2018

Common Sense on Standards of Proof

Kevin M. Clermont
Cornell Law School, kmc12@cornell.edu

Common Sense on Standards of Proof

*Kevin M. Clermont**

> *Nier l'existence des sentiments tièdes parce qu'ils sont tièdes, c'est nier le soleil tant qu'il n'est pas à midi. La vérité est tout autant dans les demi-teintes que dans les tons tranchés.*
>
> —Gustave Flaubert to Louise Colet, December 11, 1846**

ABSTRACT

The law speaks clearly on the standards of proof, but listeners often misunderstand its words. This Article tries, with some common sense and a modicum of multivalent logic, to explain how the law expects its standards to be applied, and then to show how the law thereby avoids such complications as the conjunction paradox.

First, in accordance with belief function theory, the factfinder should start at zero belief. Given imperfect evidence, the factfinder will end up retaining a fair amount of uncommitted belief. As evidence comes in, though, the factfinder will form a belief in the truth of the disputed fact but also form a disbelief, or a belief in the fact's falsity. At the close of evidence, the standard of proof requires only comparing belief and disbelief. For example, the civil standard, rather than asking whether a fact more likely than not happened according to traditional probability theory, asks whether the factfinder believes in the fact's happening more than the factfinder believes that the fact did not happen. The burdened party need not push proof above fifty percent by dispelling the phantoms of every possibility, and the opponent need not generate a competing version of truth but can instead

---

* Robert D. Ziff Professor of Law, Cornell University. I thank Zach Clopton and Dale Nance for their comments.
** GUSTAVE FLAUBERT, CORRESPONDANCE: PREMIÈRE SÉRIE (1830–1846) 417 (1926) ("To deny the existence of lukewarm sentiments because they are lukewarm is to deny the sun when it is not at noon. Truth lies as much in its shadings as it does in vivid tones."). This epigraph captures the essence of multivalent logic.

rely on denial to demand that the burdened party generate a belief.

Second, belief and disbelief being nonadditive partial truths, the mathematical result is that one cannot combine beliefs by traditional probability theory, as by using the product rule designed for conjunction of betting odds. Instead, one should use fuzzy logic, including its rule that conjoined likelihood equals the likelihood of the least likely element. Linking the elements in a chain tells a story that is as likely as its weakest link. Consequently, if each element of a claim or defense passes the standard of proof, the conjunction of elements will pass the standard of proof. The conjunction paradox thus vaporizes for factfinding, just as the law has always maintained. The law found the way to decide in accord with our best knowledge of the facts.

I. INTRODUCTION

The conjunction paradox has legs. For almost a century it has received lots of attention, and still does.[1] It is easy to grasp. In a civil case, say, the plaintiff needs to prove two elements, *a* and *b*.[2] Assume she proves element *a* to 0.70 and *b* to 0.60, whatever that means. According to American law, she wins, having proven each element by a preponderance.[3] Paradoxically, she should lose according to American law professors, supposedly because she has proven the conjunction of *a* and *b* to only 0.42.

Despite the niche notoriety of the conjunction paradox, most law professors have never thought about conjoining elements until one introduces them to the paradox. Apparently, the law's element-by-element application of the standard of proof has completely accorded with their common sense up to that very moment. Then suddenly and ironically, the law professors' "common sense" demands the application of traditional probability's product rule and induces their immediate condemnation of the law's result as being illogical. That is, almost every law professor who finally contemplates the problem accepts as a matter of common sense that the law's approach is paradoxical; indeed, a few of the hardiest will then

---

[1] The best statement of the paradox appears in Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357, 1385–88 (1985). An early rendition of the paradox lies in the innovative JEROME MICHAEL & MORTIMER J. ADLER, THE NATURE OF JUDICIAL PROOF 140–43 (1931) (also equating preponderance with the *p > 0.50* standard for the first time, according to John Leubsdorf, *The Surprising History of the Preponderance Standard of Civil Proof*, 67 FLA. L. REV. 1569, 1579 n.45 (2015)). More recent scholars reflexively invoke the paradox. *E.g.*, DALE A. NANCE, THE BURDENS OF PROOF 74–78 (2016) (accepting the paradox, although he seemed to later renounce the conjunction paradox in Dale A. Nance, *Formalism and Potential Surprise: Theorizing About Standards of Proof*, 48 SETON HALL L. REV. 1017, 1036–37 (2018)), *reviewed*, Federico Picinali, 9 JURISPRUDENCE 192 (2018), https://ssrn.com/abs tract=3033818; Ronald J. Allen & Sarah A. Jehl, *Burdens of Persuasion in Civil Cases: Algorithms v. Explanations*, 2003 MICH. ST. L. REV. 893; Edward K. Cheng, *Reconceptualizing the Burden of Proof*, 122 YALE L.J. 1254 (2013); Jason Iuliano, *Jury Voting Paradoxes*, 113 MICH. L. REV. 405 (2014); Jonathan J. Koehler & John B. Meixner, *Decision Making and the Law: Truth Barriers*, *in* 2 THE WILEY-BLACKWELL HANDBOOK OF JUDGMENT AND DECISION MAKING 749, 752 (Gideon Keren & George Wu eds., 2015); Michael S. Pardo, *Group Agency and Legal Proof: Or, Why the Jury Is an "It,"* 56 WM. & MARY L. REV. 1793 (2015). The two most recent forays argue at length that the conjunction paradox is real and that therefore the law should change to conform to the product rule. David S. Schwartz & Elliott R. Sober, *The Conjunction Problem and the Logic of Jury Findings*, 59 WM. & MARY L. REV. 619 (2017); Mark Spottswood, *Unraveling the Conjunction Paradox*, 15 LAW, PROBABILITY & RISK 259 (2016).

[2] "Element" means a finding necessary to a claim or defense under the substantive law. *See* KEVIN M. CLERMONT, STANDARDS OF DECISION IN LAW: PSYCHOLOGICAL AND LOGICAL BASES FOR THE STANDARD OF PROOF, HERE AND ABROAD 191 (2013); Kevin M. Clermont, *Conjunction of Evidence and Multivalent Logic*, *in* LAW AND THE NEW LOGICS 32, 57 (H. Patrick Glenn & Lionel D. Smith eds., 2017).

[3] *See infra* note 49.

contort themselves in efforts to find ways to live with the law's approach.[4]

This paradox has long bothered me, because I had grown up to accept the law professors' wisdom on this point.[5] Regardless of its practical impact, the conjunction paradox is of great theoretical importance. But over the years, my bother has decreased. In a series of writings over thirty years, I have tried to show that there is in fact no paradox.[6] The law is correct, and so are all the philosophers and logicians who support the law's result.[7] The law professors are therefore largely in error. Their error stems from a lack of common sense about standards of proof generally.

In his Symposium article, Professor Risinger in large part has set the professors straight by explaining how to view standards of proof in other than the traditional framework of probability theory.[8] In this comment on

---

[4] *See* Kevin M. Clermont, *Trial by Traditional Probability, Relative Plausibility, or Belief Function?*, 66 CASE W. RES. L. REV. 353, 356 (2015). The hardiest contortion is the relative plausibility theory, which posits (i) that the factfinder constructs the overall story (or stories, in some variants of the theory) that the plaintiff is spinning and another story (or stories) that the defendant is (or could be) spinning and (ii) that the factfinder then compares the two stories (or collections of stories) and gives victory to the plaintiff only if the plaintiff's version is more plausible than the defendant's. *See* Allen & Jehl, *supra* note 1, at 929–43. In addition to some real advantages, the theory has serious problems. Most notably, it does not track well the law, which tells its factfinders to proceed element-by-element and apply the standard of proof to each element, not to create holistic stories and compare them; it further diverges from the law by compelling, or at least imposing a practical obligation on, the non-burdened party to choose and formulate a competing version of the truth, even though the law allows the defendant to stand mute and still prevail. Also significantly, the theory gives no reason in logic for narrowing the focus to certain stories, and no explanation for how to compare stories when, say, the plaintiff's story is strong on all but one element. *See* Clermont, *supra*, at 358–60.

[5] *See* RICHARD H. FIELD, BENJAMIN KAPLAN & KEVIN M. CLERMONT, MATERIALS FOR A BASIC COURSE IN CIVIL PROCEDURE 669–72 (6th ed. 1990) (my introducing discussion of the paradox to our casebook, a discussion that evolved until its elimination in the eleventh edition of the book in 2014).

[6] *See* CLERMONT, *supra* note 2, at 5 n.4 (citing my seven previously published articles on standards of proof); *see also* Clermont, *supra* note 2 (subsequent article on multivalent logic); Clermont, *supra* note 4 (subsequent article on belief functions).

[7] *See* L. JONATHAN COHEN, THE PROBABLE AND THE PROVABLE 89–91, 265–67 (1977) (arguing that the conjunction of two or more propositions has the same inductive probability as the least likely conjunct); BERTRAND RUSSELL, HUMAN KNOWLEDGE: ITS SCOPE AND LIMITS 359–61 (1948) (arguing comparably that his "degrees of credibility" do not follow the product rule of traditional probability); Didier Dubois & Henri Prade, *A Set-Theoretic View of Belief Functions: Logical Operations and Approximations by Fuzzy Sets*, *in* CLASSIC WORKS OF THE DEMPSTER-SHAFER THEORY OF BELIEF FUNCTIONS 375, 403 (Ronald R. Yager & Liping Liu eds., 2008) (rejecting the application of "arguments deriving from the study of statistical experiments"); Susan Haack, *The Embedded Epistemologist: Dispatches from the Legal Front*, 25 RATIO JURIS 206, 217–18 (2012) (arguing comparably that her "degrees of warrant" do not follow the product rule of traditional probability); John MacFarlane, *Fuzzy Epistemicism* 12–22 (Oct. 9, 2008), http://johnmacfarlane.net/fuzzy-epistemicism.pdf (arguing against the product rule and in favor of the MIN rule).

[8] D. Michael Risinger, *Leveraging Surprise: What Standards of Proof Imply That We Want from Jurors, and What We Should Say to Them to Get It*, 48 SETON HALL L. REV. 965

his article, I shall sketch an alternative to traditional probabilities that is similar to his. I shall then go a step farther to show how the conjunction paradox reveals more professorial misunderstandings. Although my prior writings provided formal proofs and long explanations on the nonexistence of any paradox as to conjunction, they have unfortunately convinced virtually no one in the legal academy. I keep running into the wall of "common sense." So now I want to take another stab at a short, commonsensical explanation of standards of proof.

The pivotal point of my common-sense explanation entails the realization that factfinders do not announce the odds of an element being true, that is, of its existing in an external and potentially discernible real world.[9] Factfinders instead give their degree of belief in the element's truth based on actual proof-based knowledge. Odds and beliefs are different quantities, subject to different rules for conjunction. Beliefs are not some sort of squishy personal feelings. Instead, factfinders form their beliefs based on the available evidence. Those beliefs express a degree of certainty about the state of the real world.[10] Factfinders might believe a fact to a greater degree than they believe its opposite, even if they would not be willing to bet on it assuming the truth were somehow going to be revealed. The reason is that the factfinders' allocation between belief and disbelief might leave most of their belief uncommitted, while the bettor must commit total belief between yes and no.

II. SINGLE ELEMENT

We first need a word of orientation. I am talking about applying a standard of proof to find a fact. I refer to factfinding in its broad sense, as covering anything that a court or any institution subjects to a proof process in order to establish what the system will treat as truth. The subject includes not only yes-or-no facts but also vague and normative terms like "negligent," as well as many other nonbinary opinions and applications of law to fact. But let me focus mainly on the task of finding a historical fact that constitutes a single element of a claim or defense.

I am not talking here about the factfinder's initial processing of the evidence to form a belief of some strength. We do not know exactly how this processing task is performed, although introspection, logic, and psychology suggest that the factfinder takes into account the weight and credibility of the evidence and then applies some intuitive techniques in nonquantitative and approximate fashion to form the belief. The precise path

---

(2018).

[9] *See* D. Michael Risinger, *Searching for Truth in the American Law of Evidence and Proof*, 47 GA. L. REV. 801 (2013) (treating the necessary philosophical assumptions).

[10] *See* NANCE, *supra* note 1, at 42–57 (discussing "epistemic probability").

does not matter for present purposes. Instead, my interest is in the subsequent evaluation of the processed evidence against the standard of proof, the required measure of sureness in an uncertain world, that the law has specified for the particular element.[11]

I am talking about logic in connection with the standard of proof. We tend to forget that every logic system rests on a series of naked assumptions. The way a logician begins to work is to assume a basic representation of the world, and then to specify a small group of operators, like conjunction, that suffice to generate an internally sound and complete logic system. The logician finally tests the system to see if it makes sense of our world and hence can be useful. All thinkers accept that there are multiple valid logic systems. For any task, we have to choose the most appropriate logical system.[12]

A. *Bivalence*

One thing is clear: the confused law professors are using classical bivalent logic, which posits that all propositions in the world are either true or false.[13] Their usage is proven by their turn to traditional probability for treatment of random uncertainty as to bivalent outcomes. In particular, their usage is proven by their intuitive application of the probabilist's product rule, which governs in a bivalent world. Their use of bivalent logic is perfectly natural because such logic is dominant in our received manner of thinking. Its use is quite helpful in picturing the factfinding process. But it will get one in trouble if used for more than an image, as when one starts using it as a basis for calculations.

If one wants to go farther than simple imaging of the factfinding process, one needs to think more about what factfinding produces. We would love to know if a fact happened or did not happen, but the fact being the subject of trial means we cannot know for sure what really happened. The factfinder will never be able to know a historical fact as being absolutely true or false. Thus, the hypothetical odds of the fact being completely true rather than completely false is a measure of little usefulness to legal actors.

---

[11] *See* CLERMONT, *supra* note 2, at 124–27.

[12] *See* THEODORE SIDER, LOGIC FOR PHILOSOPHY 1–2, 6–11, 72–73 (2010).

[13] The semantic principle of bivalence holds that a proposition *P* is *either true or false*. Thus, *not-P* is not divided into further sets, but constitutes a single set called false. This principle underlies two-valued logic. It is indeed the intended (although not necessary) semantic of all classical logic, so that classical logic tends to induce bivalent thinking. Classical logic's slightly different, syntactic law of the excluded middle holds merely that a proposition *P* is *either true or not true*. Bivalence implies the law of the excluded middle, but the converse is not true. *See* Peter Suber, *Non-Contradiction and Excluded Middle* (1997), http://www.earlham.edu/~peters/courses/logsys/pnc-pem.htm.

What legal actors need is a way to represent their state of knowledge as to the fact. They need the best evidence-based estimate of truth in a world where uncertainties will persist. They need to move on with life while making the best decisions about what happened in the past. In establishing facts in the courtroom we will never know the full truth, because all we can ever find is a degree of truth. We are not concerned with whether something is fully true—whether a truth value of 1.00 will turn up—but whether it seems to have a sufficient truth value to support liability. Thus, the factfinder should form a belief in how true it is. The belief does not assert the fact to be completely either true or false, but rather holds it to be partly true and partly false. This measure of belief in *a* is useful, while the percentage chance of *a* = 1.00 in a bivalent world is not.

Let us say that our knowledge of the world, and rather our justified belief as to *a* and to *b*, is that *a = 0.70* and *b = 0.60*, speaking unrealistically but just for the moment in terms of decimals. These numbers represent our degree of beliefs in *a* and *b*, not of the odds that somehow *a* will turn out to be 1.00 or *b* will turn out to be 1.00. The beliefs depend on what was proven, not the probability of divining the unknowable. The beliefs are all that one can say about the past, and therefore acting on that state of our knowledge is our best course.

There are a variety of routes to systematizing this insight.[14] But the least steep route is multivalent logic, including its most familiar variant called fuzzy logic, which expresses propositions as true to a degree between 0 and 1. Here is a logician's neat anecdote that expresses how multivalent logic fulfills the factfinder's needs:

> Consider Albert Einstein. As the great physicist lay dying, he uttered a sentence just before he passed away. Einstein had dying words. He spoke them in his native tongue, German, the language in which he was by far the most comfortable expressing himself.Unfortunately, he passed away in New Jersey and the only person who heard him was a nurse who spoke no German. So, we will never know what Einstein's last words were.But there is a fact of the matter. There is a sentence which is the last one ever spoken by Einstein. That's a truth of the world; it's just a truth we're unable to know. What was Einstein's final thought on Earth? Any claim that it was some particular sentence is either true or false. But there is a difference between what it is and what we know about it. [Multivalent] logic does not seem apt for describing what is, but it does seem appropriate for talking about our knowledge—

---

[14]   The most systematic alternatives, which lead to the same results as mine, are Cohen's inductive probabilities (*see* COHEN, *supra* note 7), and Shafer's belief functions (*see* GLENN SHAFER, A MATHEMATICAL THEORY OF EVIDENCE (1976)).

and our lack thereof.[15]

B. *Multivalence*

Thus, another thing is perfectly clear: we need not invent a system for thinking about beliefs in multivalent terms, because the accepted non-classical system called multivalent logic already exists. Often it is deployed to handle vague terms, like "tall." But it extends to degrees of truth.

The distinction between bivalence and multivalence appears through the distinction between crisp and fuzzy sets.[16] If a classical probabilist says, "There is a 30% chance that Tom Cruise is tall," the speaker supposes that Tom is either tall or short in a world with just two crisp sets, and on the available evidence he thinks that it is only 30% likely that upon accurate measurement Tom would end up in the tall category rather than the short category. But when a multivalent logician says, "Tom's degree of membership within the set of tall men is 0.30," the speaker means that Tom is not very tall at all. Both statements can be accurate and informative when appropriately deployed. But the difference between them is real and considerable. It derives from the fact that the probabilist is assuming bivalence, so that one is tall or short, while the multivalent logician is speaking of a world where one can be more or less tall.

Within its realm, nobody contests how multivalent logic should operate, at least in general terms. Multivalent logic nicely handles the meaning of vague and even normative terms. The sticking point is law professors' resistance to choosing multivalent logic for measuring and manipulating likelihood of past facts.[17]

Nonetheless, one can express a sort of multivalent likelihood of truth, just as one can multivalently express tallness. Even the purest past facts, ones that look to something like whether an event occurred, are more complicated than they appear. One begins to unpack them by recognizing that past facts most often appear to us not as true or false, but instead as something falling between completely true and completely false, which we can call a partial truth. A partial truth can be expressed in terms of its degree of membership in the fuzzy set of true facts. Such expression is not predicting the probability of absolute truth, but instead is measuring a factfinder's judgment, based on the evidence available, about how likely true a particular proposition is. It represents not a percentage chance, but a partial truth. A 60% belief in an act's occurrence, when we have no way of discovering the absolute truth about the one-shot past event, is epistemically

---

[15] STEVEN GIMBEL, AN INTRODUCTION TO FORMAL LOGIC 530–31 (2016).

[16] *See* SIDER, *supra* note 12, at 12–23 (introducing set theory). Fuzzy set theory is a common way to represent and work with fuzzy logic.

[17] *See, e.g.*, Schwartz & Sober, *supra* note 1, at 637.

indistinguishable from a 0.60 membership in the set of true occurrences, or for that matter a 0.60 membership in the set of tall men.[18]

The law reveals its acceptance of multivalence's partial truths in its instructions to factfinders. It makes no distinctions among kinds of facts, but instead treats historical facts just as it treats vague or normative findings. It treats random uncertainty just as it treats imprecision. This approach—evaluating all facts in the same manner regardless of their nature and regardless of which of the various kinds of uncertainty the law faces—makes sense only if the law has accepted multivalent logic.

The law must eventually act on the partial truths it accepts. To facilitate this step, there exists a compatible mathematical system for handling multivalent beliefs, called belief function theory.[19] Its imagery nicely captures and expresses in an easy and comprehensible way the uncertainty resulting from scarce or conflicting evidence concerning the likelihood of past facts. It also captures and expresses the uncertainty that flows from conceptual imprecision. The degree of membership in a fuzzy set expresses some of all that uncertainty, but fuzzy set theory treats second-order uncertainty, that is, uncertainty about the actual estimate of degree of membership, in the third dimension of a so-called ultra-fuzzy set.[20] Belief functions instead bring second-order uncertainty front and center into a single measure of indeterminacy, just as legal factfinders intuitively do. One could say for belief functions, "same theoretical basis, better image for legal factfinding."

Under belief function theory, a belief in *a* can range anywhere between 0 and 1. Likewise, belief in *not-a*, which is disbelief of *a*, or equivalently active belief in *a*'s contradiction, falls between 0 and 1. Defects in evidence

---

[18] *See* CLERMONT, *supra* note 2, at 163–64 ("The inclusiveness of fuzzy logic suggests its use for each fact and for combining evaluations of different facts. Indeed, it can effortlessly express even a traditional probability as membership in a set, that is, probability is the degree to which the imagined universe of all tests would belong to the set of positive results. This compatibility is essential. It allows easy combination of a randomly uncertain set with an imprecise set, in accordance with fuzzy logical operators." (internal citation omitted)); Clermont, *supra* note 2, at 42–48; Nicholas J.J. Smith, *Degree of Belief Is Expected Truth Value*, *in* CUTS AND CLOUDS: VAGUENESS, ITS NATURE, AND ITS LOGIC 491, 494 (Richard Dietz & Sebastiano Moruzzi eds., 2010) (contending that "uncertainty-based degrees of belief and vagueness-based degrees of belief" must be treated as equivalents).

[19] *See* Clermont, *supra* note 4, at 362–69. The mathematics of belief functions are basically compatible with the set theory underlying multivalent logic, as some have shown by use of so-called possibility theory. *See* CLERMONT, *supra* note 2, at 202–03; Franz Huber, *Belief and Degrees of Belief*, *in* DEGREES OF BELIEF 1, 10–15 (Franz Huber & Christoph Schmidt-Petri eds., 2009); *cf.* Alessandro Saffiotti, *A Belief-Function Logic*, *in* AAAI PROCEEDINGS OF THE TENTH NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE 642, 644 n.3 (1992), https://www.aaai.org/Papers/AAAI/1992/AAAI92-099.pdf (constructing a hybrid logic that attaches a quantified notion of belief to classical first-order logic).

[20] *See* CLERMONT, *supra* note 2, at 162–63.

or its absence will affect the degree of belief in *a* and in *not-a*. Also, on the basis of incomplete, inconclusive, ambiguous, dissonant, or untrustworthy evidence, some of the factfinders' belief should remain indeterminate. Thus, in factfinding, we ask how much we believe *a* to be a real-world truth based on the evidence, as well as how much we believe *not-a*—while remaining conscious of indeterminacy and so recognizing that part of our belief will remain uncommitted. In other words, a belief and a belief in its contradiction will normally add to less than one, making belief function theory a so-called nonadditive system.[21]
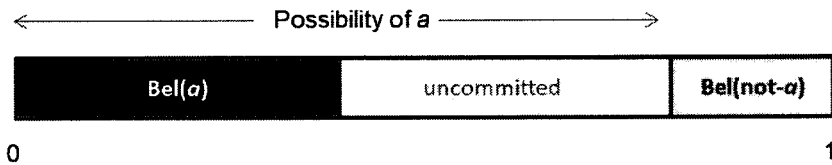
Let *a* be an element of the claim. The zone between *Bel(a)* and *Bel(not-a)* represents the uncommitted belief. Indeed, any case starts with the whole range of belief standing as uncommitted. The proper representation of lack of proof is zero belief in the plaintiff's position—but also zero belief in the defendant's position. As the plaintiff introduces proof, some of the factfinder's uncommitted belief should start to convert into a degree of belief in *a*'s existence, and almost inevitably the plaintiff's proof will also have the inadvertent effect of generating an active belief in at least the slightest possibility of its nonexistence. If the defendant introduces proof, whether in the form of negation or as part of an alternative and inconsistent account, the degree of active belief in *a*'s nonexistence would presumably grow. Indeed, the very clash of beliefs could diminish the degrees of belief in *a* and *not-a*.

When we say at the end of the evidence that *Bel(a) = 0.40*, we are not saying that the belief in *not-a = 0.60*. We are saying only that the proof is such that to a degree of 0.60, *a* has not been proven to be true, which could represent uncommitted belief. Imperfect evidence means that part of our belief will remain uncommitted, with only the remainder divided between *Bel(a)* and *Bel(not-a)*. So, the belief in *a*'s falsity would usually be smaller than 0.60. In the diagram, *Bel(not-a) = 0.20*. It does not equal *1−Bel(a)*, a measure that expresses only the maximal possibility, or plausibility, of *not-a*.

---

[21]   *See* CLERMONT, *supra* note 2, at 151, 187, 203; Rolf Haenni, *Non-Additive Degrees of Belief, in* DEGREES OF BELIEF 121, 121–27 (Franz Huber & Christoph Schmidt-Petri eds., 2009); Ron A. Shapira, *Economic Analysis of the Law of Evidence: A Caveat*, 19 CARDOZO L. REV. 1607, 1613–16 (1998) (distinguishing additive from nonadditive).

        If you wanted to get the odds for betting on *a*, you would somehow
have to allocate the uncommitted belief to *a* and *not-a*. You can bet with
very little information in hand, but you nevertheless must allocate all belief
between the two possible outcomes, in order to place your bet accordingly.
You would do so by performing what is called a pignistic (or betting)
transform from the credal (or belief) stage.

        A step that looks like a pignistic transform is to normalize the
beliefs, by scaling the quantified *Bel(a)* and *Bel(not-a)* up proportionately so
that together they equal 1, and thereby moving the uncommitted belief back
into another dimension. The partial truths are then expressed as normalized
likelihoods. In the diagram, normalizing *Bel(a) = 0.40* and *Bel(not-a) = 0.20*
would yield *a = 0.67* and *a*'s complement *= 0.33*. Although these numbers
look like odds, they are not odds. They are an expression of the beliefs in
fuzzy terms. That is, 0.67 is the best first-order estimate of the provability
of *a*, with the second-order uncertainty relegated to a different dimension.
Thus, fuzzy logic is still a nonadditive system.[22]

        As I shall show in the next section, one need not quantify beliefs in
order to work with them, and indeed usually should not. But if one desired
to quantify a particular proposition, one could express a belief as its degree
of membership in the set of true facts. Given humans' limited ability to
evaluate likelihood, one should express the belief in words drawn from a
coarsely   gradated   scale   of   likelihood,   rather   than   speaking   in
misrepresentative terms of decimals. Here is an appropriately gradated scale
that utilizes natural language and captures the fuzzy imprecision of beliefs:
(1) slightest possibility, (2) reasonable possibility, (3) substantial possibility,
(4) equipoise, (5) probability, (6) high probability, and (7) almost certainty.[23]

        Indeed, the intuitive normalizing process explains how we law
people can so easily speak in percentage terms and translate the standard of
proof into a form such as *p > 50%*. Resort to the nonnumerical scale would
make people sound even more lawyerly. The bottom line is belief function
theory adapts itself easily to imaging proof in just the way that law professors
have traditionally done. Nothing too radical so far!

        The reason for the air of familiarity is that multivalent logic in any
of its forms can be made to look much like bivalence plus probability.
Multivalent logic has this capacity because it is the more general system. It
includes bivalent logic as a special case. That is, a world of things having to
be either black or white is merely a special and extreme case of worlds
actually shaded in grays.

---

[22]   *See supra* note 21.
[23]   *See* CLERMONT, *supra* note 2, at 166–68.

C. *Standard of Proof*

The discussion of belief functions leads naturally to a discussion of standards of proof. Belief functions enable us to restate the accepted view of the standards with a precision that heightens understanding. *In a civil case, rather than misleadingly asking whether a fact is more likely than not true, the preponderance standard is more precisely asking whether the factfinder believes the fact is true more than the factfinder believes that the fact is false.*[24]

This approach does not differ radically from traditional conceptions of proof. Nor is it telling the factfinder to think a new way. It instead tries to represent how a factfinder actually thinks, which is the way that current law seems to call for thinking and is, moreover, a good way to think. So, my more-likely-than-not restatement conveys an important point: compare belief and disbelief, rather than speaking in terms of an abstract possibility that the allegation might not be true. The factfinder should ask the natural question that the law seems to pose: do you believe the burdened party's allegation more than you disbelieve it?

Moreover, this approach to the civil standard of proof does not require the non-burdened party to formulate a competing version of the truth, other than negation through denial. My restatement here conforms to the law too. The law, of course, says that the non-burdened party need not produce any proof or any story. That party can let develop or make develop

---

[24] *See* Clermont, *supra* note 4, at 373–75. Clear-and-convincing would be *Bel(a)* >> *Bel(not-a)*. *See id.* at 375–76. This standard is easy to apply because we are used to the idea of being clearly convinced of something, in law and in life. If one wanted to be more explicit, one could say that the factfinder needs to be more convinced than believing that the likelihood of *a* exceeds the likelihood of *not-a* on the scale of seven gross categories of likelihood described in the text, believing that it is at least two categories higher. Proof beyond a reasonable doubt seems different in kind in its demandingness. *See id.* at 376–77. My conjectured formulation requires both (1) that the likelihood of guilt exceed the plausibility of innocence, so that *Bel(a)* > *1−Bel(a)*, and (2) that there be no reasonable possibility of innocence, so that no reasonable person could hold that *Bel(not-a)* exceeds the lowest category. The first requirement ensures that the proof of guilt is strong enough that the factfinder is not swimming in uncommitted belief, thus requiring a completeness of evidence not required on the civil side. The second requirement expresses the idea implicit in the criminal standard that the factfinder cannot retain a reasonable belief in innocence or, equivalently, a reasonable doubt as to guilt. *See* CLERMONT, *supra* note 2, at 36–38 (explaining the meaning of equivalent standards as the burden shifts). Maybe the two requirements satisfy Professor Risinger's contention that the criminal standard "is intended to make even the juror who thinks that the defendant 'did it,' in everyday terms, think twice." Harold A. Ashford & D. Michael Risinger, *Presumptions, Assumptions, and Due Process in Criminal Cases: A Theoretical Overview*, 79 YALE L.J. 165, 199 (1969). In any event, recognizing this difference in kind of the criminal standard refutes those who would treat it as an ordinary but very high standard and then argue that the standard is much too demanding to be optimal. *E.g.*, Ronald J. Allen & Larry Laudan, *Deadly Dilemmas*, 41 TEX. TECH. L. REV. 65, 68 (2008).

a belief in the falsity of the burdened party's version of the truth in the course of trial.

More specifically, the restated standard calls for constructing separate beliefs for *a* and *not-a*, while leaving some belief uncommitted, and then comparing the beliefs in *a*'s truth and falsity. That is, a preponderance of the evidence means that *Bel(a)* > *Bel(not-a)*, not that *Bel(a)* > *0.50*. Indeed, finding an element to exist will sometimes entail a smallish belief found to exceed an even smaller belief in its contradiction. To continue my example, the factfinder should find *a* if *Bel(a)* = *0.40* when *Bel(not-a)* = *0.20*, with the uncommitted belief equaling *0.40*.

Actually, the factfinder need not quantify the likelihoods. Because all the factfinder needs to do is compare the strengths of belief and disbelief, the factfinder need never place the belief on a scale. But if sometime motivated to specify a fact's likelihood, the factfinder should express it in the coarsely gradated terms of slightest possibility, reasonable possibility, substantial possibility, equipoise, probability, high probability, and almost certainty. Thus, the factfinder in my running example might say that *Bel(a)* > *Bel(not-a)* because the substantial possibility of *a* exceeds the reasonable possibility of *not-a*. That phrasing reveals that the factfinder is not drawing a fine line or making a close call, but saying that belief in the partial truth palpably exceeds belief in its falsity, that is, it is at least a whole step upward in likelihood.

This reformulation in terms of belief functions has considerable advantages over the usual *p* > *50%* formulation.[25] First, the reformulated burden of persuasion does not require the always troublesome task of quantification.[26] Second, the civil standard invokes the factfinder's considerable powers of relative judgment, rather than the weak powers of absolute judgment on some scale of likelihood.[27] Third, it does not require the completeness of proof that would be necessary for the burdened party to get belief above 50%, but instead is willing to follow the law by resting decisions on the evidence presented.[28] Fourth, the burdened party need not

---

[25]   *See* Clermont, *supra* note 4, at 355–56.

[26]   There are the routine objections to speaking of proof in numerical terms, such as that percentages of likelihood are not how people normally think about legal cases and that use of numbers can mislead the factfinder and produce imagined conundrums such as the conjunction paradox. *See* CLERMONT, *supra* note 2, at 75–78, 113–14.

[27]   *See* CLERMONT, *supra* note 2, at 62–66.

[28]   *See* Larry Laudan, *Strange Bedfellows: Inference to the Best Explanation and the Criminal Standard of Proof*, 11 INT'L J. EVIDENCE & PROOF 292, 304–05 (2007) ("The trier of fact cannot say, 'Although plaintiff's case is stronger than defendant's, I will reach no verdict since neither party has a frightfully good story to tell'. Under current rules, if the plaintiff has a better story than the defendant, he must win the suit, even when his theory of the case fails to satisfy the strictures required to qualify his theory as the best explanation."). On how to justify disregarding the uncommitted belief when the proof of both the fact's existence and nonexistence is weak, see Clermont, *supra* note 4, at 374, 381–82.

fight imaginary fights, trying to disprove every alternative possibility, but instead can focus on elevating *a* and depressing *not-a*.   Fifth, this reformulation of the civil standard of proof conforms to the law by saying that the non-burdened party does not need to develop a competing version of the truth, but can rely on negation of any essential fact.   Sixth, the non-burdened party benefits from a tangible burden of production, because the burdened party must produce evidence sufficient to generate a belief of reasonable possibility.   Seventh, the reformulated standard better achieves preponderance's goal of minimizing error costs, while it gives the parties mainly equal treatment.[29]

Conversely, a traditionally probabilistic approach with its absolute numerical scale is undesirable and unworkable.   It also does not conform to the way the law works.   On the one hand, a probabilistic civil standard would seem impossibly difficult, as the plaintiff must rise above 50% while disproving a virtually unlimited number of alternative states of the world.[30] On the other hand, a probabilistic civil standard simultaneously would seem improperly easy, in that the plaintiff, starting with a 50/50 chance, could claim victory with a feather's weight of evidence.[31]

Professor Risinger's Symposium article delivers a convincing rejection of such probabilism.[32]   Speaking only of yes-or-no past facts, which interestingly are the kind of facts that American law professors instinctively feel are most suited to traditional probability theory, he develops an approach similar to mine.   He focuses on the surprise that the factfinder would feel if the found fact hypothetically were revealed to be false, while I stress the factfinder's beliefs in truth and falsity.   He states: "My central claim is that people believe something to be true to the extent that they would be surprised

---

[29]   *See* Clermont, *supra* note 4, at 375.   Nevertheless, Professor Nance has criticized my reformulation.   *See* NANCE, *supra* note 1, at 175–78.   His principal criticism of my preponderance standard was that expressing the test as $Bel(a) > Bel(not-a)$, or equivalently as $Bel(a) / Bel(not-a) > 1$, is not superior to his expressing the test in terms of probability, p, as $p / (1-p) > 1$.   *See id.* at 176 n.188.   But that criticism misses the points of the textual paragraph.   His test entails the difficult task of quantifying p, and it also entails the six other theoretical shortcomings of factfinders' having to speak in terms of traditional probabilistic odds.   But through his more recent study of possibility theory, Professor Nance seems to be coming around to a comparative belief-based standard. *See* Nance, *supra* note 1, at 17 & n.42.

[30]   *See* Michael S. Pardo, *Second-Order Proof Rules*, 61 FLA. L. REV. 1083, 1093 (2009).

[31]   A huge difficulty for traditional probability is fixing the starting point for factfinding. *See* CLERMONT, *supra* note 2, at 204.   The probabilist might assume that when you know nothing, the rational starting point is 50%.   Then, if the plaintiff offers a feather's weight of evidence, he in theory would thereby carry not only his burden of production, but also his burden of persuasion.   The real-life judge, however, hands only defeat to the plaintiff with nothing more than a feather's weight of evidence, and does so by summary means.   Why is that?   The reason is that the well-behaved factfinder starts not at 50% but at 0.   Then, to get past the judge and reach the factfinder requires more than a feather's weight.   It requires a showing not of slightest possibility, but one of reasonable possibility.   That insight makes sense of the notions of the burden of production and the burden of persuasion.

[32]   *See* Risinger, *supra* note 8.

to find out it was false."[33] He calls for the factfinder's "mental experiment—not by asking what one would bet on the truth of a belief, but asking directly how surprised one would be to find out that the thing believed was false."[34] He would then formulate the standard of proof as a measurement of that surprise against an "inherently imprecise scale," albeit one perhaps too finely gradated: "such as mildly surprised, surprised, quite surprised, greatly surprised, astonished, shocked, etc."[35] He explains: "One interesting thing about such scales is that they can be easily understood as rank-ordered, and therefore having the transitive property of numbers, but as not having any of the other properties of mathematization such as additivity."[36] In other words, standards of proof should focus on nonnumerical, nonadditive beliefs as to the facts rather than on hypothetical bets.

III. MULTIPLE ELEMENTS

In describing factfinding on a single element, the difference between dealing in probabilities and dealing in beliefs was rather subtle. I have maintained that multivalent beliefs are a sounder representation, because they do not rest on the faulty assumption of bivalence. But as one switches to factfinding on multiple elements, one must choose definitively whether to combine probabilities or beliefs. They calculate very differently.

Even if one accepts beliefs as the sounder representation, one must seriously resolve to keep thinking in multivalent terms when combining multiple factual elements. One has to maintain a great mental effort not to slip back into the binary thinking that we are so used to employing. When one faces uncertainty as to whether Tom was the perpetrator and as to whether a tort was committed, one will find it almost irresistible to start multiplying odds. Here is the vaccine, again: odds and beliefs are fundamentally different measures.

As for odds, which assume bivalence, the complement of the chance of a's being revealed as true is the chance of a's being revealed as false. The chance of a's being revealed as false interacts with the chance of b's being revealed as false, so that the chance of a or b being revealed as false goes up by a multiplicative amount.

To resist reversion to bivalent thinking, focus on what a belief is. In a sea of uncertainty, it is an island of credence. A belief is not a probabilistic prediction of a complete truth or complete falsity being revealed, but instead is a measure of how well proven a proposition is. It in no way suggests that its complement is the belief's falsity. That Tom is a 0.70 member of the set of tall men does not imply to any degree that he is short. That any a has been

---

[33]   *See* Risinger, *supra* note 8, at 981.
[34]   *See* Risinger, *supra* note 8, at 981.
[35]   *See* Risinger, *supra* note 8, at 982.
[36]   *See* Risinger, *supra* note 8, at 982.

proven to a 0.70 belief does not imply that there is any belief in *not-a*. Instead, the complement of the degree proven is the degree not proven, an uncertainty that could very well eventually resolve in a way that strengthens the belief.

So, when one combines beliefs, one does not combine the complements as the product rule does. The degree of *a*'s not being proven does not have any effect on or interaction with the degree of *b*'s not being proven. If one believes *a* and one believes *b*, then one believes *a* and *b* together, although, of course, not more than one believes *a* or *b* separately. In other words, conjoining knowledge is fundamentally different from placing a bet.

Recall that past facts unknowable with certainty are epistemically the same as vague concepts.[37] Imagine ten or more historical facts each believed to 0.90, which link as necessary elements to form a story. Should you accept that story? Yes, if you are a historian, even though the product rule would put the odds at 35% or less. Rejecting that path, and instead accepting the product rule, would lead to such idiocies as Holocaust denial: the denier would wonder how anyone could think that so many events conjoined to produce the Holocaust?[38] Every accepted story from history consists of linked, not multiplied, beliefs. A legal factfinder should follow the historians' lead in accepting multivalent logic's lesson to treat beliefs differently from odds.

To sum up, if you believe Tom was the perpetrator and you believe a tort was committed, then logically you believe that Tom committed the tort. As I shall next explain, when you want to know how well *a* and *b* are proven on the current proof—and so whether you should proceed with belief in the conjunction—you should look to how well proven the weaker element is.

A. *Product Rule*

The product rule says that the odds of independent events occurring together is the product of their individual odds.[39] As I have argued, truth exists as to past facts, but it is unknowable with certainty; and this reality generates no reason for the law to pay attention to the odds of something happening that will never happen—the truth of an unknowable fact revealed. Then, why should the law use the product rule to find out the odds of repeated

---

[37] *See supra* text accompanying note 18.

[38] *See* Michael Salter, *Countering Holocaust Denial in Relation to the Nuremberg Trials*, *in* HOLOCAUST AND GENOCIDE DENIAL: A CONTEXTUAL PERSPECTIVE 21, 22 (Paul Behrens et al. eds., 2017). *See generally* EVIDENCE AND INFERENCE IN HISTORY AND LAW: INTERDISCIPLINARY DIALOGUES 2–62 (William Twining & Iain Hampsher-Monk eds., 2003); WILLIAM TWINING, RETHINKING EVIDENCE: EXPLORATORY ESSAYS 352–53 (2d ed. 2006).

[39] For interdependent events, the probability operation for conjunction is *p(a)* multiplied by *p(b|a)*, so it is still multiplicative.

miraculous revelations of truth? Instead, the law should work with the knowledge we have, which comes in the form of beliefs.

Mathematically, the reason that the product rule is inappropriate for factfinding is that it is premised on a bivalent world with fully committed belief. In that world, a probability gives the additive odds of *a* being somehow revealed to be true rather than false. The product rule assumes that saying *a = 0.70* means *not-a = 0.30*. If *b = 0.60*, then *not-b = 0.40*. When asked the odds of *(a AND b)*, the result is 0.42, while *(not-a OR not-b)* is 0.58.

However, beliefs in facts do not work that way. Beliefs are nonadditive, so that *a* plus *not-a* does not necessarily equal 1 or, in other words, a belief in a fact does not imply a certain belief in its falsity. In technical terms, nonadditivity makes the product rule inappropriate mathematically.[40]

Multivalent logic provides the appropriate mathematical rule, to be described in the next section. It is easiest to comprehend in an example involving vague terms. Think of the sets of tall men and of smart men. Let *A* be the set of tallness among men and *B* be the perhaps independent set of smartness among men. Tom is a 0.30 member of *A* and a 0.60 member of *B*, which means something like "Tom is not so tall" and "Tom is fairly smart." To what degree is Tom a member of the set of (tallness AND smartness)? One can resolve this question as a matter of intersection of sets. The question becomes how much does Tom belong to the set of tall and smart men.

On the one hand, Tom does not get to carry over the higher set-membership number, because his lack of height drags him down. On the other hand, no one would ever think of multiplying vague measures. There is no reason that the product rule should apply, dragging his membership all the way down to 0.18. This probabilistic result is lower than both his tallness level and his smartness level. It seems to yield, "Because Tom is not so tall and Tom is fairly smart, Tom is very likely a very short and/or very dumb man." Mere application of the product rule seems to have made him very short and dumb. The reason that it calculates the conjunction so low is that multiplication of probabilities gives the irrelevant chance that Tom is both completely tall and completely smart. It does not tell us the degree to which he is both tall and smart.

Instead, Tom's degree of membership in the intersecting set should not decrease below the lower of his tallness and smartness levels. Tom seems a 0.30 member of the set of tall and smart men. The conjunction of tall and smart says, "Because Tom is not so tall and Tom is fairly smart, Tom is not such a tall, smart man."

Perhaps this will appear even more clearly if tallness is 0.70 and smartness is 0.70, so that Tom seems obviously a 0.70 member of the set of (tallness AND smartness). And, if his smartness is only 0.60, Tom's

---

[40]   *See* Clermont, *supra* note 2, at 60–64.

membership in the set of (tallness AND smartness) drops toward 0.60, not precipitously to 0.42.

The inappropriateness of the product rule will become still more obvious as one combines more and more elements in the calculation. The product calculation will approach 0.00, even if all of the values are high. What is Tom's membership in the set of tall, smart, rich, famous, and strange men? The product rule implausibly suggests well under 0.05. Instead, the intersection of sets suggests that the membership should go no lower than the minimum membership around 0.30.[41]

My argument still holds when one switches from combining vague characteristics to combining beliefs.[42] A historical fact unknowable with certainty can be expressed only as a partial truth. The partial truth is saying how strongly we believe it, often leaving a lot of belief uncommitted. It expresses how true we will treat a fact to be. The fact is partially true and partially not true. It says nothing about the odds that it will be revealed as true, because any such statement would require committing that uncommitted belief. So, how do you combine partial truths? It would be bizarre that you could just apply the product rule that was developed for odds. It would be especially bizarre to bring over the product rule when a whole field of logic developed to deal with partial truths and when that field rejects the product rule. For conjunction of beliefs *a* and *b* as partial truths, it is not appropriate to think of multiplication of odds.[43]

---

[41] The product rule would lead to other silliness, such as the aggregation paradox. *See* Alon Harel & Ariel Porat, *Aggregating Probabilities Across Cases: Criminal Responsibility for Unspecified Offenses*, 94 MINN. L. REV. 261 (2009) (building on the aggregation paradox); Ariel Porat & Eric A. Posner, *Aggregation and Law*, 122 YALE L.J. 2 (2012) (same). Theorists would have us convict defendants on the basis of a series of almost proven crimes. *But see* Kevin M. Clermont, *Aggregation of Probabilities and Illogic*, 47 GA. L. REV. 165 (2012) (dismantling the aggregation paradox).

[42] *See supra* text accompanying note 18.

[43] The basic difference between beliefs and odds and the applicability of multivalent theory even to pure yes/no factfinding are the two biggest things that Professor Allen misses in his *End-of-History*-entitled conception of the conjunction paradox as "a feature of the world." Ronald J. Allen, *The Declining Utility of Analyzing Burdens of Persuasion*, 48 SETON HALL L. REV. 995, 2 (2018) (page citations and quotations are drawn from the draft of Sept. 28, 2017, on file with author). In it, he references our earlier e-mail exchange that I failed to heed. *Id.* at 1003 n.31. In that exchange, he announced: "So, my advice to you is to scrap the grand ambitions to explain to people like me why we've missed the boat for the last 30 or so years notwithstanding prodigious effort, including effort to understand precisely what you're discussing here—because I think you're just wrong about that . . . ." E-mail from Ronald J. Allen to author (Dec. 22, 2011, 19:02 EST) (on file with author). He is now sticking to his guns, proclaiming and re-proclaiming that I am "simply wrong." Allen, *supra*, at 1003, 1007.

1. He gets off to a rocky start by misconceiving my whole project. I do not maintain that "changing one's views of probability will change the external world." *Id.* at 996. Instead, I am proceeding from the sound premise that for any particular task in treating the external world, one must choose the appropriate probability tools.

2. He then attacks my tools. He takes particular delight in my water-bottle example, which was designed to demonstrate the difference between crisp sets and fuzzy sets. It

B. *MIN Rule*

When combining beliefs in the legal elements of claim or defense, the sound route is to focus on first-order estimates of how well-proven each belief is, while ignoring second-order uncertainties. The mathematical

---

involves two bottles marked as having either a 91% probability of being nonpoisonous or "a .91 membership in the fuzzy set of potable liquids." CLERMONT, *supra* note 2, at 165. But he misunderstands the example, being misled by his conviction that I want to change the external world by methodological commitment. Incredibly, he believes I am saying that if you think of a bottle in fuzzy logic terms, you can magically make "the contents of the bottle change." Allen, *supra*, at 1005. No. I am saying that the two bottle labels are describing two different contents. The parched person crawling in the desert would be smart to attend to that difference.

3. He next opens his "technical" discussion by accusing me of having "conflated (contorted, perhaps) the distinct ideas of probabilistic conjunction in mathematics and set intersection in set theory." *Id.* at 1005. Given that my central point is that these two ideas are fundamentally different, it is no surprise that his technical discussion is muddled, drawing false distinctions between multivalent logic and belief functions, fuzzy logic and fuzzy set theory, and propositional logic's conjunction and set theory's intersection. His excursion by string-cite to the subject of fuzzy numbers, which is but an old extension of multivalent logic for sophisticated calculations, can have no other purpose than to spread fog. His introduction of fuzzy probability theory as proving me wrong is harder to explain, because I was in fact trying to use fuzzy probability theory in my dealing "with mixed probabilistic/non-probabilistic uncertainty." Michael Beer, *Fuzzy Probability Theory*, in ENCYCLOPEDIA OF COMPLEXITY AND SYSTEMS SCIENCE 4047, 4048 (Robert A. Meyers ed., 2009) ("The significance of fuzzy probability theory lies in the treatment of the elements of a population not as crisp quantities but as set-valued quantities or granules in an uncertain fashion, which largely complies with reality in most everyday situations."). Although relying almost exclusively on Beer, Professor Allen ignores Beer's embrace of the MIN rule coming from so-called imprecise probability theory. *See id.* at 4053.

4. He takes comfort in the loneliness of my position among law professors. *See* Allen, *supra*, at 1013. But I am not all alone. Bertrand Russell is a good ally. Yet Russell might not impress someone who treats disagreement by dismissing the late, great Oxford logician Jonathan Cohen's work as "idiosyncratic." *Id.* at 1004 n.35. *But see* CLERMONT, *supra* note 2, at 180 (discussing "Cohen's almost impenetrably brilliant book").
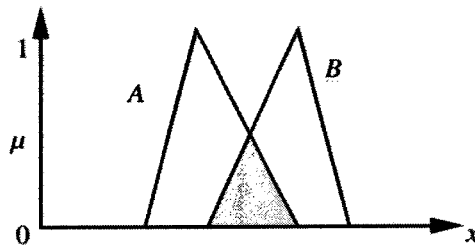
5. After misstating vagueness theory, multivalent logic, and my position in a number of ways, he comes to a glimmer of realization that we are not so far apart. *See* Allen, *supra*, at 1014 ("If Clermont would just drop his fallacious argument about the magical qualities of multivalent logic and set theory, we would welcome him into our camp."). Here is the irony. My approach and his relative plausibility theory both call for comparing the plaintiff's and defendant's positions, without concern for the conjunction paradox. I try to show by logic that the law, with its element-by-element standard of proof, is correct in ignoring the supposed conjunction paradox—and that the law's element-by-element approach is equivalent to comparing the plaintiff's overall account to the defendant's contradictory position. He gets to the same point by fiat, running roughshod over law and logic to decree that henceforth the factfinder should ignore judicial instructions, repeal probability theory, and just compare best stories. *See supra* note 4 (criticizing his relative plausibility theory). He resorts to fiat because he has failed to realize that when applied to standards of proof, every variety of inference to the best explanation, including his theory, implicitly accepts the MIN rule to show that the best explanatory story's conjunction is superior to all competing accounts.

In sum, I am not saying that Professor Allen "missed the boat." In fact, his prodigious rowing has brought us all a long way. It is just that, like Columbus, he does not know where he landed.

sophistication of belief function theory, which results from accounting for second-order uncertainties, is unnecessarily complicated for the law's purposes. It is simpler to return to fuzzy logic and to partial truths expressed as normalized degrees of likelihood.[44] We can then readily prove how to conjoin the partial truths. The conjunction operator for fuzzy logic systems, its alternative to the product rule, is the so-called MIN rule:

$$truth(a\ AND\ b) = minimum(truth(a),\ truth(b))$$

The appropriateness of this operator can be formally proven.[45] But the accompanying figure may help at least to visualize this operator for conjunction.[46] It indicates the shaded intersection of the two sets A and B, where μ gives the degree of membership of a point along the x-axis. Let the x-axis represent some ordering of conceivable states of proof in a particular case. Let z, being some point along the x-axis, represent the proof actually admitted in some case. If z falls within the intersection, the degree of membership therein will be the degree of membership in A or in B, whichever has the lower membership line at that point z. As you can see, at z the membership in *(A AND B)* is *MIN(A, B)*, which in the figure would be less than 0.50.

If after proof in a different case the membership in *A* is 0.70 and the membership in *B* is 0.60, then by fuzzy logic the membership in the intersection would be 0.60. Thinking of intersection of sets might not be that helpful, though. A better image is the links of a chain.[47] The proof tells a story consisting of elements, which can be strung element-by-element as links. The strength of the chain is the strength of its weakest link. In other words, the conjunction of many fairly true propositions is fairly true, not almost completely untrue. If *truth(a) = 0.70* and *truth(b) = 0.60*, then *truth(a AND b) = 0.60*.

Accepting this MIN rule for routine use in law makes everything simpler. The same MIN rule applies whether the elements are independent or interdependent. The MIN rule works the same on facts within elements as between elements, making any particular division of a claim or defense into elements nondeterminative.[48] If the standard of proof were applied to the whole case, the standard of proof would be satisfied if and only if each element meets the standard of proof. If the factfinder in actual practice approaches the case holistically by constructing an overall story, the outcome will equate to the outcome from applying the standard element-by-element. Therefore, the conjunction paradox disappears, whichever of the leading rational or psychological models for processing evidence that one accepts.

This is not to suggest there is no role for the product rule. It remains appropriate for many sorts of situations, including some legal situations. If the event in question is not vague, and thus readily distinguishable from its opposite, and its expected occurrence is subject only to random uncertainty, then traditional probability might be the appropriate vehicle to generate odds.

For a probability example from a non-legal setting: will I twice pick a black ball from an urn containing only pure white and pure black balls? However, if the fact is vague, and most facts in the world are vague, multivalent logic is the way to go. For a contrasting example that calls for multivalence: how black are these grayish balls?

In the legal arena, even after accepting the MIN rule in general, the factfinder will find uses for the product rule, either consciously or intuitively. A judge deciding on a preliminary injunction will need to calculate the odds of multiple future events, a task for which the judge would find the product

---

[47]   *See* Clermont, *supra* note 2, at 64–66.

[48]   Within elements, *see supra* note 2, the factfinder uses the usual intuitive techniques in nonquantitative and approximate fashion to find facts. But the process of combining facts within elements is not dissimilar to the proof process between elements. To join separate facts within the elements (as opposed to reinforcing or contradicting evidence), the factfinder uses the same multivalent operator for conjunction. *See* CLERMONT, *supra* note 2, at 158–59, 191–92; *cf.* Branion v. Gramly, 855 F.2d 1256 (7th Cir. 1988) (rejecting multiplication of odds). The apparent criticality of the number of elements thus melts away. The law splits the case into parts that it calls elements and tells its factfinder to proceed element-by-element, but does so only to make the factfinder's path to decision more comprehensible and careful.

rule to be appropriate. Or a jury might have to make predictions in connection with remedial issues. Another kind of example would arise when the factfinder needs to manipulate and evaluate statistical evidence. Or if one had to determine whether Einstein's actual deathbed statement was a particular q, the multitude of possible things he might have said would probabilistically cause *Bel(not-q)* to balloon while the uncommitted belief remains considerable.

Nonetheless, our lawmakers have explicitly chosen multivalent logic to provide the default operator for combining facts. By phrasing the instruction to factfinders so as to require applying the standard of proof element-by-element, the law reveals that it has adopted the MIN rule.[49] That instruction does not ask the factfinder any overall question. It does not ask what the overall likelihood is. It does not give instructions on how to combine independent or interdependent findings, which would otherwise be necessary. Instead, a rule of law tells the factfinder to determine whether an element is likely enough and, if so, to proceed afresh to the next element. It is a rule of law, and not the factfinder's function, that says the plaintiff shall recover once the factfinder finds each element to exist under the standard of proof.

C. *Multivalent Conjunction*

Multivalent logic, then, calls for the MIN rule to calculate conjunctive likelihood. And the law has heard that call. This legal move has brought the benefits of simplicity. But the question remains whether the law's approach produces the "right" answers.

The right answer depends on the purpose of standards of proof. They could serve a moral end, enhance public acceptance of outcomes, or further some other process value. But I accept the dominant view that the standards aim at the appropriate error distribution. In particular, the civil

---

[49] *See, e.g., In re* Corrugated Container Antitrust Litig., 756 F.2d 411, 416–17 (5th Cir. 1985); 3 KEVIN F. O'MALLEY, JAY E. GRENIG & WILLIAM C. LEE, FEDERAL JURY PRACTICE AND INSTRUCTIONS: CIVIL § 104.01 (6th ed. 2011):

> Plaintiff has the burden in a civil action, such as this, to prove every essential element of plaintiff's claim by a preponderance of the evidence. If plaintiff should fail to establish any essential element of plaintiff's claim by a preponderance of the evidence, you should find for defendant as to that claim.

See also Allen & Jehl, *supra* note 1, at 897–904 (criticizing Dale A. Nance, Commentary, *A Comment on the Supposed Paradoxes of a Mathematical Interpretation of the Logic of Trials*, 66 B.U. L. REV. 947, 949–51 (1986) (finding the pattern jury instruction ambiguous)). Of course, the pattern jury instruction also verbalizes the process implicitly imposed on the judge when acting as factfinder.

standard of preponderance aims at minimizing errors and error costs through the pursuit of accuracy.[50]

Using the MIN rule yields the most accurate story, indeed, the story more likely than all the other possible stories combined. That is, the conjoined story is as likely as the weakest element, which has already satisfied the standard of proof if each element is more likely true than false. Moreover, the disjunction of the falsities that combines all other possible stories is only as likely as the most plausible of the falsities, which has already failed the standard of proof.[51]

This resolution may sound like magic. But it follows ineluctably from the realization that linking the elements in a chain tells a story that is as likely as its weakest link, and so all we have to do is ensure that each link is more true than false. Multivalent logic instructs that this is the way to proceed if you want to make the most accurate decision in accord with our best knowledge of the facts.

Getting the most accurate answer gives the most efficient answer, and hence the "right" answer. If instead the standard of proof asked on whom the factfinder would bet to win if only we were somehow to stumble upon perfect bivalent knowledge, the factfinder should employ the product rule. But there is no reason whatsoever that society should have interest in that question. Basing decisions on the product rule would mean that plaintiffs will lose some strong cases that they should have won. Moreover, defendants at fault would not be receiving enough corrective messages. The product rule would lead to more errors. These errors would detrimentally affect economic efficiency.

Understandably, because society wants to act upon the most accurate view of the facts given the knowledge we have, its legal system has chosen to go with the MIN rule.

CONCLUSION

What the law mandates on standards of decision, if clearly understood and followed, will produce optimal results. First, the standard of proof on an element of a claim or defense should call for a comparison of the factfinder's belief in its truth and belief in its falsity. Second, the standard should be applied separately to each element.

Indeed, one does not have to carry a lot of multivalent theory around in one's head. Once one has stopped worrying about the conjunction paradox, one

---

[50] *See* Clermont, *supra* note 2, at 64–66.

[51] Disjunction is calculated by the so-called MAX rule: *truth(a OR b) = maximum(truth(a), truth(b))*. Thus, if *truth(a) = 0.70* and *truth(b) = 0.60*, the membership in the union would be 0.70. Just as intersection of sets does not cause likelihood to plummet wildly, union does not wildly increase likelihood. The degree of *a*'s not being proven does not have any effect on, or interaction with, the degree of *b*'s not being proven.

no longer needs to master or apply fuzzy logic. Although belief functions, with their uncommitted belief, define the standards of proof, one can understand the resultant standards without flights of theoretical fancy. The civil-case standard asks merely whether the factfinder believes that the fact is true more than the factfinder believes that the fact is false.

　　In any event, nowhere in the analysis of standards does traditional probability theory lend a helpful hand. Instead, the law here is, more or less, a matter of common sense.