily have the same probative value.<sup>137</sup> Second, whatever positive features make eyewitness testimony (or any similar evidence) more probative than the base-rate in Blue Bus will also be possessed by some items of statistical evidence.<sup>138</sup> Thus, whatever the defect is in examples such as Blue Bus, the defect is not that it is statistical in form. Third, and relatedly, whatever negative features make the evidence in Blue Bus problematic will also be possessed by items of non-statistical evidence. Thus there is no general distinction to be

---

<sup>132</sup> See sources cited *supra* note 130.

<sup>133</sup> Wells, *supra* note 13, at 749.

<sup>134</sup> See Enoch & Fisher, *supra* note 3, at 558.

<sup>135</sup> See *id.* at 579 (arguing legal system sacrifices accuracy when excluding statistical evidence based on perceived epistemic inferiority to individualized evidence); Daniel Shavero, *Statistical-Probability Evidence and the Appearance of Justice*, 103 HARV. L. REV. 530, 552-53 (1989) (discussing bias from lawyers and judges to favorably portray legal system by downplaying statistical evidence).

<sup>136</sup> For additional arguments challenging this distinction, see Callen, *supra* note 128, at 1531-32; Shavero, *supra* note 135, at 530-31 (explaining courts’ aversion to “overtly probabilistic evidence that expressly states the risk of error”).

<sup>137</sup> See Allen & Pardo, *supra* note 110, at 114-16. Scholars have pointed to a number of features that may make the evidence in the eyewitness case more probative than the ownership statistic. See, e.g., Thomson, *supra* note 104, at 203-05 (exploring perceived additional value of eyewitness evidence indicating causal connection).

<sup>138</sup> See NANCE, *supra* note 3, at 111 (arguing that “paradoxical element” in examples is present “in all fact-finding” and not merely cases involving statistical evidence: “whatever analysis is needed to resolve them is also needed to confront the ubiquitous use of generalizations in the process of inference”); ALEX STEIN, FOUNDATIONS OF EVIDENCE LAW 64-106 (2005) (arguing that reliable evidence must be adequate both qualitatively and quantitatively); Callen, *supra* note 128, at 1531-32 (explaining that context is necessary feature of statistical evidence).



drawn—whether in terms of general epistemic features or probative value—between statistical and non-statistical evidence.<sup>139</sup> Fourth, and finally, both examples (ownership and eyewitness) are significantly underdetermined because many potentially relevant details are left unspecified.<sup>140</sup> Even if we draw distinctions in the examples in terms of *probative value*, we cannot tell whether either piece of evidence is sufficient by itself until we specify what the standard of proof means and requires. It is presupposed that the eyewitness evidence would be sufficient, given a “beyond 0.5” conception of the proof standard. Perhaps this is correct—but general doubts about that conception should carry over to the eyewitness case as well.<sup>141</sup> These reasons cut against drawing any general conclusions about statistical evidence from the examples.

The caselaw provides concrete support for these reasons,<sup>142</sup> revealing a pattern much more complex than those indicated by the paradoxical examples.<sup>143</sup>

---

<sup>139</sup> Non-epistemic based policy recommendations, *see supra* note 104 and accompanying text, likewise cut across the distinction between statistical and non-statistical evidence.

<sup>140</sup> The simplified and unrealistic nature of these examples makes drawing conclusions about law problematic. Actual cases involve more evidence and information, including arguments about why no other evidence is forthcoming, as well as competing explanations of the evidence and events, which we do not have in the examples. *See Callen, supra* note 128, at 1532 (“Some courts have thought that they confronted the issue of whether they should permit a verdict based on ‘naked’ statistical evidence, but they had a great deal more information than a single, solitary number. Generally the question in such cases was whether, when the evidence would not be sufficient to support a verdict without the statistical probability evidence, the evidence would be sufficient if the statistical information were added.”).

<sup>141</sup> *See MOSS, supra* note 3, at 216-20, for an illuminating discussion.

<sup>142</sup> *See NANCE, supra* note 3, at 108 n.13 (“It is often claimed or assumed . . . that courts would not in fact allow a judgment for the plaintiff under the assumed facts. Insofar as such claims are considered descriptions of what courts in fact do or predictions of what courts will do, they are not supported by substantial authority in the case law.”); Koehler, *supra* note 131, at 385-400 (analyzing factors affecting admissibility of base-rate evidence). The discussion below focuses on the use of statistical evidence, as in the examples, to prove contested facts of consequence in particular cases. Other potential uses, such as alternations in doctrine, are outside the scope of the discussion. *See, e.g., Sindell v. Abbott Labs., 607 P.2d 924, 936-38 (Cal. 1980)* (considering statistical evidence in context of market-share liability).

<sup>143</sup> *See, e.g., Halliburton Co. v. Erica P. John Fund, Inc., 573 U.S. 258, 280 (2014)* (noting that either side may present statistical analyses to prove whether stock price was affected by publicly available information); *Ricci v. DeStefano, 557 U.S. 557, 586-87 (2009)* (stating statistical disparity may be sufficient to prove prima facie case of discrimination); *United States v. Utah v. Evans, 536 U.S. 452, 464-79 (2002)* (discussing census sampling method used to appoint congressional representatives); *United States v. Fior D’Italia, Inc., 536 U.S. 238, 246-49 (2002)* (considering statistical methods for estimating tax liability); *Int’l Bhd. of Teamsters v. United States, 431 U.S. 324, 339 (1977)* (discussing importance of statistical evidence for proving discrimination); *Brown Shoe Co. v. United States, 370 U.S. 294, 340-46 (1962)* (considering statistical analysis of shoe sales in antitrust context); Michael D. Green, D. Michal Freedman & Leon Gordis, *Reference Guide on Epidemiology*, in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 333, 384 (2d ed. 2000) (discussing use of

It is therefore a mistake to infer that the examples capture something essential about legal doctrine.

The Supreme Court's recent opinion in *Tyson Foods, Inc. v. Bouaphakeo*<sup>144</sup> provides an instructive example.<sup>145</sup> The case, a class action, involved a dispute concerning the time that it took employees to "don and doff" protective gear, which was relevant to plaintiffs' claim for overtime pay. The plaintiffs (employees) proffered statistical evidence on the average time that it took an employee to don and doff the gear, based on a study conducted by an expert witness. The statistical average was then applied to the individual plaintiffs.<sup>146</sup> The Court held that the evidence was properly admitted and relied upon.<sup>147</sup> In doing so, the Court explained that there is no categorical rule regarding statistical evidence, whether in class actions or individual lawsuits. Rather, the Court explained, "like all evidence," the permissibility of statistical evidence depends on "the degree to which the evidence is reliable in proving or disproving the elements of the relevant cause of action."<sup>148</sup> Moreover, the Court clarified that the statistical evidence was not only permissible, but that it would also be *sufficient* to support a finding for the plaintiff in any individual employee's lawsuit.<sup>149</sup>

The statistical evidence in *Tyson Foods* is one example from a wide array of permissible uses of statistical evidence. Other prominent examples of potential uses include DNA evidence<sup>150</sup> as well as uses in discrimination, antitrust, and toxic-tort cases.<sup>151</sup> Indeed, for some issues, such as proving "disparate impact" in discrimination cases or causation in tort cases, statistical evidence will be not only permissible but also essential.<sup>152</sup> Even in cases that most closely resemble

---

"relative risk" analyses to prove causation in toxic tort cases: "a relative risk of 2.0 . . . implies a 50% likelihood that an exposed individual's disease was caused by the agent.").

<sup>144</sup> 136 S. Ct. 1036 (2016).

<sup>145</sup> For illuminating discussions of the case, see Robert G. Bone, *Tyson Foods and the Future of Statistical Adjudication*, 95 N.C. L. REV. 607 *passim* (2017); Jonah B. Gelbach, *The Triangle of Law and the Role of Evidence in Class Action Litigation*, 165 U. PA. L. REV. 1807, 1809-15 (2017).

<sup>146</sup> *Tyson Foods*, 136 S. Ct. at 1039.

<sup>147</sup> *Id.* at 1043-46 ("A representative or statistical sample, like all evidence, is a means to establish or defend against liability."). The Court rejected the argument that such evidence was improper because of possible variations involving individual employees. *Id.* at 1046.

<sup>148</sup> *Id.*

<sup>149</sup> *Id.* at 1048-49.

<sup>150</sup> See Andrea Roth, *Safety in Numbers? Deciding when DNA Alone Is Enough to Convict*, 85 N.Y.U. L. REV. 1130, 1184 (2010) (concluding DNA match statistics are under certain conditions sufficient to convince juries of guilt).

<sup>151</sup> See Bone, *supra* note 145, at 636-37 (positing *Tyson Foods* holding can apply to other kinds of cases).

<sup>152</sup> See *Ricci v. DeStefano*, 557 U.S. 557, 587 (2009) (indicating statistical evidence is sufficient to show disparate impact in discrimination cases); *Wells v. SmithKline Beecham Corp.*, 601 F.3d 375, 380 (5th Cir. 2010) ("[T]his court has frowned on causative conclusions

the unrealistic paradoxical examples—namely, those attempting to prove liability or guilt based on base-rate evidence—courts have expressed mixed views about the evidence.<sup>153</sup> In a survey of the caselaw, Professor Jonathan Koehler explains when courts are likely to admit such evidence.<sup>154</sup> The admissibility and sufficiency of such evidence (like all evidence) will depend on its details and will vary from case to case.<sup>155</sup> In short, there is no general rule for statistical evidence. The paradoxes cast a misleading impression of the law's stance toward statistical evidence rather than projecting a true reflection of doctrine.

In sum, despite being phrased as ostensibly about *statistical* evidence, nothing of significance for evidence law, theory, or policy follows about statistical evidence (*qua* statistical) from the examples. The real import of the examples concerns the issues they raise about probative value and standards of proof. Although they help to uncover implicit assumptions about these legal concepts, the examples do not answer the questions they help to uncover. This perhaps explains, in part, why the paradoxes remain a source of fascination and frustration—why they remain, in other words, paradoxical. Understanding probative value and standards of proof requires connecting the statistical-evidence paradoxes with the other types of paradoxes.

### III. THE CONJUNCTION PARADOX

Although lesser known than the statistical-evidence examples, the “conjunction” paradox, which applies to any kind of evidence, has also played a prominent role in evidence scholarship. As Professor Kevin Clermont explains: “almost every law professor who finally contemplates the problem accepts as a matter of common sense that the law’s approach is paradoxical.”<sup>156</sup> The discussion first explains the paradox, then turns to possible solutions, and finally

---

bereft of statistically significant epidemiological support.”).

<sup>153</sup> Compare *Guenther v. Armstrong Rubber Co.*, 406 F.2d 1315, 1318 (3d Cir. 1969) (evidence defendant made seventy-five to eighty percent of tires at issue not sufficient to survive summary judgment), with *Kaminsky v. Hertz Corp.*, 288 N.W.2d 426, 429 (Mich. Ct. App. 1980) (holding evidence that defendant owned ninety percent of trucks at issue was sufficient to survive summary judgment and to create rebuttable presumption of ownership).

<sup>154</sup> Koehler, *supra* note 131, at 402 (explaining factors that affect probative value, including when evidence is offered to rebut “chance” hypotheses, are derived from narrow or refined reference classes, and when other evidence is unobtainable).

<sup>155</sup> See *Tyson Foods, Inc. v. Bouaphakeo*, 136 S. Ct. 1036, 1048 (2016) (“[T]he study here [providing an average number of overtime hours] could have been sufficient to sustain a jury finding as to hours worked if it were introduced in each individual action.”). For a discussion of the complex relationship between statistical evidence and legal standards of proof, see Michelle M. Burtis, Jonah B. Gelbach & Bruce H. Kobayashi, *Error Costs, Legal Standards of Proof and Statistical Significance*, SUP. CT. ECON. REV. (forthcoming) (analyzing statistical significance tests’ relationship to standards of proof and discussing consequences).

<sup>156</sup> Clermont, *Common Sense*, *supra* note 3, at 1059.

explains the significance of the paradox and its relationship with the statistical-evidence paradox.

A. *What Is the Conjunction Paradox?*

The conjunction paradox arises from a tension between how the law applies burdens of proof and standard probabilistic reasoning.<sup>157</sup> Legal doctrine and jury instructions apply burdens and standards of proof to the individual elements of a claim, crime, or affirmative defense.<sup>158</sup> Therefore, a civil plaintiff will, for example, have to prove each of several elements by a preponderance of the evidence to prevail.<sup>159</sup> Similarly, the prosecution must prove each element of a crime beyond a reasonable doubt.<sup>160</sup> If the standards of proof are probabilistic thresholds, then parties with the burden of proof will win if they surpass the threshold for each element.<sup>161</sup>

Applying the burden to each element, however, appears to reject standard probabilistic reasoning.<sup>162</sup> When there are two or more propositions, the probability of their combination (conjunction) will not only depend on the probabilities of the individual propositions, it will also depend on the relationship between the propositions. If two propositions, *A* and *B*, are *independent* of each other (i.e., the probability of one does not affect the probability of the other), then the probability of the conjunction (*A* & *B*) will be *A* multiplied by *B*.<sup>163</sup> For example, if the probability of a coin landing heads up is 0.5, then the probability of getting two heads in a row is 0.25 (0.5 x 0.5).<sup>164</sup> This theorem in probability theory—sometimes referred to as the “product rule”—states that the probability of any two independent propositions is their product.<sup>165</sup>

When the propositions are *dependent*, then the probability of both propositions being true (the conjunction) will depend on the conditional

<sup>157</sup> Modern discussions of the conjunction problem arise in COHEN, *supra* note 2, at 58-67.

<sup>158</sup> See Ronald J. Allen & Sarah A. Jehl, *Burdens of Persuasion in Civil Cases: Algorithms v. Explanations*, 2003 MICH. ST. L. REV. 893, 898-902 (surveying jury instructions); Allen & Pardo, *supra* note 57 (manuscript at 6) (explaining burdens and standards of proof).

<sup>159</sup> See 3B KEVIN F. O'MALLEY, JAY E. GRENIG & HON. WILLIAM C. LEE, FEDERAL JURY PRACTICE AND INSTRUCTIONS § 168:200 (6th ed. 2017).

<sup>160</sup> *In re Winship*, 397 U.S. 358, 361 (1970).

<sup>161</sup> See *supra* note 65 and accompanying text.

<sup>162</sup> See sources cited *supra* note 58.

<sup>163</sup> See HACKING, *supra* note 58, at 58-68.

<sup>164</sup> This outcome is one of four equally likely outcomes (heads-heads, heads-tails, tails-heads, and tails-tails).

<sup>165</sup> See COLIN AITKEN, PAUL ROBERTS & GRAHAM JACKSON, FUNDAMENTALS OF PROBABILITY AND STATISTICAL EVIDENCE IN CRIMINAL PROCEEDINGS 87 (“The product rule for independent events for calculating conjunctive probabilities should be applied only to verifiably independent events. Independence should never be a default assumption in criminal proceedings . . . .”); *supra* text accompanying note 18.

relationship between the propositions.<sup>166</sup> For example, it might be impossible for two propositions to both be true. Therefore, even if each has a probability of 0.5, the probability of their conjunction ( $A \& B$ ) may be 0.<sup>167</sup> On the other hand, one proposition may entail the other and thus the probability of the conjunction ( $A \& B$ ) will be the same as probability of  $A$ .<sup>168</sup> The theorem for dependent propositions states that the probability of the conjunction ( $A \& B$ ) is equal to  $A$  multiplied by  $B$  given  $A$ .

This apparent conflict between the law's approach to burdens of proof and probabilistic reasoning produces "paradoxical" conclusions. For the purposes of the examples to follow, this Article will assume (for simplicity's sake only) independence among the elements.<sup>169</sup> Suppose a civil plaintiff's claim is based on two elements,  $X \& Y$ , and plaintiff proves each to 0.6. The plaintiff will win. If these elements are independent, however, then the plaintiff's claim is only 0.36 likely to be true.<sup>170</sup> The addition of more elements only seems to make the problem worse: a claim with four elements, each proven to 0.6, is only 0.1296 likely, and so on. Applying the burdens of proof in this manner appears to award parties with victories even though their claims as a whole appear to fall well below the standard of proof.<sup>171</sup>

---

<sup>166</sup> In other words, the probability of  $A \& B$  will equal the probability of  $A$  times the probability of  $B$  given  $A$  ( $A \& B = A \times B|A$ ).

<sup>167</sup> Putting aside abstract hypotheticals, for example, the probability that the coin landed heads ( $A$ ) and tails ( $B$ ) on the same toss would be zero.

<sup>168</sup> Suppose we are using a Lincoln penny, again putting abstract hypotheticals to the side: the probability the coin landed heads ( $A$ ) and with a representation of Lincoln ( $B$ ) would be the same as the probability of  $A$ .

<sup>169</sup> See Allen, *supra* note 2, at 405 n.18 (making same simplifying assumption).

<sup>170</sup> See Nesson, *supra* note 2, at 1390 (pointing out that combination  $A \& B$  in such circumstances will be more probable than other three possible combinations:  $A \& \text{not-}B$  (0.24),  $\text{not-}A \& B$  (0.24), and  $\text{not-}A \& \text{not-}B$  (0.16)).

<sup>171</sup> Proceeding in this manner is inconsistent with the underlying goals of the standards regarding accuracy and the risk of error. See *supra* note 38 and accompanying text. For example, when a plaintiff's claim is 0.1296 likely, then we would expect that in one hundred similar cases only about thirteen of the claims will be true. Hamer, *supra* note 38, at 75 ("If ten plaintiffs proved their cases to a level of 60 per cent, all would succeed. Six of these verdicts would be expected to be factually correct, and four factually incorrect . . ."). Awarding victories to all one hundred plaintiffs (because they all prove each element to 0.6) would thus result in thirteen correct decisions and eighty-seven errors. This arrangement appears to unfairly place more of the risk of error onto defendants. See *id.* at 75-76. The same is true with regard to defendants and affirmative defenses. Moreover, applying the burden in this manner generates additional paradoxical implications. For example, consider again a two-element claim in which the plaintiff proves each element to 0.6 and wins. Compare this with a second case: this plaintiff proves one element to 0.9 and the other to 0.5. This plaintiff loses even though their claim is more likely to be true than the first claim (0.45 vs. 0.36). The plaintiff with the stronger claim loses and the plaintiff with the weaker claim wins. The comparative aspect of proof, see *supra* notes 78-87 and accompanying text, also generates paradoxical implications. Suppose in a two-element claim, the plaintiff proves each to 0.4 and

As with the statistical-evidence paradox, it will help to make explicit the assumptions that underlie the conjunction issue. The paradox arises from the inconsistent conclusions that follow from two lines of reasoning. We can illustrate these two lines with the example of a two-element claim in which the plaintiff proves each to 0.6.

According to one line of reasoning (law's practices):

**Premise 1 (Elements):** the plaintiff must prove each element by a preponderance of the evidence.

**Premise 2 (Standard):** the preponderance standard requires a probabilistic threshold of greater than 0.5.

**Premise 3 (Example):** the plaintiff has proven each element to a probability greater than 0.5.

**Conclusion:** Plaintiff wins.

According to a second line of reasoning (probabilistic reasoning):

**Premise 1\* (Conjunction):** the goals underlying the standards of proof require the standard be applied to the conjunction of the elements (i.e., the case as a whole).<sup>172</sup>

**Premise 2\* (Standard):** the preponderance standard requires a probabilistic threshold of greater than 0.5.

**Premise 3\* (Product Rule):** The probability of two (independent) elements is their product.<sup>173</sup>

**Premise 4\* (Example):** The plaintiff has not proven the conjunction of the elements to a probability of greater than 0.5.<sup>174</sup>

**Conclusion\*:** Plaintiff loses.

Given these inconsistent conclusions, something above has to go.

---

the defendant presents an alternative account that is 0.2 probable for each element. Plaintiff loses even though the plaintiff's version of the events is four times more likely to be true than the defendant's (0.16 vs. 0.04). This result also conflicts with the goals of accuracy and equally allocating the risk of error. *See supra* note 38 and accompanying text. The conclusion follows from another rule of probability theory, the complementation axiom. *See supra* notes 58-63. This is the idea that if the probability of *A* being true is 0.6, then the probability of *A* being false is 0.4. In other words, any proposition *X* plus not-*X* sum to 1. In the litigation context, this assumption implies that the party with the burden of proof has to show more than that their account is more likely than the other side's—they have to show their account is more likely than the combination of every possibility that supports the other side. For criticism of this assumption in the context of litigation, see Allen & Pardo, *supra* note 57 (manuscript at 9-13).

<sup>172</sup> *See supra* note 38 and accompanying text.

<sup>173</sup> Again, the independence assumption is merely to simplify the discussion. Similar (and additional) problems arise when elements are probabilistically dependent. *See* Allen, *supra* note 2, at 405 n.18.

<sup>174</sup>  $0.6 \times 0.6 = 0.36$ .

### B. *Proposed Solutions*

Faced with these inconsistent lines of reasoning, scholars have proposed several ways to resolve the tension. These responses fall into four main categories, each of which is discussed below. Before turning to these responses, it is important to clarify one important observation shared by most writers on the conjunction paradox. Although the probabilistic line of reasoning suggests that the law is making an egregious error in applying the burden to individual elements (and not to their conjunction), the near universal scholarly consensus is that the conjunction issue does not appear to cause any significant problems in practice.<sup>175</sup> Part of this explanation is that some cases will involve only one disputed element (or will involve two highly dependent elements). But even when several elements are disputed, there is little evidence that errors with conjunctive reasoning are causing inaccurate or unjust outcomes.<sup>176</sup>

For these reasons, theoretical disputes about conjunction are less about resolving problems in litigated cases, and more about attempting to harmonize or ameliorate these apparently inconsistent lines of reasoning. This is not to suggest, however, that the conjunction paradox is not of practical significance. Quite the contrary: as with other paradoxes, the conceptual questions raised by the conjunction paradox do force one to come to grips with serious issues. As we will see, in this instance, these issues have to do with how best to characterize the standards of proof and their requirements. And how these issues are resolved will have practical consequences for every litigated case.<sup>177</sup>

---

<sup>175</sup> See, e.g., Allen & Jehl, *supra* note 158, at 929 (“[T]he conjunction paradox probably does not have perverse effects in the real world.”); Clermont, *Death of Paradox*, *supra* note 3, at 1072 (“[T]he law currently does a much better job in structuring fact-finding than one would guess by reading its many critical commentators.”); Schwartz & Sober, *supra* note 3, at 628 (“[W]e see no basis to treat the conjunction paradox as a problem that needs to be solved through theory-driven legal reform.”). Professor Spottswood, by contrast, proposes jury instructions to correct for the conjunction effect. Spottswood, *supra* note 3, at 294 (“[T]he best way to deal with the conjunction paradox is to instruct juries to account for the reduced likelihood of conjunctions and the increased likelihood of disjunctions when deliberating.”).

<sup>176</sup> See KEVIN M CLERMONT, *STANDARDS OF DECISION IN LAW* 189 (2013) (“Admittedly, the conjunction paradox turns out to be not such a serious problem in practice.”). Moreover, the realities of litigation are in some ways the opposite of what the conjunction effect implies. As a formal matter, more elements should be to the plaintiff’s advantage because it lowers the conjunctive probability required to win. In practice, however, more elements are typically to a defendant’s advantage. See David A. Moran, *Jury Uncertainty, Elemental Independence and the Conjunction Paradox: A Response to Allen and Jehl*, 2003 MICH. ST. L. REV. 945, 946-50 (“Anyone who has ever litigated a real case knows the exact opposite of the conjunction paradox is true: the more disputed elements the plaintiff has to prove, the *less* likely the plaintiff is to prevail.”).

<sup>177</sup> See *supra* Section I.A (explaining central features of proof process and importance to overall case outcomes).

## 1. Denying Premise 1 (Elements)

One proposal for reconciling the apparent inconsistency is to deny that proving each of the individual elements is *sufficient* for satisfying the burden of proof. This interpretation in effect rejects the inconsistency between the law's practices and the probabilistic line of reasoning. In denying Premise 1 of the first line of reasoning, this proposal also rejects the conclusion that the plaintiff should win.

This proposal was advanced by Professor Dale Nance and later elaborated on by Professors David Schwartz and Elliott Sober.<sup>178</sup> The argument underlying this proposal depends on a particular interpretation of jury instructions that tell jurors "each" or "every" element must be proven to the applicable standard of proof.<sup>179</sup> They claim that such instructions are ambiguous on whether proving the elements is merely *necessary* or whether it is also *necessary and sufficient*.<sup>180</sup> They claim the first reading is not ruled out by some instructions that refer to proof of individual items, and thus the instructions permit the probabilistic line of reasoning to go forward.<sup>181</sup> According to their interpretation then, proving individual elements is not sufficient for satisfying the burden of proof. In addition to proving the individual elements, parties must also prove that the *conjunction* meets the standard (in other words, they accept Premise 1\*).<sup>182</sup>

<sup>178</sup> Dale A. Nance, *A Comment on the Supposed Paradoxes of a Mathematical Interpretation of the Logic of Trials*, 66 B.U. L. REV. 947, 950 (1986) ("[T]he simultaneous satisfaction of each condition in a set of necessary conditions is not necessarily a sufficient condition."); Schwartz & Sober, *supra* note 3, at 647-52.

<sup>179</sup> See sources cited *supra* note 16.

<sup>180</sup> In particular, they argue that this interpretation is "consistent" with some instructions but is ruled out by others. See *infra* notes 181-83.

<sup>181</sup> Schwartz and Sober give the following as an example: "For you to find the defendant guilty of this crime, you must be convinced that the government has proved each of the following beyond a reasonable doubt." Schwartz & Sober, *supra* note 3, at 682-83 (quoting COMM. ON PATTERN JURY INSTRUCTIONS, DIST. JUDGES ASS'N FIFTH CIRCUIT, PATTERN JURY INSTRUCTIONS (CRIMINAL CASES) § 2.19 (2015)). They assert: "The most natural reading of this is that the finding based on the elements is permissive—that is, meeting burden of proof on each element is necessary but not sufficient to find guilt." *Id.* at 683. Moreover, they claim other instructions employ "aggregation language" when they state language such as "Your verdict must be for the plaintiff . . . if all the following elements have been proved." *Id.* (quoting COMM. ON MODEL JURY INSTRUCTIONS FOR THE EIGHTH CIRCUIT, MANUAL OF MODEL CIVIL JURY INSTRUCTIONS FOR THE DISTRICT COURTS OF THE EIGHTH CIRCUIT § 3.04 (2017)). Finally, they note that other instructions list the individual elements for claims without specifying whether the burden of proof applies to the elements individually or the whole claim. *Id.* at 683-84 (citing CIVIL PATTERN JURY INSTRUCTIONS COMM., ALASKA CIVIL PATTERN JURY INSTRUCTIONS § 3.01 (2017)).

<sup>182</sup> *Id.* at 673 ("[M]eeting the probability threshold for the burden of proof for each element is a necessary but not sufficient condition for the claimant to win the claim.").



Although this interpretation is not ruled out by some pattern jury instructions,<sup>183</sup> it is implausible as a general description of the burden of proof. The interpretation faces two problems. First, it is inconsistent with several aspects of legal doctrine, including how courts instruct jurors in actual cases, characterize the burden, and evaluate the sufficiency of evidence.<sup>184</sup> Moving beyond pattern jury instructions, courts do not require that evidence be sufficient to meet each element *plus the conjunction*—they evaluate evidence only with regard to the elements individually.<sup>185</sup> If the proposal were true, then there would be cases where evidence was sufficient to satisfy each element but not the conjunction.<sup>186</sup> But there do not appear to be any such cases; rather, courts end

---

<sup>183</sup> Schwartz and Sober survey ninety-two pattern instructions and conclude that their preferred interpretation is ruled out by twenty-eight percent of the instructions and that the other seventy-two percent are consistent with the proposition that proving “each element” is necessary but not sufficient (i.e., plaintiffs must also prove the conjunction to the probabilistic threshold). *Id.* at 688. Their interpretation is also inconsistent with special verdicts, which do not require findings on the conjunction of elements. *Id.* In criminal cases, they find their interpretation ruled out by 33.3% of the instructions surveyed but consistent with the other 66.7%. *Id.* at 688 n.243. Logical consistency, however, is a very weak constraint. *See* Allen & Jehl, *supra* note 158, at 900 (“Consistency is not very helpful in understanding the meaning of [jury instructions]; requiring proof of each element is consistent with an infinite number of propositions that are themselves inconsistent with each other . . .”); Allen & Pardo, *supra* note 57 (manuscript at 43) (“Consistency of propositions alone carries virtually no justificatory force.”). For example, the sampled instructions are also “consistent” with consulting astrological charts to decide the cases (and thus they do not rule out the “astrological” interpretation). Moreover, their interpretation is logically inconsistent with several instructions.

<sup>184</sup> These aspects are discussed in more detail in Allen & Pardo, *supra* note 57 (manuscript at 41-46).

<sup>185</sup> *See, e.g.,* Pioneer Ctrs. Holding Co. Emp. Stock Ownership Plan & Tr. v. Alerus Fin., N.A., 858 F.3d 1324, 1337 (10th Cir. 2017) (noting plaintiff “assumed the burden of proof on each element of its claim”), *cert. dismissed*, 139 S. Ct. 50 (2018); Ga. Pac. Consumer Prods. v. Von Drehle Corp., 618 F.3d 441, 455 (4th Cir. 2010) (finding plaintiff met burden as to “each element” of their claims); Brady v. Fort Bend County, 145 F.3d 691, 710-11 (5th Cir. 1998) (accepting jury instructions requiring that plaintiff “have established each element of their claims” to relevant standard); Spierer v. Evans Inc., No. 94-50339, 1995 WL 29305, at \*1 (5th Cir. Jan. 13, 1995) (“Following a four-day trial, the district court included the following in its jury instructions: The plaintiff must prove by a preponderance of the evidence each of the following . . .”). The same is true in criminal cases. *See, e.g.,* United States v. Booker, 543 U.S. 220, 230 (2005) (noting constitutional requirement that jury find defendant “guilty of all the elements of the crime”); Jackson v. Virginia, 443 U.S. 307, 316 (1979) (“[N]o person shall be made to suffer the onus of a criminal conviction except upon sufficient proof—defined as evidence necessary to convince a trier of fact beyond a reasonable doubt of the existence of every element of the offense.”); United States v. Delgado, 672 F.3d 320, 331 (5th Cir. 2012) (“[D]ue process requires the government present evidence sufficient to prove each element beyond a reasonable doubt . . .”).

<sup>186</sup> I cannot prove the negative (i.e., that no such cases exist), but I have never found such a case. Nor have I found any cases where a party has even *argued* that the individual elements are proven but the conjunction is not.

sufficiency analysis when the elements are satisfied. Second, special verdicts require jurors to find only whether each element is satisfied; they do not also include a separate question on the conjunction.<sup>187</sup> This is significant. As a matter of legal doctrine, the burden of proof remains constant regardless of which form the verdict takes.<sup>188</sup> Because the burden applies to the *elements* in special-verdict cases, *a fortiori*, it similarly applies in cases with general verdicts.<sup>189</sup> These inconsistencies render the necessary-but-not-sufficient proposal implausible as a description of how actual burdens of proof operate in law. In other words, the proposal does not succeed in knocking out Premise 1 as an account of law's practices.<sup>190</sup>

## 2. Rejecting Premise 1 (Elements)

Rather than denying that Premise 1 describes legal practice, a second possible solution takes a normative stand and rejects Premise 1.<sup>191</sup> According to this path, the proper way to resolve the apparent tension is to fully accept the second line of reasoning, reject the first, and revise any legal practices that are inconsistent with the second line.

Although this proposal would resolve the tension, we should be cautious about adopting such a revisionary approach to legal practice. First, as noted above, there is little evidence that there are any conjunction-related problems *in practice* that need to be addressed and fixed.<sup>192</sup> At the very least, the burden should be on advocates of such a proposal to demonstrate actual problems in need of resolution along with reasons why the proposal is likely to resolve them.<sup>193</sup>

<sup>187</sup> Schwartz and Sober concede that special verdicts are inconsistent with their interpretation. Schwartz & Sober, *supra* note 3, at 674.

<sup>188</sup> See *Fritz v. Wright*, 907 A.2d 1083, 1093 (Pa. 2006); see also *Santosky v. Kramer*, 455 U.S. 745, 757 (1982) (explaining that standards of proof apply to general categories of cases and cannot be altered on case-by-case basis). Moreover, trial courts are given discretion to decide whether to use general or special verdicts. See FED. R. CIV. P. 49; CHARLES A. WRIGHT & ARTHUR R. MILLER, *FEDERAL PRACTICE AND PROCEDURE* § 2502 (3d ed. 2002), Westlaw FPP (updated Sept. 2018) (reporting that there does not appear to have ever been reversal on this ground). If the standard of proof did shift based on the choice of verdict type, then trial courts would have virtually unlimited discretion to alter the standard of proof on a case-by-case basis. Cf. *Whitman v. Am. Trucking Ass'ns, Inc.*, 531 U.S. 457, 468 (2001) ("Congress . . . does not, one might say, hide elephants in mouseholes.").

<sup>189</sup> As a formal matter, if this were not true, then it would imply that plaintiffs should always seek special verdicts and defendants should oppose them, again this is the opposite of what is true in practice. See *supra* note 176 and accompanying text.

<sup>190</sup> See *supra* notes 158-61 and accompanying text.

<sup>191</sup> See Richard D. Friedman, *Answering the Bayesioskeptical Challenge*, 1 INT'L J. EVID. & PROOF 276, 280 ("[T]he Bayesian approach would show that the law was wrong, not the other way around.").

<sup>192</sup> See sources cited *supra* note 175.

<sup>193</sup> There are a number of practical challenges facing such a change. How would jurors

Moreover, even on theoretical grounds, the proposal to reject Premise 1 faces three additional problems. First, doing so causes the burden for each element to rise in potentially troubling ways. Legislative decisions about how best to define crimes and causes of action may affect the burden of proof in arbitrary ways.<sup>194</sup> Moreover, it would produce requirements that are inconsistent with the goals of the legal system—for example, the prosecution’s burden for proving “intent” to commit first-degree murder may be much lower than the burden for proving “intent” to commit relatively minor crimes.<sup>195</sup> This would imply that the government could impose the most serious of penalties (including the death penalty) based on a lower degree of proof than it would need to impose fines for minor crimes.<sup>196</sup> Second, the non-comparative aspect of this proposal is likely to produce outcomes counter to the goals of the standards of proof regarding accuracy and the risk of error. Even when the conjunction of the plaintiff’s claim is below 0.5, it may nevertheless be considerably more likely than the conjunction of the defendant’s version (or versions) of the facts.<sup>197</sup>

The third problem concerns dependence among elements. Rather than being a way out of the conjunction paradox, recognizing that some elements are probabilistically dependent instead raises even more difficulties. When elements are dependent, factfinders need to know not only the probabilistic relationships between the evidence and the elements but also the dependence relationships among the elements. This requirement recreates similar problems that exist regarding evidence in a probabilistic framework: factfinders need numbers. Such numbers may either be objective (based on actual data) or subjective (based on the beliefs of factfinders). The first path is unworkable,<sup>198</sup> and the second path (allowing jurors and judges to make up numbers for these relationships) is unlikely to advance law’s goals regarding accurate fact-finding.<sup>199</sup>

---

know the dependence relationships among the elements? How would parties discover and present evidence on such matters?

<sup>194</sup> Adding, or disputing, elements will cause the burden of proof needed for unrelated issues to rise. See Cheng, *supra* note 3, at 1263 (“It seems odd . . . that merely disputing another element of the tort not only creates a burden on the plaintiff regarding that element, but also raises the standard by which the plaintiff must prove [other] elements at issue.”); Chris William Sanchirico, *Evidence: Theoretical Models*, in *PROCEDURAL LAW & ECONOMICS* 203, 222 (Chris William Sanchirico ed., 2d ed. 2012) (“Particularly troubling is the fact that the implied threshold probability for a charge, claim, or defense decreases (quite rapidly) in the number of elements it contains . . .”).

<sup>195</sup> See Ronald J. Allen & Brian Leiter, *Naturalized Epistemology and the Law of Evidence*, 87 VA. L. REV. 1491, 1504-05 (2001).

<sup>196</sup> *Id.* at 1505 (“To convict for theft requires on average that intent to steal be established to a higher probability than intent to kill for a murder conviction. This strikes all legal observers as both unacceptable and absurd.”).

<sup>197</sup> See Nesson, *supra* note 2, at 1390.

<sup>198</sup> See *supra* text accompanying notes 110-12.

<sup>199</sup> See *supra* text accompanying notes 115-20.

Rejecting as irrational aspects of the law that appear to clash with conventional probabilistic reasoning is a tempting proposal. The reasons above, however, suggest that this potential siren song should be avoided until the benefits of such revisions have been more clearly established.

### 3. Alternative Probabilistic Accounts

A third category of solutions rejects one or more aspects of the probabilistic line of reasoning, while otherwise retaining a probabilistic framework. Two such solutions concern Premise 3\* (Product Rule) and Premise 2\* (Standard). According to these solutions, the law's current practices can be harmonized with probabilistic reasoning when one (or both) of these premises are rejected. This Section discusses each in turn.

The most conceptually radical proposal on the conjunction issue is Professor Kevin Clermont's rejection of the product rule in the context of legal proof.<sup>200</sup> Relying on "fuzzy logic," he argues that the law's element-by-element process avoids the conjunction problem because the conjunction of multiple elements is not their product.<sup>201</sup> Instead, he relies on the "MIN rule" for combining elements.<sup>202</sup> According to this alternative rule, "the conjoined probability of elements, whether or not independent" is "equal to the least likely element."<sup>203</sup> Thus, in a two-element lawsuit in which the plaintiff proves one element to 0.6 and the other to 0.6, the probability of plaintiff's claim would be 0.6 (not .36).

Unfortunately, the conjunction effect cannot be sidestepped so easily. It is not just a theoretical construct—it reflects features of the world.<sup>204</sup> We can see this

<sup>200</sup> CLERMONT, *supra* note 176, at 191 ("The truth of the conjunction equals the minimum of the truths of the elements.").

<sup>201</sup> *Id.* at 159 ("I propose considering the broad version of fuzzy logic as the legal model for human expression of uncertainty, in preference to assuming a probability-based bivalent view."); *see, e.g.*, Timothy J. Ross & W. Jerry Parkinson, *Fuzzy Set Theory, Fuzzy Logic, and Fuzzy Systems*, in *FUZZY LOGIC AND PROBABILITY APPLICATIONS* 29, 31 (Timothy J. Ross, Jane M. Brooker & W. Jerry Parkinson eds., 2002); L.A. Zadeh, *Fuzzy Sets*, 8 *INFO. & CONTROL* 338, 338-39 (1965). The framework underlying Clermont's proposal raises a host of complicated technical questions that are outside the scope of the discussion below. For a discussion and critique of Clermont's reliance on the framework, see Ronald J. Allen, *The Declining Utility of Analyzing Burdens of Proof*, 48 *SETON HALL L. REV.* 995, 1003-15 (2018).

<sup>202</sup> CLERMONT, *supra* note 176, at 172. The MIN operator rule provides a rule of conjoining different "fuzzy sets." The sets have degrees of membership. For example, suppose "Tom is a 0.3 member in the set of tall men" (this means he is not very tall) and "Tom is a 0.4 member in the set of smart men" (this means he is not very smart, below average at least). The MIN operator rule tells us that Tom is a 0.3 member in the set of tall *and* smart men because this is the "least likely element." *Id.* at 171. The product rule, by contrast, would dictate that the probability that Tom is tall and smart is 0.12 (0.3 x 0.4).

<sup>203</sup> *Id.* at 172.

<sup>204</sup> *See* Allen, *supra* note 201, at 1004 ("[T]he conjunction effect is a feature of the world, not of our logic or mathematics. It does not change as we embrace new or different ways of thinking.").

most clearly with events like coin flips: the probability of getting two (or more) heads in a row is not 0.5.<sup>205</sup> A similar phenomenon exists for (at least some) legal elements.<sup>206</sup> Suppose a tort claim depends on two contested elements: the identity of the person responsible for an accident, and whether the accident caused plaintiff's injuries. Suppose the jury thinks each element is 0.6 probable. Let us focus on the first element, identity. This 0.6 probability that the defendant is the perpetrator of the accident (and thus the 0.4 probability someone else did it) must interact with the second element, causation. And it must do so in one of two ways: either the accident caused the injuries or it did not. Here the probabilities are 0.6 that it caused the injuries (and 0.4 that it did not). The only way that the probability of both being true is 0.6—as Clermont's theory declares—is if the accident could have caused the injuries only when the defendant is the perpetrator, and vice versa. To put this another way, when assessing the combination of these elements, there are two possibilities that must be considered: (1) the accident caused the injuries *but* someone else caused the accident, and (2) the defendant caused the accident *but* the accident did not cause the injuries. The only way the conjunction is 0.6 is if the probability of each of these two possibilities is 0. This would be like saying that flipping a "heads" the first time guaranteed a "heads" the second time (and that a "heads" the second time was only possible if it was "heads" the first time). This may well be true with some legal elements, but it cannot be declared by fiat for all possible combinations. The reality underlying the product rule will not go away by accepting fuzzy logic. And, therefore, the underlying features that produce the "paradoxical" lines of reasoning will exist even if one employs "least likely element" rule.

A second possibility for tinkering with the probabilistic line of reasoning is to question Premise 2\*. This premise also interprets the preponderance rule as a threshold of 0.5—but if the preponderance rule means something else, the inconsistent conclusions in the two lines of reasoning no longer necessarily follow. Two recent arguments suggest this possibility for resolving the conjunction problem while otherwise retaining a probabilistic framework. The first is also offered by Clermont,<sup>207</sup> and the second offered by Professor Edward

---

<sup>205</sup> See *supra* note 164.

<sup>206</sup> Fuzzy logic and fuzzy sets work well in explaining problems having to do with vagueness, and Clermont's framework may have utility in explaining some elements that require the application of vague legal terms (e.g., "reasonable") to other factual elements. See Allen & Pardo, *supra* note 57 (manuscript at 70) ("[T]hinking of the meaning of vague legal terms from the perspective of fuzzy set theory is interesting and perhaps useful.")

<sup>207</sup> Although this aspect and the one discussed previously are related in Clermont's framework, the two aspects are conceptually distinct (i.e., it is possible to accept one without accepting the other) and they respond to different features of the conjunction problem. See CLERMONT, *supra* note 176, at 150. This second aspect of Clermont's account does not alter the critique in the previous paragraph above.

Cheng, who conceptualizes the standards of proof based on the comparative likelihood of plaintiff's and defendant's stories.<sup>208</sup>

Clermont relies on the notion of "belief functions" to argue against the "beyond 0.5" interpretation of the preponderance rule.<sup>209</sup> In doing so, Clermont rejects a second rule of conventional probabilistic reasoning—the complementation axiom, providing that a proposition (*A*) and its negation (not-*A*) must sum to 1.<sup>210</sup> Under this rule, if *A* is 0.6 probable, then not-*A* is 0.4 probable. We assumed this, for example, when discussing the examples above: proving each element to 0.6 meant that the negation of that element was 0.4 probable. Clermont argues this assumption should also be rejected.<sup>211</sup> Instead, for each element there are three possibilities for a factfinder's belief: belief in *A*, belief in not-*A*, and uncertain belief.<sup>212</sup> This third category, under Clermont's theory, favors neither party. Under this interpretation, the preponderance rule requires only that belief in *A* outweigh not-*A*, regardless of whether belief in *A* surpasses 0.5.<sup>213</sup> So, for example, if the factfinder concludes *A* is 0.3 probable, not-*A* is 0.2, and the remaining 0.5 is uncertain, then plaintiff has satisfied the preponderance rule.<sup>214</sup>

In terms of the law's goals, this move toward a more comparative approach is a step in the right direction. As we saw above, the "0.5" rule places all the risk of unknown possibilities onto the plaintiff and appears to be inconsistent with the goals of accuracy and equalizing the risk of error.<sup>215</sup> However, this interpretation has other difficulties. First, it ultimately must rely on subjective beliefs and therefore faces similar problems as other probabilistic approaches.<sup>216</sup> Second, the comparison between belief in *A* and belief in not-*A* applies to each

---

<sup>208</sup> Cheng, *supra* note 3, at 1259 ("[T]he preponderance standard is better characterized as a probability ratio, in which the probability of the plaintiff's story of the case is compared with the defendant's story of the case.").

<sup>209</sup> CLERMONT, *supra* note 176, at 149 (citing GLENN SHAFER, *A MATHEMATICAL THEORY OF EVIDENCE* (1976)). The aspect of "belief functions" on which Clermont relies is the idea that belief in a proposition and its negation need not sum to 1. *Id.* at 203 ("Under the scheme of belief functions, [a proposition and its negation] do not necessarily add to 1, because normally some belief remains uncommitted."). In other words, someone may believe that *A* is 0.4 likely and not-*A* is 0.3 likely, with the remaining 0.3 unaccounted for.

<sup>210</sup> See *supra* notes 164, 209 and accompanying text.

<sup>211</sup> CLERMONT, *supra* note 176, at 219 ("All the factfinder must do is compare belief and disbelief . . . while some of the factfinder's belief remains uncommitted . . .").

<sup>212</sup> *Id.*

<sup>213</sup> *Id.* ("[A]ll that preponderance of the evidence requires is that the strength of the factfinder's belief that Tom is at fault must exceed his belief that Tom was not at fault.").

<sup>214</sup> *Id.*

<sup>215</sup> See *supra* note 38 and accompanying text.

<sup>216</sup> See *supra* text accompanying notes 115-20.

element, not the case a whole, and thus recreates the conjunction effect (because of the features of reality reflected by the product rule).<sup>217</sup>

Cheng's proposal is more explicitly comparative. Under his account, the preponderance rule requires that plaintiff's story be more likely than defendant's alternative story.<sup>218</sup> This interpretation thus rejects the "0.5 rule"—the plaintiff's story may be more likely than the defendant's story, even though it is lower than 0.5.<sup>219</sup> In rejecting Premise 2, this interpretation avoids the requirement that the conjunction of elements must exceed 0.5. For example, in our two-element example, even if the plaintiff's case is only 0.36 probable (the conjunction of 0.6 and 0.6), this may still be more likely than the defendant's alternative on these elements ( $0.4 \times 0.4 = 0.16$ ).<sup>220</sup> Although this move toward comparative assessments, similar to Clermont's, appears to avoid the inconsistency that generates the conjunction paradox, it too faces conceptual difficulties. First, by operating within an explicitly probabilistic framework, it must rely on subjective probabilities and the consequent problems.<sup>221</sup> Second, because it applies element-by-element rather than to the case as a whole, it also potentially recreates the conjunction problem it was trying to avoid.<sup>222</sup>

In sum, one approach for resolving the conjunction problem is to reconsider one or more premises in the probabilistic line of reasoning, while otherwise remaining within a broadly probabilistic framework. Although solutions of this type face difficulties, they suggest two important points. First, a move toward more comparative assessments appears to better fit the goals of the rules. And second, Premise 2 (and Premise 2\*)—that the preponderance standard means a probabilistic threshold of 0.5—is a potentially weak link in the probabilistic line of reasoning.

---

<sup>217</sup> For example, suppose the factfinder concludes that the plaintiff has proven two elements to 0.3 and that the negation of these two elements is each 0.2. The plaintiff wins under Clermont's account even though the plaintiff's claim is 0.09 probable ( $0.3 \times 0.3$ ), which is lower than the factfinder's disbelief in either element (0.2).

<sup>218</sup> See Cheng, *supra* note 3, at 1259.

<sup>219</sup> *Id.* (characterizing assumptions underlying 0.5 probability threshold as "sharply at odds with current legal practice").

<sup>220</sup> See *id.* at 1265; Nesson, *supra* note 2, at 1389-90.

<sup>221</sup> See *supra* text accompanying notes 115-20.

<sup>222</sup> See Cheng, *supra* note 3, at 1264 (explaining that under his comparative probability framework, defendants win by showing their account is more likely on any element). *But see* Ronald J. Allen & Alex Stein, *Evidence, Probability, and the Burden of Proof*, 55 ARIZ. L. REV. 557, 598 (2013) ("[W]hen the plaintiff's suit has two elements with probabilities amounting to 0.9 and 0.4, and the defendant's probabilities are their reciprocals, the defendant will win the case. The plaintiff thus gets no credit for his overwhelming advantage on the first element of the suit (0.9 against 0.1). A genuinely comparative system, however, should give this credit to the plaintiff. The plaintiff's overwhelming advantage on the first element makes the overall probability of his case ( $0.9 \times 0.4 = 0.36$ ) six times higher than the overall probability of the defendant's case ( $0.1 \times 0.6 = 0.06$ ).").

#### 4. The Explanatory Account

The explanatory account also rejects Premise 2 (and 2\*). Moreover, it does so in a manner that both ameliorates the problems created by the conjunction of elements and avoids the problems faced by other possible solutions.

The explanatory account explains the standards of proof in terms of explanatory (as opposed to probabilistic) thresholds.<sup>223</sup> According to this account, factfinders do not assign probabilities to individual elements; rather, they evaluate competing explanations of the evidence and events as a whole, with differing thresholds corresponding to different standards of proof.<sup>224</sup> For example, under the preponderance standard, factfinders infer whether the best of the available explanations favors the plaintiff or the defendant. In a typical civil case, the plaintiff will offer an explanation that includes all of the elements, and the defendant will offer an alternative, contrasting explanation that omits one or more of the elements.<sup>225</sup> This inferential process ameliorates the potential problems created by the conjunction effect by assessing the strength of cases as a whole and by applying to both parties' explanations.<sup>226</sup> The pernicious consequences identified by the conjunction paradox are ameliorated not merely because the law does not require factfinders to make probabilistic judgments—this by itself would be a facile way to avoid the problem. Rather, it is because the holistic process of explanation evaluation, including distributing any conjunction effects to both sides' explanations, brings the law's practices in line with the probabilistic account's case-as-a-whole assumption about how to achieve the law's underlying goals.<sup>227</sup> In other words, the explanatory account explains how Premise 1 and Premise 1\* cohere.

Moreover, in rejecting Premise 2 (and 2\*), the explanatory account better fits with legal doctrine and the practices of judges, juries, and parties than

---

<sup>223</sup> See *supra* Section I.B.2 (discussing the explanatory account).

<sup>224</sup> See *supra* text accompanying notes 77-82. In other words, factfinders select (or construct) an explanation in light of the standard of proof and then compare that explanation with the formal legal elements.

<sup>225</sup> If the factfinder selects the plaintiff's explanation, or another explanation that includes all of the legal elements, then the plaintiff has satisfied the burden of proof. If the selected explanation omits one or more of the elements, then the plaintiff has failed to satisfy the burden of proof.

<sup>226</sup> In other words, the proof process is not one in which one side argues for each element as an independent proposition (*A*, *B*, etc.) and the other side argues for the negation of these propositions (not-*A*, not-*B*, etc.). Rather, each side typically offers an explanation that includes or omits these propositions, and any conjunction effects are distributed among both parties' explanations. To be clear, nothing in the account prevents parties from relying on alternative explanations or disjunctive explanations that attempt to aggregate alternatives when they believe it is to their advantage to do so. See Pardo, *supra* note 21, at 1831 (explaining that explanatory account operates "at the level of cases as a whole" and by "selecting among competing explanations").

<sup>227</sup> See *supra* note 38 and accompanying text.



probabilistic accounts.<sup>228</sup> It also avoids the problems that affect other possible solutions. First, it is comparative and thus better fits with the goals of the rules.<sup>229</sup> Second, in comparing explanations as a whole, factfinders conform to the assumption that, to achieve their stated goals, the standards of proof must apply to cases as a whole (and not just individual elements).<sup>230</sup> Third, the approach approximates these goals without the need to calculate individual probabilities or to know the dependence relationships among the elements.<sup>231</sup> Fourth, it allows for assessment of cases as a whole without requiring the standard for individual elements to rise (or fall) in arbitrary and otherwise problematic ways based on the number of other contested elements.<sup>232</sup> Fifth, it does not reject the product rule.<sup>233</sup> Because conjunctions are not calculated, it has no need for the rule in deciding cases. But, importantly, it does not ignore or deny the underlying features of reality. The explanatory reasoning posited by the approach already accounts for these features.<sup>234</sup> Sixth, unlike solutions that rely on subjective beliefs, the explanatory account relies on more objective relationships between evidence and explanations.<sup>235</sup>

### C. *The Significance of the Conjunction Paradox*

The significance of the conjunction paradox concerns what it reveals about standards of proof in general. The inconsistency between the two lines of reasoning reveals that something is amiss in how the law's practices are conceptualized. The practices themselves do not appear to be creating problems; rather, what creates conceptual problems is the idea that factfinders apply probabilistic thresholds to individual elements. The explanatory account ameliorates these problems because of two features: explanations are evaluated as a whole, and the process is comparative. Accepting the explanatory account

---

<sup>228</sup> See Allen & Pardo, *supra* note 57 (manuscript at 16-18) (describing how explanatory account better fits policy goals and realities of legal practice).

<sup>229</sup> See *supra* note 38 and accompanying text. The explanatory account thus has the benefits of Clermont's and Cheng's proposals, while avoiding the conceptual problems underlying them.

<sup>230</sup> See *supra* note 171 and accompanying text.

<sup>231</sup> See Williamson, *supra* note 76, at 267 (explaining how inference to best explanation can be "good heuristic to use when—as often happens—probabilities are hard to estimate" and "may be the closest we can get to probabilistic epistemology" in such situations).

<sup>232</sup> See *supra* text accompanying notes 194-96.

<sup>233</sup> See Allen & Pardo, *supra* note 57 (manuscript at 33-34). Rather, it distributes the consequences of conjunction among parties' alternative explanations.

<sup>234</sup> But see Spottswood, *supra* note 3, at 292-93 (suggesting improvements in jury instructions).

<sup>235</sup> Explanatory criteria refer to features of evidence that are distinct from subjective beliefs. See Pardo, *supra* note 10, at 605 (explaining that "there are objective ways to examine explanatory relationships between evidence and competing propositions").

of standards of proof requires giving up the “0.5” interpretation in particular and the notion that the standards express probabilistic thresholds in general.<sup>236</sup>

The conjunction paradox forces one to come to grips with serious issues: what do the standards of proof mean, and what do they require in individual cases? These issues are of tremendous practical importance throughout both civil and criminal litigation. The issues have consequences for every civil and criminal trial—and, indeed, they affect issues throughout the litigation processes.<sup>237</sup> Every application of a standard of proof requires some assumption (implicit or explicit) about what the standard means and requires. Whether one adopts (or presupposes) a probabilistic or explanatory conception will therefore inform and affect those applications. The conjunction paradox helps to make the implications of that choice explicit. The explanatory account explains how the law accommodates the conjunction effect. The probabilistic accounts, by contrast, face conceptual difficulties in trying to explain the law.<sup>238</sup>

How one understands burdens of proof also has implications for the statistical-evidence paradox.<sup>239</sup> A central assumption generating “paradoxical” results in the statistical-evidence examples is that the proof standards are probabilistic thresholds (surpassed on their face by the data proffered as evidence). The conjunction paradox, however, provides reasons to reject that assumption. If we reject that assumption, then the “paradoxical” results in the statistical-evidence examples are similarly ameliorated. Explanatory thresholds,

---

<sup>236</sup> See Cheng, *supra* note 3, at 1279 (arguing that “courts and attorneys [should] stop using the misleading 0.5 rule as a shorthand for the preponderance standard”).

<sup>237</sup> See *supra* Section I.A (explaining primacy of central features of proof process in litigation). One example are toxic tort cases in which courts have required epidemiological evidence showing a “relative risk” of 2.0 or greater in order to have sufficient evidence to meet the preponderance standard (understood as a 0.5 threshold), excluding experts and/or granting summary judgment when plaintiffs fail to produce such evidence. See, e.g., *Henricksen v. ConocoPhillips Co.*, 605 F. Supp. 2d 1142, 1158 (E.D. Wash. 2009) (stating that epidemiology studies can be probative of specific causation if relative risk is greater than 2.0); *Cook v. United States*, 545 F. Supp. 306, 308 (N.D. Cal. 1982) (“Whenever the relative risk to vaccinated persons is greater than two times the risk to unvaccinated persons, there is a greater than [fifty-percent] chance that a given GBS case among vaccinees of that latency period is attributable to vaccination, thus sustaining plaintiff’s burden of proof on causation.”); *Watts v. Radiator Specialty Co.*, 990 So. 2d 143, 150 (Miss. 2008); Russellyn S. Carruth & Bernard D. Goldstein, *Relative Risk Greater than Two in Proof of Causation in Toxic Tort Litigation*, 41 *JURIMETRICS* 195, 197-99 (2001) (listing cases discussing role of a relative risk of greater than 2.0 in toxic tort cases); Steve C. Gold, *The “Reshaping” of the False Negative Asymmetry in Toxic Tort Causation*, 37 *WM. MITCHELL L. REV.* 1507, 1523 (2011) (“[C]ourts have equated more than a doubling of relative risk in an exposed group to a more-likely-than-not probability of causation in an exposed individual plaintiff.”).

<sup>238</sup> This is one of several reasons why the explanatory account provides a better general explanation of legal proof. See Allen & Pardo, *supra* note 57 (manuscript at 9) (“[T]he conventional probabilistic account faces several problems that make it empirically implausible as a general explanation of juridical proof.”).

<sup>239</sup> See *supra* Part II (discussing statistical-evidence paradox).

however, are not as precise as probabilistic ones. Thus, the former may not generate clear answers in the (highly unrealistic) fact patterns.<sup>240</sup>

This state of affairs points toward a diagnosis. If actual standards of proof are better explained as explanatory thresholds, as this Part argues, then this potential uncertainty in the statistical-evidence examples provides a reason why the examples both continue to resist consensus about how best to resolve them and why they continue to fascinate. In other words, the examples misleadingly draw attention to the *form* of the evidence, but this is not a meaningful distinction for legal evidence.<sup>241</sup> Much more important are assumptions about what the standards of proof mean and require. Until one comes to grips with those issues, the statistical-evidence paradox will resist resolution. And even if one does resolve those issues, the examples may nevertheless remain underdetermined because of the nature of the standards and the limited amount of information in the examples.

#### IV. THE VERDICT PARADOX

A third paradox has received considerably less attention in the evidence literature.<sup>242</sup> Discussions on the statistical-evidence and conjunction paradoxes, as with discussion of legal evidence and proof more generally, typically proceed from the assumption that the “factfinder” is a single, unified actor.<sup>243</sup> The examples above, for example, all assumed that the factfinder (judge or jury) did or did not find particular facts proven to the applicable standard of proof. This assumption makes sense when judges act as factfinders and it often helps to simplify otherwise complex discussions (such as those above) by assuming the jury members agree on their findings. Sometimes, however, jurors disagree on aspects of cases, and these disagreements can also produce seemingly paradoxical results.<sup>244</sup>

---

<sup>240</sup> See Callen, *supra* note 128, at 1530-32 (discussing problems with drawing conclusions from evidence in statistical-evidence hypotheticals).

<sup>241</sup> See *supra* Section II.C (arguing that these paradoxes do not support meaningful distinction between statistical and individualized evidence).

<sup>242</sup> See Pardo, *supra* note 21, at 1845 (“The issue of juror agreement has received more attention in the context of criminal cases, from both courts and academic commentators, than in civil cases. Nevertheless, considerable confusion remains.”).

<sup>243</sup> *Id.* at 1796 (“[T]he jury does indeed possess a type of group agency that is distinct from the agency of its individual members.”).

<sup>244</sup> In addition to disagreeing on whether specific elements are proven, jurors may also disagree about the underlying reasons why the elements are proven (while agreeing on whether the elements are proven). The discussion below focuses on the former type of disagreement. On the latter type of disagreement, see *Richardson v. United States*, 526 U.S. 813, 815, 824 (1999) (finding that conviction under a criminal statute required juror unanimity for each individual violation, not just a determination that series of violations occurred); *Schad v. Arizona*, 501 U.S. 624, 630 (1991) (declining to require juror unanimity about theory behind capital conviction); *Westen & Ow*, *supra* note 22, at 160 (“[T]he Constitution

One such paradox can arise in cases that do not require unanimous verdicts. The source of this paradox is not unique to legal proof; rather, the jury context is just one example of a larger issue that can arise whenever a decision-making group faces multiple, related issues.<sup>245</sup> As with the statistical-evidence and conjunction paradoxes, the “paradoxical” results in the verdict example arise because of inconsistent lines of reasoning. Moreover, as with the other categories, an underlying source of problems is the probabilistic interpretation of the standards of proof. Rejecting this assumption ameliorates the paradoxical consequences in the verdict context, just as it does with the conjunction paradox.

Some jurisdictions allow for verdicts without juror unanimity.<sup>246</sup> The details vary considerably based on jurisdiction, type of case, and consent of the parties, and are subject to a few constitutional limitations.<sup>247</sup> The details are not important for illustrating the paradox. Rather, it is simply the *possibility* of such verdicts that produces the potential for paradoxical consequences. Non-unanimous verdicts may generate such consequences because juror votes on individual elements may produce general verdicts that jurors would otherwise reject.<sup>248</sup>

For example, consider a crime or cause of action consisting of four elements (*A*, *B*, *C* & *D*) and a verdict rule that requires an affirmative vote of a supermajority of nine of twelve jurors.<sup>249</sup> *These conditions could produce a verdict that every juror rejects.* This could arise when jurors 1-3 reject element

---

sometimes does and sometimes does not require that jurors concur on the alternative means by which defendants commit offenses.”).

<sup>245</sup> See CHRISTIAN LIST & PHILIP PETTIT, GROUP AGENCY: THE POSSIBILITY, DESIGN, AND STATUS OF CORPORATE AGENTS 92 (2011) (“[M]ajority voting may fail to ensure rational group judgments on interconnected propositions . . . .”); Lewis A. Kornhauser & Lawrence G. Sager, *The One and the Many: Adjudication in Collegial Courts*, 81 CALIF. L. REV. 1, 58 (1993) (“[C]ases are paradoxical when rationales and outcomes are set in conflict with each other.”).

<sup>246</sup> See ROTTMAN & STRICKLAND, *supra* note 23, at 233-37 tbl.42.

<sup>247</sup> See *id.*

<sup>248</sup> See *Gourley v. Neb. Methodist Health Sys., Inc.*, 663 N.W.2d 43, 57 (Neb. 2003) (“[A] juror who disagreed with the determination of who was liable provided the 10th vote necessary to decide the amount of damages and how to apportion the defendants’ negligence.”); *Williams v. James*, 552 A.2d 153, 160-61 (N.J. 1989) (reversing non-unanimous general jury verdict where jury instructions failed to explain that jurors need not vote consistently on each question); *O’Connell v. Chesapeake & Ohio R.R. Co.*, 569 N.E.2d 889, 898 (Ohio 1991) (vacating judgment in which jurors who found neither party negligent also participated in apportionment of comparative negligence, rendering general verdict “wholly unreliable”); *Fritz v. Wright*, 907 A.2d 1083, 1086 (Pa. 2006) (finding inconsistency between general verdict and answers of some jurors on interrogatories).

<sup>249</sup> See ROTTMAN & STRICKLAND, *supra* note 23, at 233-36 (listing states with verdict rules requiring supermajority of jurors). To the extent the Supreme Court would draw the constitutional line at ten of twelve jurors, see *supra* note 49 and accompanying text, then this same result (i.e., a guilty verdict that every juror rejects) could arise with five elements (and groups of two jurors each rejecting a different element).

A but not the others, 4-6 reject element *B*, 7-9 reject element *C*, and 10-12 reject element *D*. Each element would be proven to the requisite nine of twelve, but every juror concluded that the defendant is not guilty or not liable.

The mismatch between conclusions on individual elements and overall conclusions is an example of a general decision-making problem, of which verdicts are but one example.<sup>250</sup> Whenever judgments about individual propositions are aggregated based on a non-unanimous voting rule, conclusions are possible that a majority (or all) individuals reject. This general feature plays out in the proof context because standards of proof apply to individual elements and not to claims as a whole.<sup>251</sup>

In the verdict context, the following simplified example illustrates the potential paradox. Suppose a civil claim in which the plaintiff must prove two elements, *A* and *B*, by a preponderance of the evidence.<sup>252</sup> Moreover, suppose a verdict requires nine of twelve jurors to vote in the affirmative. In the example, suppose six jurors conclude that *A* and *B* are each proven. Of the remaining six, three conclude that *A* is proven but *B* is not, and three others conclude that *B* is proven but *A* is not. On one hand, nine jurors have concluded that *A* is proven and nine have concluded that *B* is proven, so it appears that the plaintiff should win.<sup>253</sup> On the other hand, only six jurors have concluded that the plaintiff has proven its case, and six others have concluded that the plaintiff has not, so it appears that a verdict has not yet been reached.<sup>254</sup> These conflicting lines of reasoning are spelled out below.

According to one line of reasoning:

**Premise 1 (Elements):** the plaintiff must prove each element by a preponderance of the evidence.

**Premise 2 (Standard):** the preponderance standard requires a probabilistic threshold of greater than 0.5.

---

<sup>250</sup> See LIST & PETTIT, *supra* note 245, at 92 (“A member of a tenure committee, for example, may be better at making a correct judgment on each of the separate premises about teaching and research than on the overall conclusion about tenure.”).

<sup>251</sup> See cases cited *supra* note 16.

<sup>252</sup> For example, suppose *A* is identity (i.e., the defendant is the person responsible for the alleged conduct) and *B* is causation (i.e., the alleged conduct caused the plaintiff’s injuries, which are not disputed).

<sup>253</sup> See, e.g., *Gourley*, 663 N.W.2d at 60 (“Even though a juror, who disagreed on the question of who was liable, provided the 10th vote necessary on the damages and apportionment questions, the verdict was valid.”).

<sup>254</sup> See, e.g., *O’Connell v. Chesapeake & Ohio R.R. Co.*, 569 N.E.2d 889, 898 (Ohio 1991) (“Only *six jurors* found the appellant seventy-percent negligent in causing her own injuries . . . . Because jurors Bryson and Hall should have been disqualified from taking part in the apportionment question, their votes as to apportioning fault are invalid. As such, fewer than three-fourths of the eight jurors agreed to the apportionment of fault. The Ohio Constitution requires the concurrence of three-fourths of the jury for a valid verdict.”).

**Premise 3 (Aggregation Rule):** A verdict requires agreement from nine out of twelve jurors.

**Premise 4 (Example):** six jurors find that *A* and *B* are each proven; three jurors find that *A* is proven but *B* is not; and three other jurors find that *B* is proven but *A* is not.

**Conclusion:** a verdict has been reached and the plaintiff wins.<sup>255</sup>

According to a second line of reasoning:

**Premise 1\* (Claim):** to carry out its policy goals, the standard of proof requires the plaintiff to prove their claim (i.e., the *conjunction* of the elements) by a preponderance of the evidence.<sup>256</sup>

**Premise 2\* (Standard):** the preponderance standard requires a probabilistic threshold of greater than 0.5.

**Premise 3\* (Aggregation Rule):** a verdict requires agreement from nine out of twelve jurors.

**Premise 4\* (Example):** six jurors find that *A* & *B* are proven, and six jurors find that *A* & *B* are not proven.

**Conclusion:** no verdict has been reached.

One source of tension is between Premise 1 and Premise 1\*. Notice that this is similar to the tension that arises for the conjunction paradox.<sup>257</sup> Accordingly, similar issues arise as well as possible solutions. In particular, proving individual elements serially may produce results inconsistent with the policy goals of the standards of proof when coupled with probabilistic interpretations of the standards (Premise 2 and Premise 2\*). Rejecting Premise 2 (and Premise 2\*) and adopting the explanatory conception, by contrast, ameliorates the paradoxical effects.

To illustrate these effects in the verdict context, consider the two-element example above. It might be thought that the first line of reasoning (and the probabilistic conception it contains) will maximize accuracy and thus better accord with the goals of the standard of proof.<sup>258</sup> Because a majority of jurors concludes each element surpasses the threshold for each element, so the thought goes, the elements are each more likely to be true than false and thus the plaintiff ought to win (regardless of the differences among some jurors on the elements and their views about the claim as a whole).<sup>259</sup> This argument, however, ignores the conjunction effect.<sup>260</sup> Adopting the first line of reasoning, and the

<sup>255</sup> Similar considerations apply to affirmative defenses on which defendants have the burden of proof.

<sup>256</sup> See *supra* note 171 and accompanying text.

<sup>257</sup> See discussion *supra* Part III (discussing conjunction paradox).

<sup>258</sup> See *supra* note 38 and accompanying text.

<sup>259</sup> See sources cited *supra* note 38.

<sup>260</sup> See *supra* text accompanying notes 162-70.

probabilistic interpretation of the standard, is likely to frustrate the goals of the standards of proof (as with the conjunction paradox).

For example, suppose that the six jurors who conclude that *A* and *B* are each proven all find that *A* and *B* are each 0.6 likely. Assuming independence,<sup>261</sup> the plaintiff's claim is only 0.36 likely (considering only the findings by these six jurors).<sup>262</sup> What about the other six jurors? Because they each think that one of the two elements is not proven, we know that they all think the plaintiff's claim is not greater than 0.5.<sup>263</sup> Thus, as with the other six jurors, they also conclude that the conjunction of *A* & *B* does not meet the probabilistic threshold.<sup>264</sup> Thus, the probabilist approach to standards of proof, when applied to individual elements, again (as with the original conjunction problem) implies verdicts that are inconsistent with the goals of the rules.

The explanatory account addresses the verdict paradox by rejecting the idea that the standard is a probabilistic threshold (Premises 2 and 2\*). Instead, the standard is better explained as an explanatory threshold. Under this conception: (1) the standard applies to the case as a whole,<sup>265</sup> and (2) the selected explanation is then compared with the verdict categories.<sup>266</sup> How does this account handle the above example? The fundamental proof issue is whether the plaintiff's explanation, which includes (or entails) *A* & *B*, is better than any defense explanations.<sup>267</sup> Viewing the example through this lens, six jurors accept an explanation that includes *A* and *B*, and six jurors reject this explanation.<sup>268</sup> Thus,

---

<sup>261</sup> As with the conjunction paradox, this assumption is merely for the sake of simplicity. For example, suppose the six jurors conclude that it is 0.6 probable that the defendant is responsible for the alleged conduct and that it is 0.6 probable that the alleged conduct caused the plaintiff's injuries.

<sup>262</sup> Moreover, it might be assumed that adding three additional votes for each element will not change this result. This assumption, however, raises a number of complex issues. If the conditions for the Condorcet voting theorem were satisfied, then adding additional votes might raise the overall likelihood of the conclusions being true. See Levmore, *supra* note 70, at 734-39 (discussing relationship between product rule, Condorcet theorem, and likelihood of jury conclusions). The theorem states conditions under which increasing group size also increases the probability that the majority's decision will be true; however, for reasons why these conditions (most importantly, the independence of each individual's decision) do not apply to modern trials, see Allen & Jehl, *supra* note 158, at 906-14.

<sup>263</sup> Because of the conjunction effect, even if the jurors conclude the other element is 1.0 probable, then the maximum probability for the claim as a whole would be 0.5 (0.5 x 1.0).

<sup>264</sup> Indeed, a conjunctive probabilistic approach implies that the defendant should win even though nine of twelve conclude each element is proven.

<sup>265</sup> As with the conjunction paradox, this feature ameliorates the conjunction effect and distributes its consequences between the parties. See Allen & Pardo, *supra* note 57 (manuscript at 18).

<sup>266</sup> See *supra* note 224.

<sup>267</sup> See *supra* note 224. For example, suppose the two elements at issue are identity and causation.

<sup>268</sup> Note that these jurors may reject this explanation for different reasons: three reject *A*

a verdict has not yet been reached.<sup>269</sup> This resolution appears to better accord with the goals of the rules.<sup>270</sup> Once again, as with the original conjunction paradox, rejecting a probabilistic approach ameliorates paradoxical consequences.<sup>271</sup>

The significance of the verdict paradox, as with the conjunction paradox, is that it raises fundamental questions about the standards of proof—what they mean and what they require. Different conceptions of the standards imply conflicting results about whether verdicts have been reached. The probabilistic account predicts results that are inconsistent with the goals of the rules.<sup>272</sup> By contrast, the explanatory account explains standards of proof in a manner that better fits with these goals. Thus, as with the conjunction paradox, the verdict paradox provides further support for the explanatory account—not only for

---

and three reject *B*.

<sup>269</sup> At this point, no explanation has been accepted by nine jurors.

<sup>270</sup> As noted above, allowing verdicts in the face of such disagreement is likely to frustrate the goals of the rules. Because the explanatory threshold applies to the claim as a whole, it better aligns the decision-making process with the goals of accuracy and allocating the risk of error. *See supra* note 38 and accompanying text. Thus, until the requisite number of jurors coalesce around an explanation, then the verdict requirements have yet to be satisfied. To be clear, nothing in this account requires jurors to proceed in any particular manner when juror votes on particular issues produce results that diverge from juror votes on overall conclusions. Jurors—like multi-member appellate courts—are generally free to structure deliberations in whatever manner they conclude makes more sense, as least in general verdict cases. *See, e.g.,* *Fritz v. Wright*, 907 A.2d 1083, 1093, 1096 (Pa. 2006) (“These questions permit a jury unlearned in the law to frame and structure their deliberations, if they so choose . . . . The fact that two jurors dissented on one of the preceding interrogatories, in effect disputing nothing more than the path the jury followed to reach the consensus, is irrelevant to the fact that ten jurors agreed on the final verdict.”); *cf* *LIST & PETTIT*, *supra* note 245, at 88–102 (discussing criteria that may affect structure of jury deliberations); Kornhauser & Sager, *supra* note 245, at 30–33 (endorsing “meta-rule” in which multi-members decide for themselves whether to proceed issue-by-issue or by overall conclusion when two paths produce different results). Special verdicts may create additional problems. *See* discussion *infra* note 271.

<sup>271</sup> As with the conjunction paradox, the explanatory account in effect ameliorates the tension between Premise 1 and Premise 1\*. Jason Iuliano illustrates how special verdicts and non-unanimous verdicts might produce outcomes that jurors did not intend. *See generally* Iuliano, *supra* note 3. His proposed solution, however—which relies on “conditional voting” that binds juror votes on subsequent issues based on their previous answers—relies on a probabilistic conception that recreates the conjunction problem as well as other problems affecting such accounts. *See* Pardo, *supra* note 21, at 1832–33 n.156 (critiquing Iuliano’s proposed solutions). The potential problem with special verdicts in such circumstances is that they may force jurors down a decision-making path that they may not otherwise choose. *See supra* note 270. Instructions explaining the implications of their answers may help to reinsert more decision-making freedom into their deliberations.

<sup>272</sup> Other inconsistencies between the probabilistic conception and verdict requirements are discussed in Pardo, *supra* note 21, at 1857 (discussing inconsistencies and concluding that “[t]urning to the group-level perspective reveals a number of conceptual problems for a probabilistic conception of legal evidence and proof”).



examples that give rise to the verdict paradox but as a general account of the standards of proof whenever they are employed throughout the law.

#### CONCLUSION

Paradoxes force one “to come to grips with key issues.”<sup>273</sup> The histories of mathematics, science, and philosophy provide numerous examples. Perhaps most prominently, Zeno’s paradoxes questioned fundamental assumptions about time, space, and motion,<sup>274</sup> and Russell’s paradox likewise challenged fundamental assumptions about the foundations of mathematics.<sup>275</sup> Similar examples continue to this day, where, for example, Newcomb’s paradox raises questions for decision theory and for issues such as “backward causation,”<sup>276</sup> and various skeptical paradoxes raise questions about “epistemic closure” and other issues related to knowledge and evidence.<sup>277</sup> And the various paradoxes raised by the contested concept of “free will” continue to unearth issues with potentially deep implications for a wide variety of issues in science, law, and other domains.<sup>278</sup> In each of these contexts, the value of the paradoxes is their ability to unearth issues and assumptions with wide implications. Resolving the

---

<sup>273</sup> SAINSBURY, *supra* note 1, at 1.

<sup>274</sup> *See id.* at 4-20 (examining Zeno’s paradoxes); SORENSEN, *supra* note 6, at 48 (“Many of the mathematicians and physicists who present Zeno’s paradoxes assure their readers that Zeno is not crazy. They say he is just challenging us to clarify our ideas.”).

<sup>275</sup> *See* SAINSBURY, *supra* note 1, at 316-32 (examining Russell’s paradox); Andrew David Irvine & Harry Deutsch, *Russell’s Paradox*, STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Oct. 9, 2016), <https://plato.stanford.edu/entries/russell-paradox/> [<https://perma.cc/9Z33-Y9UQ>].

<sup>276</sup> *See* SAINSBURY, *supra* note 1, at 69-82 (examining Newcomb’s paradox); Robert Nozick, *Newcomb’s Problem and Two Principles of Choice*, in *ESSAYS IN HONOR OF CARL G. HEMPEL* (Nicholas Rescher et al. eds., 1969) (using contemporary decision theory to examine Newcomb’s paradox); Alex Bellos, *Newcomb’s Problem Divides Philosophers: Which Side Are You On?*, *GUARDIAN* (Nov. 28, 2016), <https://www.theguardian.com/science/alex-adventures-in-numberland/2016/nov/28/newcombs-problem-divides-philosophers-which-side-are-you-on> [<https://perma.cc/NDU3-NUZJ>] (presenting both possible solutions to Newcomb’s Paradox).

<sup>277</sup> *See* PRITCHARD, *supra* note 26, at 166 (stating that “closure-based radical skepticism” requires “idea that one can have rationally grounded knowledge of the denials of radical skeptical hypotheses, and such a claim appears epistemically immodest in the extreme”).

<sup>278</sup> *See, e.g.*, Adam J. Kolber, *Free Will as a Matter of Law*, in *PHILOSOPHICAL FOUNDATIONS OF LAW AND NEUROSCIENCE* 9, 27 (Dennis Patterson & Michael S. Pardo eds., 2016) (concluding that “[g]iven the apparent centrality of free will to criminal law, it is perhaps surprising that the law fails to address the topic explicitly”); Joshua Greene & Jonathan Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, 359 *PHIL. TRANS. ROYAL SOC’Y LONDON B*. 1775, 1775 (2004) (“New neuroscience will change the law, not by undermining its current assumptions, but by transforming people’s moral intuitions about free will and responsibility.”); Stephen J. Morse, *The Non-Problem of Free Will in Forensic Psychiatry and Psychology*, 25 *BEHAV. SCI. & L.* 203, 204 (2007) (“The only practical free will problem is the confusion among forensic practitioners and others who think that free will is a problem or who speak and write as if it is.”).

issues (and endorsing or rejecting various assumptions in doing so) has profound implications for the respective domains more generally.

The paradoxes of legal proof function similarly. Although each type discusses particular problems of proof that can arise when certain conditions are met, they unearth fundamental issues concerning the legal doctrine, practices, and concepts that structure the process of legal proof. The most important of these issues concerns standards of proof. The paradoxical results arising from each of the three paradoxes all depend on disputed assumptions about the law's proof requirements and the meaning of the various standards. These issues have important implications not merely for the examples themselves but rather for every legal trial and for manifold issues throughout the processes of civil and criminal litigation.<sup>279</sup>

The analysis of the proof paradoxes in this Article provides two primary lessons, one general and one specific. The general lesson concerns the interrelated nature of the paradoxes. Assumptions about how to resolve any of them will have implications for the others. This suggests a caveat for legal commentators—single-shot attempts to “solve” any of the paradoxes may fail to engage with the key proof issues or may otherwise rely on questionable assumptions with undesirable implications.<sup>280</sup> For example, analyses regarding statistical evidence or juror disagreements on the elements that rely on questionable assumptions about burdens and standards of proof (exposed by the conjunction issue) may be of limited utility precisely because they have mismodeled the legal phenomena at issue.

The specific lesson concerns the content of the standards of proof. For each of the paradoxes, assuming a probabilistic interpretation of the proof standards produces the paradoxical conclusions and a host of undesirable implications. In each case, the explanatory account provides a better explanation of the law and a way out of the paradoxical conclusions. The explanatory account avoids the troubling conclusions that appear to follow in the statistical-evidence examples, ameliorates the troubling consequences of the conjunction effect regarding legal elements, and explains how to align juror disagreements on the elements with the legal system's goals regarding standards of proof. Explaining how the explanatory account avoids the paradoxical consequences that arise from the probabilistic account are just a few of the many reasons why the explanatory account remains a better overall explanation of legal proof than any others that have been advanced.<sup>281</sup>

Finally, the explanatory account points the way toward a “happy-face” solution by identifying the premise that is the “odd man out” and should be

---

<sup>279</sup> See *supra* Section I.A.

<sup>280</sup> See Fletcher, *supra* note 7, at 1292 (“If we wish to avoid disabling contradictions, we must reach a deeper understanding of the legal premises that guide our thinking.”).

<sup>281</sup> For other reasons, see Allen & Pardo, *supra* note 57.

rejected.<sup>282</sup> For each paradox this is the assumption that the standard of proof is a probabilistic threshold—for example, that the preponderance standard means proof beyond 0.5 and the beyond-a-reasonable-doubt standard means something around 0.9 or higher.<sup>283</sup> A happy-faced solution not only identifies the premise that should be rejected; it also “shows us why this spurious proposition deceived us, strips from it its patina of truth, so that we’re not taken in by it again.”<sup>284</sup> This Article aims to be a step in that direction. It is not difficult to understand how the “spurious proposition” that the standards are probabilistic thresholds became so entrenched in legal thought.<sup>285</sup> Probability theory has been enormously successful in explaining and formalizing aspects of uncertainty in other domains. Moreover, as Professor Cheng notes, “every first-year law student” learns that the preponderance standard “requires that a plaintiff establish the probability of her claim to greater than 0.5” and that “the criminal beyond-a-reasonable-doubt standard is akin to a probability greater than 0.9 or 0.95.”<sup>286</sup> Despite the intuitive appeal of these notions, he too argues that they are inaccurate and misleading.<sup>287</sup> The explanatory account, and its resolution of the proof paradoxes, exposes why the proposition that standards of proof are probabilistic thresholds is both descriptively false and normatively unattractive. Moreover, in better explaining the relevant legal phenomena, the explanatory account provides the conceptual resources for resisting the charms of this deceptive yet spurious proposition.<sup>288</sup>

---

<sup>282</sup> Schiffer, *supra* note 6, at 178-79.

<sup>283</sup> See *supra* note 60 and accompanying text.

<sup>284</sup> Schiffer, *supra* note 6, at 179.

<sup>285</sup> Indeed, law played an important role in the early development of probability theory. See IAN HACKING, *THE EMERGENCE OF PROBABILITY: A PHILOSOPHICAL STUDY OF EARLY IDEAS ABOUT PROBABILITY, INDUCTION, AND STATISTICAL INFERENCE* 85-91 (2d ed. 2006) (discussing law as focal point for early ideas about probability).

<sup>286</sup> Cheng, *supra* note 3, at 1256. Despite this popular way of talking about the standards, it is important to note that the thresholds are not themselves incorporated into doctrine and jury instructions. Jury instructions on the preponderance standard are quite varied, some consistent with the 0.5 conception and some inconsistent with it. See John Leubsdorf, *The Surprising History of the Preponderance Standard*, 67 FLA. L. REV. 1569, 1571-77 (2015) (surveying jury instructions and noting variations, including “greater weight of the evidence,” “more likely than not,” “actual belief,” and “balance of probabilities,” which imply different outcomes). Courts generally do not attempt to quantify reasonable doubt. See, e.g., *United States v. Copeland*, 369 F. Supp. 2d 275, 286 (E.D.N.Y. 2005) (“Typically, courts have not quantified burdens of proof.”).

<sup>287</sup> Cheng, *supra* note 3, at 1258 (“[E]vidence scholars need [to] let go of their love for  $p > 0.5$ .”); *id.* at 1279 (“Courts and attorneys [should] stop using the misleading 0.5 rule as a shorthand for the preponderance standard . . .”); see also Sullivan, *supra* note 105, at 44 (“Efforts to defend the Bayesian probability approach against [the] paradoxes have been strained and unpersuasive.”).

<sup>288</sup> For a detailed discussion of why this account provides the best available empirical description of legal proof and its various components, see Allen & Pardo, *supra* note 57.