Cornell University Law School

Scholarship@Cornell Law: A Digital Repository

Cornell Law Faculty Publications

Faculty Scholarship

Winter 2020

A Theory of Factfinding: The Logic for Processing Evidence

Kevin M. Clermont

A THEORY OF FACTFINDING:
THE LOGIC FOR PROCESSING EVIDENCE

KEVIN M. CLERMONT*

Academics have never agreed on a theory of proof. The darkest corner of anyone's theory has concerned how legal decisionmakers logically should find facts. This Article pries open that cognitive black box. It does so by employing multivalent logic, which enables it to overcome the traditional probability problems that impeded all prior attempts. The result is the first-ever exposure of the proper logic for finding a fact or a case's facts.

The focus will be on the evidential processing phase, rather than the application of the standard of proof as tracked in my prior work. Processing evidence involves (1) reasoning inferentially from a piece of evidence to a degree of belief and of disbelief in the element to be proved, (2) aggregating pieces of evidence that all bear to some degree on one element in order to form a composite degree of belief and of disbelief in the element, and (3) considering the series of elemental beliefs and disbeliefs to reach a decision. Zeroing in, the factfinder in step one should connect each item of evidence to an element to be proved by constructing a chain of inferences, employing multivalent logic's usual rules for conjunction and disjunction to form a belief function that reflects the belief and the disbelief in the element and also the uncommitted belief reflecting uncertainty. The factfinder in step two should aggregate, by weighted arithmetic averaging, the belief functions resulting from all the items of evidence that bear on any one element, creating a composite belief function for the element. The factfinder in step three does not need to combine elements, but instead should directly move to testing whether the degree of belief from each element's composite belief function sufficiently exceeds the corresponding degree of disbelief. In sum, the factfinder should construct a chain of inferences to produce a belief function for each item of evidence bearing on an element,

and then average them to produce for each element a composite belief function ready for the element-by-element standard of proof.

This Article performs the task of mapping normatively how to reason from legal evidence to a decision on facts. More significantly, it constitutes a further demonstration of how embedded the multivalent-belief model is in our law.

> *[I]n our reasonings concerning matter of fact, there are all imaginable degrees of assurance, from the highest certainty to the lowest species of moral evidence. A wise man, therefore, proportions his belief to the evidence.*
> — *David Hume*[1]

---

1. DAVID HUME, AN ENQUIRY CONCERNING HUMAN UNDERSTANDING 80 (Peter Millican ed., Oxford Univ. Press 2007) (1777).

I. INTRODUCTION

Factfinding is foundational for law. The great minds of Locke,[2] Bentham,[3] and Wigmore[4] laid the modern foundation.[5] Of late, factfinding has become the subject of theoretical innovation[6] and even comparative study.[7] Still, here in the United States, we cannot even agree on how to spell it: Should it be one word, hyphenated, or two words?[8]

As the very first step in understanding factfinding at a level deeper than the orthographic, I need to locate the subject. By "fact," I mean it loosely to suggest anything out in the real world that a court, other institution, or person subjects to a proof process in order to establish whether to treat it as truth. The subject extends way beyond the legal context, and certainly far beyond the jurors' task. The subject includes not only yes-or-no facts but also vague and partly normative terms like "fault" and many other applications of law to fact, and

---

2. *See* Richard W. Wright, *Haack on Legal Proof,* 68 ESTUDIOS FILOSÓFICOS 517, 520 (2018) (quoting Locke's treatment of varying degrees of belief).

3. *See* TERENCE ANDERSON, DAVID SCHUM & WILLIAM TWINING, ANALYSIS OF EVIDENCE 80 n.4 (2d ed. 2005) (crediting Bentham as the source of the Rationalist Tradition); *see also* WILLIAM TWINING, RETHINKING EVIDENCE: EXPLORATORY ESSAYS 75–80 (2d ed. 2006) (describing the common-law system's so-called Rationalist Tradition).

4. *See* ANDERSON, SCHUM & TWINING, *supra* note 3, at 87–88 (crediting Wigmore as the inspiration of the New Evidence movement); *see also* Richard Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof,* 66 B.U. L. REV. 439, 440–50 (1986) (describing the common-law system's so-called New Evidence).

5. The modern law of evidence rests on the free evaluation of the evidence: the relevant evidence comes in, subject to some exceptions, and the factfinder rationally processes it without legal restraints. This approach supplanted the medieval formal theory of evidence, or *la preuve légale*: Medieval legal proof had assigned weights to specified classes of evidence, such as admissions and oaths, and prescribed exactly when a set of evidence amounted to full proof. *See* MIRJAN DAMAŠKA, EVALUATION OF EVIDENCE: PREMODERN AND MODERN APPROACHES 2–3, 5 (2019) (arguing that the differences between free evaluation and medieval proof are not as pronounced as conventional wisdom would have it).

6. *See* Lempert, *supra* note 4, at 440–50 (describing the field of New Evidence).

7. *See* R.R. VERKERK, FACT-FINDING IN CIVIL LITIGATION: A COMPARATIVE PERSPECTIVE (2010).

8. Writing "fact finding" as two words is now considered archaic. "Fact-finding" is most common. "Factfinding" is considered the trend, albeit an incipient trend. *See* BRYAN A. GARNER, A DICTIONARY OF MODERN LEGAL USAGE 237 (1987). Yet the GPO STYLE MANUAL 94 (rev. ed. 1973), long treated as authoritative on such matters by the *Bluebook*, says to write "factfinding." *See* THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION R. I.2 (Colum. L. Rev. Ass'n et al. eds., 17th ed. 2000) (referring "punctuation, capitalization, compounding, and other matters of style" to the GPO); *cf. id.* R. 8(c) (20th ed. 2015) (now mentioning capitalization only).

even a variety of nonbinary[9] opinions.[10] Nonetheless, discussion will be easiest if focused on the legal task for judge or jury of reaching a dichotomous finding on a historical fact material to a claim or defense.

Next, I need to deconstruct the legal factfinding process.[11] The "finding" of facts breaks down into three stages: "the fact-gathering stage, the evidence stage and the decision-making stage."[12] The first is the search for and sharing of relevant information.[13] The second is the presentation of evidence to the decisionmaker.[14] In the third or "decision stage, the decision maker will weigh

---

9. Outputs can be yes/no or can be selected among multiple choices, that is, dichotomous or nondichotomous outputs. *See* Adam J. Kolber, *Smooth and Bumpy Laws*, 102 CALIF. L. REV. 655, 666–68 (2014) (calling dichotomous outputs bumpy, and nondichotomous outputs smoother). A nondichotomous output is not only possible, but indeed common in fixing terms of injunctions, criminal sentences, or money damages. For a nondichotomous output, the factfinder could sequentially consider the decisional factors with respect to each outcome and choose the optimizing one. A cost-minimization balancing technique would guide the court in selecting the terms of any injunction. Or the factfinder could formulate its best judgment of the factors' satisfaction by placing its estimate of the situation on some commensurable scale. With the latter sort of method, the law can be especially comfortable with a range of outputs. Each possible output would correspond to a point on or a part of the scale. Sometimes the law would correlate decisional categories, like the degrees of the crime, to the scale's metric. Sometimes the scale's metric would directly produce the output, as is done with money damages. *See, e.g.*, Valerie P. Hans, *What's It Worth? Jury Damage Awards as Community Judgments*, 55 WM. & MARY L. REV. 935, 941–50 (2014) (describing a "gist model" that posits that factfinders first make a categorical gist judgment that money damages are warranted, then make an ordinal gist judgment ranking the damages deserved as low, medium, or high, and finally construct numbers to fit that magnitude). Nondichotomous outputs will therefore involve logical steps supplemental to the logic laid out in this Article.

10. I am not hereby wading into the debate on the fact/value distinction. *See* Kevin Mulligan & Fabrice Correia, *Facts*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2017), https://plato.stanford.edu/entries/facts/ [https://perma.cc/2FKN-BM6L] ("Facts, philosophers like to say, are opposed to theories and to values . . . ."); *cf. id.* § 2.4 (distinguishing fact and proposition, another distinction I do not draw). I am instead using a broad definition of "fact" so as to include all matters subjected to a proof process. *See also infra* note 159.

11. I employ "deconstruct" in its traditional sense, not in the Derridean sense. *See* Bernadette Meyler, *Derrida's Legal Times: Decision, Declaration, Deferral, and Event*, *in* ADMINISTERING INTERPRETATION: DERRIDA, AGAMBEN, AND THE POLITICAL THEOLOGY OF LAW 147, 147 (Peter Goodrich & Michel Rosenfeld eds., 2019). Ironically, my key analytic move thereafter is to deploy "multivalent logic" rather than Derrida's "binary opposition."

12. VERKERK, *supra* note 7, at 1.

13. *See generally* PAUL J. ZWIER & ANTHONY J. BOCCHINO, FACT INVESTIGATION: A PRACTICAL GUIDE TO INTERVIEWING, COUNSELING, AND CASE THEORY DEVELOPMENT (2d ed. 2015).

14. *See generally* MCCORMICK ON EVIDENCE (Kenneth S. Broun gen. ed., 7th ed. 2013). "Evidence rules" determine what the factfinder considers. They include the "principles of proof,"

the evidence and render a decision on matters of fact."[15] The law extensively treats the first two stages by the provisions of procedure and evidence law. But the law treads very lightly in the third stage.

This Article addresses only the third stage. In Figure 1, I divide that mental stage, in turn, into a processing phase and an evaluating phase.
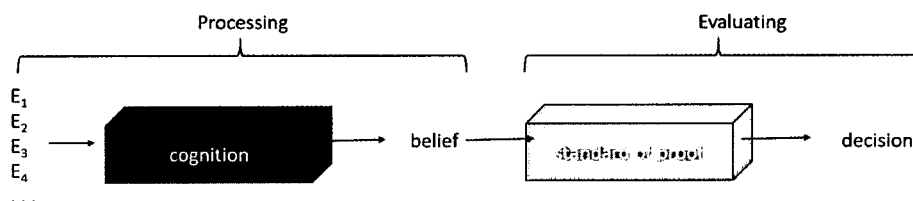

Figure 1: Decision Stage

First, law imposes virtually no enforceable restraints on factfinders' methods during the first phase's processing of pieces of evidence (E_1, E_2, etc.),[16] other than any review of the output for clear error or the like.[17] The factfinders just do it. Psychologists have made limited progress in figuring out how they do it, progress on which I shall draw. The actual process may be rational[18] or

---

which are logically anterior to "exclusionary rules." The principles are materiality, which is determined by substantive law, and probative force, which is determined by logic and general experience but which is loosely cabined by various legal prescriptions. The exclusionary rules, such as hearsay, privilege, and incompetency, are a mixed group of historically based exceptions to the general acceptance of freedom of proof. *See* ANDERSON, SCHUM & TWINING, *supra* note 3, at 289–94 (mapping the evidence rules).

15. VERKERK, *supra* note 7, at 2.

16. *See* ANDERSON, SCHUM & TWINING, *supra* note 3, at 226 (noting that "there is an almost total absence of formal regulation in respect of evaluating evidence or, to put it differently, the Anglo-American law of evidence has almost no rules of weight").

17. *See, e.g.*, FED. R. CIV. P. 52(a)(6) ("clearly erroneous").

18. *See, e.g.*, Edmund M. Morgan, *Introduction to Evidence, in* AUSTIN W. SCOTT & SIDNEY P. SIMPSON, CASES AND OTHER MATERIALS ON CIVIL PROCEDURE 941, 943–45 (1951) (surmising the logical method jurors use to process evidence). "Rational" means logical. *See* ANDERSON, SCHUM & TWINING, *supra* note 3, at 56 (saying that "a conclusion based upon the evidence can only be justified as rational through the use of [deductive, inductive, or abductive] logic").

intuitive,[19] although it should involve so-called critical common sense.[20] It may proceed atomistically[21] or holistically.[22] The factfinders need to combine into a single measure all sorts of evidence on all sorts of facts, ranging from likelihood of a contested occurrence to a vague and partly normative characterization of blameworthiness.[23] The best view based on psychology, and introspection, posits that factfinders process the weight and credibility of the evidence largely by intuition and in an approximate and nonquantified way, perhaps while looking simultaneously at the whole case. They then take a stab at forming beliefs as to the truth. Although the stabs seem to be generally reliable,[24] this kind of cognition remains a black box, both in the study of law as a logical matter and in the study of practice from a psychological perspective.

---

19. *See, e.g.*, Mark Spottswood, *The Hidden Structure of Fact-Finding*, 64 CASE W. RES. L. REV. 131, 137 (2013) (applying the dual-process psychological framework to legal factfinding). "Intuition" is analyzed in BENGT LINDELL, MULTI-CRITERIA ANALYSIS IN LEGAL REASONING 80–97 (2017).

20. *See* D. Michael Risinger, *Searching for Truth in the American Law of Evidence and Proof*, 47 GA. L. REV. 801, 813 (2013) (discussing "the notion of critical common sense and its attendant implication that participation in rational factfinding about legal issues is possible for most humans of normal intelligence").

21. American law seems to envisage an atomistic approach. *See* Marijke Malsch & Ian Freckelton, *The Evaluation of Evidence: Differences Between Legal Systems, in* LEGAL EVIDENCE AND PROOF: STATISTICS, STORIES, LOGIC 117, 130–31 (Hendrik Kaptein, Henry Prakken & Bart Verheij eds., 2009) (linking the adversary system to atomism). Judicial instructions impose on the jury the duty to proceed element-by-element. *See* 3 KEVIN F. O'MALLEY, JAY E. GRENIG & WILLIAM C. LEE, FEDERAL JURY PRACTICE AND INSTRUCTIONS: CIVIL § 104:01 (6th ed. 2011) (describing the burden "to prove every essential element"); *infra* text accompanying note 196 (quoting instruction). Theorizing on the fit between law and logic should therefore depart from an atomistic approach.

22. American practice seems to follow a holistic approach. *See* Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511, 559–69 (2004) (arguing that factfinders consider evidence holistically rather than atomistically). Holistic methodologies are typified by the well-known story model, which posits that the factfinder would judge the likelihood of the proponent's whole story rather than judging its parts separately. *See* Kevin M. Clermont, *Staying Faithful to the Standards of Proof*, 104 CORNELL L. REV. 1457, 1466 n.30 (2019) (citing sources).

23. *See* KEVIN M. CLERMONT, STANDARDS OF DECISION IN LAW: PSYCHOLOGICAL AND LOGICAL BASES FOR THE STANDARD OF PROOF, HERE AND ABROAD 157–66 (2013) (discussing the need for a common currency).

24. *See, e.g.*, Kevin M. Clermont & Theodore Eisenberg, *Trial by Jury or Judge: Transcending Empiricism*, 77 CORNELL L. REV. 1124, 1154 (1992) ("Apparently, judge trial and jury trial combine to operate a decisionmaking system that is, at least in [its ability to treat like cases alike], highly reliable."); Thomas B. Metzloff, *Resolving Malpractice Disputes: Imaging the Jury's Shadow*, LAW & CONTEMP. PROBS., Winter 1991, at 43, 43 (showing the trial system's usual competence and fairness by an empirical comparison of medical malpractice verdicts and insurers' pretrial evaluations).

Second, in the evaluation phase, whereby beliefs lead to legal decision, the standards of proof reign.   Psychologists have thus far had very little to contribute to our understanding of the standards of proof. Philosophical logic takes over here, providing a theoretical basis. Given that basis, the law specifies those standards of proof as the measure of sureness required in an uncertain world for decision about each material fact, fixing the requirement's level to achieve policy goals.

I have written extensively on the evaluation or standard-of-proof phase, from both the logical and the legal angles.[25] My conclusion was that heretofore we have understood the evaluation phase much less well than we pretended. I called for expressing factfinding in terms of degrees of beliefs, rather than the probabilistic odds of an absolute truth that might somehow be revealed.

In the huge subject of the theory of proof, I now feel the need to extend my focus from the evaluation phase to the processing phase. Not only is the processing phase the natural place to continue exploration, but also this very important subject is truly unexplored. Just as for many foundational ideas, we tend to invoke the idea of factfinding all the time without pursuing the idea all the way down. For example, we readily state the test for granting judgment as a matter of law against a party in terms of whether "a reasonable jury would not have a legally sufficient evidentiary basis to find for the party on that issue."[26] The judge therefore needs to determine the limit of reasonable factfinding. In this context, and in many others, it is undeniably appropriate for judges and lawyers to think deeply about how a jury ought to find a fact. But we do not, remaining willing to represent the process as a fairly opaque black box.

This Article's impetus stems from a natural unwillingness to leave factfinding as a black box. In particular, the Article will focus on the logical ideal for factfinders to process evidence.[27] The result is an original account of how facts should be found, a systematic logic for factfinding.

---

25. *See* Clermont, *supra* note 22, at 1464 n.23, 1505 n.161 (referencing my prior work, twelve articles and a book, on standards); *infra* note 159 (citing two subsequent articles).

26. FED. R. CIV. P. 50(a)(1); *see infra* text accompanying note 161 (specifying the standard).

27. I shall be focusing on the logic for a single factfinder. But I do not think that having a group of factfinders complicates the logical picture. The idea would be that each factfinder mentally does its own thing to produce a decision. All of them must then deliberate in order to convince a number of them, sufficient to satisfy the governing unanimity or nonunanimity rule, to agree on one outcome. On the utilization of standards of proof by a group of factfinders, see Allison Orr Larsen, *Bargaining Inside the Black Box*, 99 GEO. L.J. 1567, 1608 (2011). On the requisite agreement, whereby the requisite percentage of factfinders must agree on each element but not on the evidence, grounds, or theories underlying each element, see RICHARD H. FIELD, BENJAMIN KAPLAN & KEVIN M. CLERMONT,

Because my interest thus falls mainly on the normative side, I shall leave it to psychologists to conduct further research on how factfinding is actually done and, in particular, how practice departs from the logical. I shall therefore not focus heavily on the undoubted role of intuitions, heuristics, emotions, and biases in factfinding.[28] I do, however, observe that my multivalent-belief model, in lieu of probabilities, may make more comprehensible any future discussion of the role of intuitions, heuristics, emotions, and biases.

Nonetheless, I am not engaged in pure theory. I am trying to formulate an ideal consistent with the greater legal system's purposes and constraints, while still keeping at least one eye open to human limitations. The normative logic for factfinding will turn out to be legally acceptable and rather feasible to perform. I make no call for change to law or practice, as none proves necessary. I describe how I think good factfinders should, could, and largely do proceed.

A. Background Concepts

Clearly, the ideal logic for processing legal evidence is worth exposition. But is the search for the logical ideal even a possible quest? I previously concluded about the ideal: "Logicians have not managed to agree on how evidence should get processed."[29] Now, however, I assert that the logicians' failure was not inevitable. It resulted from inapt premises, which led inevitably to a roadblock. The logicians almost all proceeded on the basis of classical logic and its derivative of traditional probability.[30] With those tools,

MATERIALS FOR A BASIC COURSE IN CIVIL PROCEDURE 1560–61 (13th ed. 2020). In the situation of a nondichotomous decision, such as awarding damages, group deliberation might involve some sort of averaging process, as well as entail its own effects of intuitions, heuristics, emotions, and biases. *See supra* note 9 (discussing nondichotomous decisions).

28. *See, e.g.*, Hannah Phalen, Jessica M. Salerno & Janice Nadler, *Emotional Evidence in Court*, *in* RESEARCH HANDBOOK ON LAW AND EMOTION (Susan A. Bandes et al. eds., forthcoming 2020) (manuscript at 1).

29. Clermont, *supra* note 22, at 1465 (citing Symposium, *Artificial Intelligence and Judicial Proof*, 22 CARDOZO L. REV. 1365 (2001) (working on the need to formalize evidence processing for application of "computational intelligence"), and James Franklin, *How Much of Commonsense and Legal Reasoning Is Formalizable? A Review of Conceptual Obstacles*, 11 LAW PROBABILITY & RISK 225, 245 (2012) (taking a pessimistic view on formalizing common reasoning)).

30. The lone exception as to legal factfinding is L. JONATHAN COHEN, THE PROBABLE AND THE PROVABLE (1977) (arguing that the law's interest lies in provability, not probability). This great book diagnosed the problem but suggested an opaque approach that he called inductive logic. *See* Alan Hájek, *Interpretations of Probability*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY § 3.2 (Edward N. Zalta ed., 2019), https://plato.stanford.edu/entries/probability-interpret/ [https://perma.cc/J3F6-MNCC] (discussing inductive logic more generally). Belief functions are a much more telling

understanding the processing of typically uncertain evidence is indeed an impossible task.

One needs instead to address the processing problem with an alternative to that constrained logic and its probability scheme. I begin that recasting of the processing problem by presenting some required background from my prior work. I regret that this makes for a rather long introduction.

i.   Multivalent Logic

In analyzing the processing phase, traditional probability's fatal defect is that, unbeknownst to most of us, it is built on an assumption of bivalence. That is, like classical logic, it assumes that nothing lies between completely true and completely false.[31] Probability then reflects only the random uncertainty that the proposition is actually true or false, one or the other. In this black-and-white world, probability of truth, p, provides the chance of the fact being somehow revealed as true. Moreover, the probability of the fact being revealed as false is the complement, $1 - p$. This bivalence assumption pays off, in that classical logic and traditional probability turn out to be very useful oversimplifications—but ones that give wrong answers around the edges of their assumptions.

The cure for those mistakes lies in deploying a more general logic system. Many different systems of logic exist, chosen for effectiveness in the particular setting. The way a logician creates a logic system is to assume a set of "genuine logical truths" that provide a basic representation of the world;[32] then stipulate a small but adequate group of "operators," such as conjunction, disjunction, and negation, that suffices to generate an internally sound and complete logic

---

approach. *See* Timber Kerkvliet & Ronald W.J. Meester, *Assessing Forensic Evidence by Computing Belief Functions*, 15 LAW PROBABILITY & RISK 127, 129 (2016) ("We think that Shafer's approach is the most promising when it comes down to application possibilities and general acceptance, since it is conceptually simpler and closer to classical probability than the approach of Cohen."); *see also* Ronald Meester, *Classical Probabilities and Belief Functions in Legal Cases*, 19 LAW PROBABILITY & RISK 1 (2020).

31. *See* Kevin M. Clermont, *Conjunction of Evidence and Multivalent Logic*, *in* LAW AND THE NEW LOGICS 32, 40 (H. Patrick Glenn & Lionel D. Smith eds., 2017) (describing assumptions of classical logic and observing: "This multivalent form of logic boldly declines the simplification offered by two-valued, or bivalent, logic built on a foundation of true/false with an excluded middle. It instead recognizes partial truths. Both a proposition and its opposite can be true to a degree."); Clermont, *supra* note 22, at 1462–63 nn.20–22, 1487–88 n.100 (comparing and contrasting the principle of bivalence with the similar law of the excluded middle).

32. *See* THEODORE SIDER, LOGIC FOR PHILOSOPHY 2, 72–73 (2010) ("[L]ogical truth is *truth by virtue of form*. . . . One can infer a logical truth by using logic alone, without the help of any premises.").

system;[33] and finally test the system to see if it produces "genuine logical consequences" that make pragmatic sense of our world to us.[34] Nonclassical logic may look and sound much like standard logic, but it has altered some classical assumptions (and so requires slightly different operators, as we shall see).[35] Multivalent logic, which is a family of versions of special interest to law, does not assume bivalence.[36] It allows a proposition to be perceived as true to a degree, taking some value between 0 and 1.[37] It accepts that a fact can be both believed and disbelieved. Classical logic appears then as a special case of multivalent logic, usable when values can exist only as 0 or 1.[38]

The easiest examples of the effectiveness of multivalent logic come from vague terms like "tall."[39] In a bivalent world, with crisp terminology, a person would be either tall or not. With multivalence, we can describe a person as exhibiting a degree of tallness. But multivalent logic is in no way limited to

---

33. *See id.* at 9–11, 25, 35–37, 67–80 ("Say that a set of [operators] is *adequate* iff all truth functions can be symbolized using sentences containing no connectives not in that set.").

34. *See id.* at 1–2, 4–9, 72–73 ("The question of whether a given formal system represents genuine logical consequence is a philosophical one . . . .").

35. *See id.* at 72 ("There are many reasons to get interested in nonclassical logic, but one exciting one is the belief that classical logic is *wrong*—that it provides an inadequate model of (genuine) logical truth and logical consequence.").

36. *See generally* GRAHAM PRIEST, AN INTRODUCTION TO NON-CLASSICAL LOGIC: FROM IF TO IS (2d ed. 2008); Siegfried Gottwald, *Many-Valued Logic*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2015), https://plato.stanford.edu/entries/logic-manyvalued/ [https://perma.cc/8C3V-XBYR]; *cf.* MARK HELPRIN, PARIS IN THE PRESENT TENSE 184 (2017) ("[A] paradox was more the statement of two contradictory propositions, both of which, nevertheless, were true. That two contending propositions could be correct was for Jules rather easy to accept in that it was an almost ordinary facet of music, and part of what gave music its escape from worldly friction in its ability to embrace even the starkest contradictions."); MICHEL DE MONTAIGNE, THE COMPLETE WORKS OF MICHAEL DE MONTAIGNE 289 (William Hazlitt ed., London, John Templeton 1842) ("[W]e are, I know not how, double in ourselves, which is the cause that what we believe we do not believe . . . .").

37. Gottwald, *supra* note 36.

38. *Id.*

39. *See* DAVID LANIUS, STRATEGIC INDETERMINACY IN THE LAW 4–61 (2019) (discussing types of vagueness); Roy Sorensen, *Vagueness*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2018), https://plato.stanford.edu/entries/vagueness/ [https://perma.cc/7J58-9HWU] (saying that "a term is vague to the extent that it has borderline cases"); *The Big Bang Theory: The Hofstadter Isotope* (CBS television broadcast Apr. 13, 2009):

> Stuart: Ooh, Sheldon, I'm afraid you couldn't be more wrong.
> Sheldon: More wrong? Wrong is an absolute state and not subject to gradation.
> Stuart: Of course it is. It's a little wrong to say a tomato is a vegetable, it's very wrong to say it's a suspension bridge.

vagueness.[40] It can handle any sort of variable, from vague concepts all the way to sureness about the truth of a past fact. Such past facts otherwise would have to be treated as yes/no—even though they actually appear to the investigator not as true or false, but instead as something falling between completely true and completely false.

What does tallness have in common with whether the light was red? Both are unknowable with precision. The limits on natural language and on human perceptions sometimes impede us in describing, and indeed in thinking about, a person's height.[41] Meanwhile, a past event cannot be retrieved perfectly, because once an event passes into the past, it becomes inaccessible.[42] The two propositions entail different kinds of uncertainty, but they both evade perfect knowledge in the classical sense. The power of multivalent logic lies in its giving us a measure for expressing all the different kinds of uncertainty, that is, giving us a commensurable measure of how far we fall short of perfect knowledge.

Therefore, even for a past event, one can express a multivalent degree of truth, just as one can vaguely express tallness. The factfinder, in practice, adjudges facts as partly true and partly false. Without sleight of hand and with full allegiance to logic, we can reject bivalence in favor of multivalence for the purposes of factfinding. Legal factfinding should aim to measure a partial conviction as established by the evidence, instead of measuring the probability

---

40. *See* Radim Bělohlávek, George J. Klir, Harold W. Lewis III & Eileen Way, *On the Capability of Fuzzy Set Theory to Represent Concepts*, 31 INT'L J. GEN. SYS. 569, 575 (2002) (stressing the reach of fuzzy logic, which is a type of multivalent logic). Legal scholars are increasingly using fuzzy logic. *See* Michael T. Nguyen, Note, *The Myth of "Lucky" Patent Verdicts: Improving the Quality of Appellate Review by Incorporating Fuzzy Logic in Jury Verdicts*, 59 HASTINGS L.J. 1257, 1261 (2008) (listing examples). And of course, scholars in other disciplines are finding all kinds of uses for fuzzy logic, well beyond its traditional stronghold in consumer electronics, control engineering, and artificial intelligence. *See, e.g.*, CHARLES C. RAGIN, USER'S GUIDE TO FUZZY-SET/QUALITATIVE COMPARATIVE ANALYSIS 71 (2008) (describing fuzzy-set/qualitative comparative analysis, a new but accepted social-science method developed to combine case-oriented and variable-oriented analysis, in which the coding is done by fuzzy variables rather than crisp ones, and observing: "With fuzzy sets, logical AND is accomplished by taking the minimum membership score of each case in the sets that are intersected.").

41. *See* Bertrand Russell, *The Philosophy of Logical Atomism*, 28 MONIST 495, 498 (1918) (observing, before extending his point in lecture 2 to the word "Piccadilly": "Everything is vague to a degree you do not realize till you have tried to make it precise, and everything precise is so remote from everything that we normally think, that you cannot for a moment suppose that is what we really mean when we say what we think.").

42. *See, e.g.*, RICHARD A. MULLER, NOW: THE PHYSICS OF TIME (2016) (developing a theory that time, like space, is expanding and that "now" is the leading edge of the new time).

of bivalent truth. As shown in my prior writing, the law should, can, and even does use multivalent logic for factfinding.[43]

ii. Belief Degrees

The critical step in my argument, and thus a step that I have had to defend at length in my prior work, is speaking in terms of beliefs or, rather, multivalent degrees of belief.[44] This degree of belief measures the strength of belief in the truth of the fact. Even when the fact is yes or no, the factfinder's belief in the fact, given an uncertain world, is a matter of degree.

The factfinder can withhold part of its belief, leaving belief uncommitted to an extent dependent on the quality of the evidence and the nature of the issue.[45] The uncommitted belief represents uncertainty. That uncertainty can represent more kinds of uncertainty than the *aleatory* uncertainty of a system that behaves in random ways, which traditional probability expresses. Uncommitted belief can additionally reflect the second-order or *epistemic* uncertainties[46] coming from the ignorance left by incomplete, inconclusive, ambiguous, or dissonant evidence and from the indeterminacy produced by the vagueness of our concepts and expressed perceptions of the real world (or even from the subvarieties of ontological uncertainty, metaphysical indeterminacy,

---

43. Clermont, *supra* note 22, at 1477.

44. *See* Franz Huber, *Belief and Degrees of Belief*, *in* DEGREES OF BELIEF 1, 1 (Franz Huber & Christoph Schmidt-Petri eds., 2009) (exploring generally the new thinking on degrees of belief, and saying: "Degrees of belief formally represent the strength with which we believe the truth of various propositions. . . . For instance, Sophia's degree of belief that it will be sunny in Vienna tomorrow might be .52, whereas her degree of belief that the train will leave on time might be .23. The precise meaning of these statements depends, of course, on the underlying theory of degrees of belief."); *cf.* SUSAN HAACK, *Epistemology and the Law of Evidence: Problems and Projects*, *in* EVIDENCE MATTERS: SCIENCE, PROOF, AND TRUTH IN THE LAW 1, 13 (2014) (describing her epistemology as "gradational," resulting in degrees of "warranted belief"); Eric Schwitzgebel, *Belief*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY § 2.4 (Edward N. Zalta ed., 2015), https://plato.stanford.edu/archives/sum2015/entries/belief/ [https://perma.cc/8WVZ-NBDX] (describing degrees of belief, but failing to account for uncommitted belief).

45. *See* Jane Friedman, *Suspended Judgment*, 162 PHIL. STUD. 165, 165 (2013) (treating uncommitted belief as a genuine attitude, in addition to yes and no).

46. *See* Craig R. Fox & Gülden Ülkümen, *Distinguishing Two Dimensions of Uncertainty*, *in* PERSPECTIVES ON THINKING, JUDGING, AND DECISION MAKING 21, 23 (Wibecke Brun et al. eds., 2011) (distinguishing aleatory from epistemic uncertainty as the first cut). The same distinction is drawn by phraseology such as puzzles and mysteries, tame and wicked problems, and resolvable and radical uncertainty. *See* JOHN KAY & MERVYN KING, RADICAL UNCERTAINTY: DECISION-MAKING BEYOND THE NUMBERS 14, 20–23 (2020) ("resolvable" means either verifiable or reducible to precise probabilities).

and so on).[47] Traditional probability simply ignores these additional kinds of uncertainty.

Speaking in terms of degrees of belief, and uncommitted belief, does not mean that the law is accepting a quasi-belief in lieu of truth.[48] Truth still matters. The law simply recognizes that a multivalent degree of belief in truth is the best it can do, or rather that such a belief is the best representation of what the factfinder actually produces. Although such a belief does not require absolute truth, it is still not a New Age idea or a subjective sensation.[49] It is neither firm knowledge nor a squishy personal feeling. A multivalent degree of belief is instead the factfinder's attempt to express its degree of sureness about the state of the real world as represented by the evidence put before it by a reasonable process.[50]

Let me be clearer about the proper interpretation of belief. A factfinder's belief in a fact is the *degree to which the factfinder considers the fact to have been proven*, measured on a scale running from no-proof-at-all to a fully proven, and hence fully believed, fact. Given imperfect evidence, the factfinder

---

47. On the types of uncertainty, see CLERMONT, *supra* note 23, at 148–49, 153–54, 159–66; Didier Dubois & Henri Prade, A Unified View of Uncertainty Theories 1 (Mar. 7, 2012) (unpublished manuscript), http://legal1.cit.cornell.edu/private/Dubois & Prade.pdf [https://perma.cc/BAD6-UBHQ].

48. For the philosophical foundation of beliefs, see Clermont, *supra* note 22, at 1468–75 (tracking from the correspondence theory for thought and reality, down to a measurement of the degree of sureness about the state of the real world as represented by evidence).

49. *See* GLENN SHAFER, A MATHEMATICAL THEORY OF EVIDENCE 20 (1976) (defining the factfinder's belief as an act of judgment "that represents the degree to which he judges that evidence to support a given proposition and, hence, the degree of belief he wishes to accord the proposition"); Glenn Shafer, *The Construction of Probability Arguments*, 66 B.U. L. REV. 799, 801–04 (1986) (developing a constructive interpretation of probabilistic reasoning that is neither too objective nor too personalistic). *Compare* DAVID CHRISTENSEN, PUTTING LOGIC IN ITS PLACE 12–13, 69 (2004) (saying that some use "belief" as an unqualified or categorical assertion of an all-or-nothing state of belief), *with* L. Jonathan Cohen, *Should a Jury Say What It Believes or What It Accepts?*, 13 CARDOZO L. REV. 465, 479 (1991) (using "belief," for his purposes, in the sense of a "passive feeling," and arguing that factfinders should deal instead in acceptance), *and* Jordi Ferrer Beltrán, *Legal Proof and Fact Finders' Beliefs*, 12 LEGAL THEORY 293, 294 (2006) (saying that "the proof of *p* should be explained in terms of its acceptability (and not simply of its acceptance)").

50. *See* SUSAN HAACK, *Legal Probabilism: An Epistemological Dissent*, *in* EVIDENCE MATTERS, *supra* note 44, at 47, 54 ("[The] standards of proof should be understood, *not* as a simple psychological matter of the degree of jurors' belief, but as primarily an epistemological matter, the degree of belief *warranted by the evidence*."); Leonard R. Jaffee, *Of Probativity and Probability: Statistics, Scientific Evidence, and the Calculus of Chance at Trial*, 46 U. PITT. L. REV. 925, 937 (1985) (saying that the preponderance standard "is intended to assure that the factfinder will not believe an assertion of fact without evidence *adequate in logic and experience* to support the belief").

will retain some degree of belief as uncommitted. That uncommitted belief does not equate to a belief the fact is false, but simply contributes to the degree to which the fact has not been proven.

Perhaps I can further clarify by contrasting belief degrees to traditional probability. By "traditional probability," I am referring to any system conforming to Kolmogorov's axiomatization.[51] All such systems are giving the odds of truth, p, with the necessary implication that the odds of falsity are $1 - p$. These systems include the probability interpretations that most people think of as probability, that is, the classical, frequentist, and subjective versions.[52] Those three interpretations of probability all have a suggestive air of frequentism about them, so that 60% probability means that in a hundred trials, the outcome will tend to be 1 sixty times and the outcome will tend to be 0 forty

---

51. *See* Hájek, *supra* note 30, § 1 (listing the axioms of non-negativity, normalization, and additivity). Of course, some probabilists have perceived the problem of epistemic uncertainties and so have adjusted their approach to create so-called logical probability, which includes or at least borders on inductive probability and epistemic probability. Here falls the brilliant work of Keynes and Carnap, *see id.* § 3.2.1 ("Indeed, the logical interpretation, in its various guises, seeks to encapsulate in full generality the degree of support or confirmation that a piece of evidence *e* confers upon a given hypothesis *h* . . . ."), as well as the theory of imprecise probabilities, *see* Seamus Bradley, *Imprecise Probabilities*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY § 1 (Edward N. Zalta ed., 2019), https://plato.stanford.edu/entries/imprecise-probabilities/ [https://perma.cc/VB93-2P3Q] ("Among the reasons to question the orthodoxy, it seems that the insistence that states of belief be represented by a single real-valued probability function is quite an unrealistic idealization . . . ."). Yet, unless bivalence is jettisoned, these theories require mental gymnastics that go far beyond the capabilities of the law and its factfinders. *See supra* note 30 (alluding to the difficulties of applying Jonathan Cohen's inductive logic); *infra* note 142 (discussing the difficulties of deriving the principle of conjunctive closure). Nonetheless, with great effort, one can arrive at multivalent logic's conclusions through use of logical probability. *See, e.g.*, Brian Weatherson, *From Classical to Intuitionistic Probability*, 44 NOTRE DAME J. FORMAL LOGIC 111, 112 (2003) (arguing that, "where we have little or no evidence for or against *p*, it should be reasonable to have low degrees of belief in each of *p* and ¬*p*"); *cf.* COHEN, *supra* note 30, at 89–91, 220–22, 265–67 (concluding that the conjunction of two or more propositions has the same inductive probability as the least likely conjunct, but more avowedly abandoning probability's axioms and adopting a nonadditive "probability" system to do so). This point is majorly significant. Many readers cannot migrate to multivalent logic, or do not wish to do so. They do not have to do so in order to agree with this Article's conclusions. But they will have to engage probability at a deeper level than traditional probability, which is just too simplistic for the task of describing the processing of evidence.

52. *See* Hájek, *supra* note 30, § 3.1 ("The guiding idea [of classical probability] is that in such circumstances, probability is shared equally among all the possible outcomes, so that the classical probability of an event is simply the fraction of the total number of possibilities in which the event occurs."); *id.* § 3.3.2 ("Your degree of [subjective] belief in *E* is *p* iff *p* units of utility is the price at which you would buy or sell a bet that pays 1 unit of utility if *E*, 0 if not *E*."); *id.* § 3.4 ("[T]he *probability of an attribute A in a finite reference class B is the relative frequency of actual occurrences of A within B.* ").

times. Although subjective accounts of probability resort to the image of willingness to bet, even they carry the same implication that all propositions are true or false, with no room for epistemic uncertainties, so that the odds of truth and falsity add to one.

Unarguably, these beliefs and probabilities both exist. But they are different measures, as can be shown by thinking about how the factfinder would shift from a multivalent view to a bivalent view. Assume the factfinder has formed an uncertain degree of belief. Even with very little information in hand, the factfinder could bet on the truth. To do so, the factfinder would need to formulate the odds of having uncovered the truth that will eventually reveal itself. To get the odds, the factfinder would allocate its belief between the two possible outcomes of true and false by performing some contestable transform from the credal (or belief) stage to the pignistic (or betting) stage.[53] Rather than leaving some belief still uncommitted, the factfinder would commit more belief to p and more belief to $1 - p$.[54] Beliefs and probabilities therefore differ in magnitude, with the former normally being smaller than the latter. Yet taking this extra step of calculating the probability would add nothing of value to the factfinder's belief in the fact. Indeed, that extra step would destroy the information conveyed by those uncertainties that traditional probability ignores.[55]

In other words, finding facts is different from placing bets or flipping coins. The choice for law's focus in factfinding thus comes down to multivalent belief in truth versus bivalent probability of truth. Why do I say that a belief is what

---

53. *See* Barry R. Cobb & Prakash P. Shenoy, *A Comparison of Methods for Transforming Belief Function Models to Probability Models, in* SYMBOLIC AND QUANTITATIVE APPROACHES TO REASONING WITH UNCERTAINTY 255 (Thomas Dyhre Nielsen & Nevin Lianwen Zhang eds., 2003) (surveying several transform methods); Rolf Haenni, *Non-Additive Degrees of Belief, in* DEGREES OF BELIEF, *supra* note 44, at 121, 129 (discussing betting probabilities); Philippe Smets, *Decision Making in the TBM: The Necessity of the Pignistic Transformation*, 38 INT'L J. APPROXIMATE REASONING 133, 136 (2005) (discussing difficulties of the transform); *cf.* Nicholas J.J. Smith, *Degree of Belief Is Expected Truth Value, in* CUTS AND CLOUDS: VAGUENESS, ITS NATURE, AND ITS LOGIC 491, 503–05 (Richard Dietz & Sebastiano Moruzzi eds., 2010) (discussing the transform as applied to vague concepts). For an example of a transform, normalization will scale up the belief and the disbelief proportionately so that together they add to one.

54. *See* Huber, *supra* note 44, at 11 ("Subjective probabilities require the epistemic agent to divide her knowledge or belief base into two mutually exclusive and jointly exhaustive parts: one that speaks in favor of *A* and one that speaks against *A*. That is, the neutral part has to be distributed among the positive and negative parts. Subjective probabilities can thus be seen as [Dempster-Shafer] belief functions without ignorance.").

55. *See* Kerkvliet & Meester, *supra* note 30, at 138 (saying, compared to traditional probability, "that we lose nothing by working with belief functions, and that belief functions only add flexibility").

the legal factfinder ideally, feasibly, and actually produces, rather than a truth or a probability? First, any asserted fact that the law decides to subject to a proof process is not susceptible to being viewed as certainly, or even almost certainly, true or false. The system otherwise would just take the fact as a given. Second, the factual dispute is "unsettlable," or unknowable.[56] There will be no miraculous revelation of truth at the end. Third, while abandoning any sense of epistemic uncertainty, traditional probabilities do not tell the law anything additional that is of interest. Why should the law care very much that the factfinder would view the odds as 60/40, if forced to bet on very weak evidence? Fourth, analyzing in terms of probability presents all sorts of practical problems.[57] A big hurdle, right at the outset, is that probabilities do not capture how humans think of facts; beliefs are a more natural way for human factfinders to think than probabilities; factfinders ask themselves whether they believe a party, not what odds they would demand on a feigned bet.[58] Fifth, the law of proof makes sense in terms of beliefs, while using traditional probabilities would render the existing law illogical. For example, if plaintiffs had to refute the infinite range of possible alternatives to their allegations, they could never prevail at trial.[59] Sixth, beliefs also fit better with the law's purposes and words. The law asks for only the sureness measure of the factfinder's beliefs, a conviction rendered with a sharp sense of what the factfinder does not know.[60]

Therefore, the theory of proof was meant to operate with multivalent degrees of belief. The legal factfinder should, can, and does express its views of triable facts as degrees of belief. It should not deliver a misleading

---

56. COHEN, *supra* note 30, at 91.

57. *See* Clermont, *supra* note 22, at 1459–62, 1484–86 (contrasting traditional probability unfavorably with multivalent logic for purposes of factfinding).

58. *See* CLERMONT, *supra* note 23, at 68–75 (explaining why humans perform probabilistic tasks quite weakly).

59. *See* Michael S. Pardo, *Second-Order Proof Rules*, 61 FLA. L. REV. 1083, 1093 (2009) ("If the plaintiff must prove that some fact, *X*, is more probable than its negation, *not-X*, then the plaintiff should have to show not only the probability that the state of the world is such that *X* is true, but also the probability of every other possible state of the world in which *X* is not true. This would mean that in order to prevail, plaintiffs would have to disprove (or demonstrate the low likelihood of) each of the virtually limitless number of ways the world could have been at the relevant time. This would be a virtually impossible task, and thus, absent conclusive proof, plaintiffs would lose." (footnote omitted)); *cf.* Patrick Rysiew, *Epistemic Contextualism*, in STANFORD ENCYCLOPEDIA OF PHILOSOPHY § 3.3 (Edward N. Zalta ed., 2016), https://plato.stanford.edu/entries/contextualism-epistemology/ [https://perma.cc/4MWK-FYAX] (formulating a response to the skeptical argument).

60. *See* Clermont, *supra* note 22, at 1486 ("The law is looking for some sort of conviction on the part of its factfinders, not a probability.").

probability for betting on having divined the "real" truth. (P.S.: This position on factfinding—think of beliefs, not odds—does not rest on anti-probabilist prejudice. Traditional probabilities have many other roles to play properly in legal proof, as in the presentation of statistical evidence.[61] Rather, this position rests simply on the view that traditional probability's proper roles do not include measuring sureness in finding uncertain facts.)

iii. Belief Functions

Belief function theory is a version of multivalent logic, based on the mathematics of Professor Glenn Shafer, that allows imaging, evaluating, and combining beliefs and also accounts well for uncertainty.[62] The theory's key insights are that, given imperfect evidence, (1) a degree of belief can coexist with a degree of disbelief produced by the evidence, that is, a belief in the contradiction of the fact, and (2) the factfinder can leave some of its belief uncommitted when forming a degree of belief the fact is true and a degree of belief the fact is false.[63] For example, a thinking religious person would not calculate odds of the divine, but instead would strive to generate a belief that overcomes disbelief while retaining a sense of the unknown.

For another example, I contend that a civil factfinder—rather than calculating an all-or-nothing probability—should, can, and does proceed in just this way: the factfinder forms a degree of belief and a degree of disbelief, but retains a sense of uncertainty; the factfinder would say that, given this evidence, it believes one side's position more than or as much as the other side's, although it remains quite unsure; the factfinder would not take the extra step of saying, "I would wager at such-and-such odds if forced to bet in this sea of uncertainty." After the factfinder processes the evidence, the belief in a fact called *a* can range anywhere between 0 and 1. Likewise, belief in *not-a*, which is disbelief of *a* or, equivalently, an active belief in *a*'s contradiction, falls between 0 and 1. Also, given incomplete, inconclusive, ambiguous, or dissonant evidence, the factfinder should retain some belief as uncommitted. In other words, a belief and the belief in its contradiction will normally add to less than one. In sum, we ask how much the factfinder believes *a* to be a real-world truth based on the

---

61. *See infra* note 83 (discussing statistical evidence).

62. The theory of belief functions received formalization and elaboration in 1976 from Shafer, then a professor of statistics at Princeton, who built on earlier work by Harvard's Arthur P. Dempster. *See* SHAFER, *supra* note 49, at 20 (treating "evidence" in a much broader sense than legal evidence); Glenn Shafer, *Perspectives on the Theory and Practice of Belief Functions*, 4 INT'L J. APPROXIMATE REASONING 323, 325 (1990) (summarizing, simplifying, and updating his theory).

63. *See* SHAFER, *supra* note 49, at 5–7.

evidence, as well as how much it believes *not-a*—while it remains conscious of ignorance and indeterminacy, and so recognizes that part of belief should remain uncommitted as a nonbelief.[64]
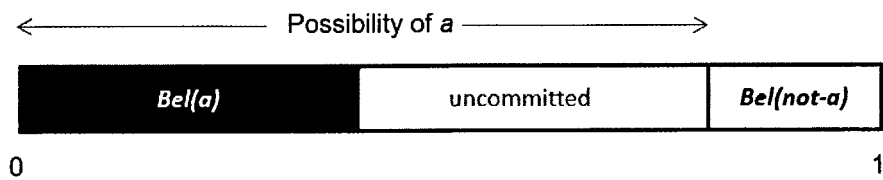

Figure 2: Belief Function

To illustrate by Figure 2, let *a* be a required finding on a factual proposition, say, that Katie is dead.[65] 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

---

64. *See* Liping Liu & Ronald R. Yager, *Classic Works of the Dempster-Shafer Theory of Belief Functions: An Introduction*, *in* CLASSIC WORKS OF THE DEMPSTER-SHAFER THEORY OF BELIEF FUNCTIONS 1, 3–4 (Ronald R. Yager & Liping Liu eds., 2008) (describing uncommitted belief); Meester, *supra* note 30, at 3 ("[A] portion of the total support provided by the evidence may be 'withheld,' or 'uncommitted' as between the mutually exclusive hypotheses *H* and not-*H*."); Hans Rott, *Degrees All the Way Down: Beliefs, Non-Beliefs and Disbeliefs*, *in* DEGREES OF BELIEF, *supra* note 44, at 301, 302 (calling uncommitted belief a nonbelief); Rajendra P. Srivastava & Glenn R. Shafer, *Belief-Function Formulas for Audit Risk*, *in* CLASSIC WORKS OF THE DEMPSTER-SHAFER THEORY OF BELIEF FUNCTIONS, *supra*, at 577, 581 ("Belief functions . . . permit uncommitted belief . . . .").

65. The whole image in Figure 2 is a "belief function." The constituent degrees of belief and disbelief are represented by *Bel(a)* and *Bel(not-a)*. *See* SHAFER, *supra* note 49, at 5–7. The competing probabilists' image, the one that I am rejecting, would be something like this:
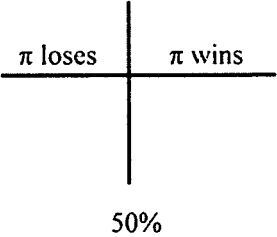
belief in at least the slightest possibility of its nonexistence, namely, that Katie is alive. The zone between *Bel(a)* and *Bel(not-a)* represents the remaining uncommitted belief. If the defendant next introduces effective proof to reduce the belief in *a*, whether the proof comes in the form of negation or as part of an alternative and inconsistent account, the degree of active belief in *a*'s nonexistence would presumably grow. Or the very clash of beliefs could diminish the degrees of belief in both *a* and *not-a*. The force of the parties' presentations, including avoidable or otherwise probative defects in evidence,[66] will affect the degree of belief in *a* and in *not-a*.

When we say after evidence processing that *Bel(a)* = 0.40, we are not saying that *Bel(not-a)* = 0.60. We are saying only that the proof is such that to a degree of 0.60, which could represent uncommitted belief in part or in whole, *a* has not been proven to be true. Imperfect evidence means that some of the belief will remain uncommitted, with the rest of the belief divided between *Bel(a)* and *Bel(not-a)*. So, the belief in *a*'s falsity would be smaller than 1−*Bel(a)* = 0.60. In Figure 2, *Bel(not-a)* = 0.20. Hence, there is a big difference between the complement of *a* and the belief in the contradiction of *a*, the difference being the uncommitted belief.[67] After all, a lack of belief and a disbelief are entirely different states of mind.[68]

---

66. *See* Kevin M. Clermont, *Standards of Proof Revisited*, 33 VT. L. REV. 469, 480–81 (2009) ("[T]he common-law fact-finder is not supposed to hold an *unavoidable* paucity of evidence against the burdened party, but is instead in such a situation supposed to decide the likelihood based on the evidence." (emphasis added)). An avoidable or otherwise probative gap in evidence would best be treated as an item of evidence itself, generating a chain of inferences that supports or undermines the element.

67. Parenthetically, amidst all these decimals, bear in mind that one need not quantify beliefs in order to work with them, and indeed usually one should not. Because all the factfinder usually needs to do is compare the strengths of belief and disbelief, the factfinder need almost never place the fact on a quantified scale of likelihood. Even if one desired to quantify a particular proposition, given humans' limited ability to evaluate likelihood, one should quantify the belief only in words drawn from a coarsely gradated scale of likelihood, rather than speaking in misrepresentative terms of decimals. *See* Clermont, *supra* note 22, at 1480 (providing such a scale). The scale illuminates why, when the factfinder finds *Bel(a)* > *Bel(not-a)*, it is not drawing a fine line or making a close call, but rather saying that the degree of belief is at least a whole step upward in likelihood from the corresponding degree of disbelief.

68. The difference is recognized in FED. R. CIV. P. 11(b)(4), which authorizes denials that, "if specifically so identified, are reasonably based on belief or a lack of information." Such a pleading would be appropriate when the pleader has sufficient information to form a substantial lack of belief in the truth of the opponent's position, but does not feel comfortable asserting that the negative is true. *See* Boykin v. KeyCorp, 521 F.3d 202, 215 (2d Cir. 2008) (approving such pleading when facts are not within the knowledge of the pleader).

Beliefs can alternatively be expressed in set theory.[69] A belief becomes a degree of membership in the set of fully believed facts, namely, the belief function's lower bound, which is termed a "necessity."[70] The unallocated zone between a belief of *Bel(a)* and the contradictory belief of *Bel(not-a)* represents uncommitted belief owing to uncertainty.[71] The belief function's upper bound, which demarks disbelief, represents *a*'s "possibility"; it is the extent to which the factfinder thinks *a* is possible or entertainable; that is, it is the sum of the affirmative belief plus the uncommitted belief or, equivalently, it equals one minus disbelief.[72]

---

69. *See* Huber, *supra* note 44, at 10–15 (discussing the use of possibility theory for this purpose). For an introduction to possibility theory, which builds a bridge between belief functions and fuzzy logic, see Didier Dubois & Henri Prade, *Possibility Theory*, *in* SCHOLARPEDIA (2007), http://www.scholarpedia.org/article/Possibility_theory [https://perma.cc/B5ZL-YASA] (explaining that possibility theory turns on an upper and lower probability called possibility and necessity, respectively). It is important to note that belief function theory is basically consistent with possibility theory, as well as with possibility theory's progenitor fuzzy logic. All these logic systems are consistent, being alternative versions of multivalent logic. *See* Didier Dubois & Henri Prade, *A Set-Theoretic View of Belief Functions: Logical Operations and Approximations by Fuzzy Sets*, *in* CLASSIC WORKS OF THE DEMPSTER-SHAFER THEORY OF BELIEF FUNCTIONS, *supra* note 64, at 375, 403 (linking belief functions and fuzzy sets); Dale A. Nance, *Formalism and Potential Surprise: Theorizing About Standards of Proof*, 48 SETON HALL L. REV. 1017, 1036–37 (2018) (using possibility theory to untangle the conjunction paradox); Ron A. Shapira, *Economic Analysis of the Law of Evidence: A Caveat*, 19 CARDOZO L. REV. 1607, 1614 (1998) ("In the legally relevant literature, it was Professor Glenn Shafer who introduced fuzzy measures as appropriate formalizations of epistemic functions."); L.A. Zadeh, *Fuzzy Sets as a Basis for a Theory of Possibility*, 1 FUZZY SETS & SYS. 3 (1978) (deriving possibility theory from fuzzy sets); Lotfi A. Zadeh, *Reviews of Books: A Mathematical Theory of Evidence*, AI MAG., Fall 1984, at 81, 83 (reviewing GLENN SHAFER, A MATHEMATICAL THEORY OF EVIDENCE (1976) and treating belief function theory as a version of fuzzy logic's possibility theory); *cf.* DAVID A. SCHUM, THE EVIDENTIAL FOUNDATIONS OF PROBABILISTIC REASONING 266–69 (1994) (observing that one can fuzzify belief functions); John Yen, *Generalizing the Dempster-Shafer Theory to Fuzzy Sets*, *in* CLASSIC WORKS OF THE DEMPSTER-SHAFER THEORY OF BELIEF FUNCTIONS, *supra* note 64, at 529 (showing how to form beliefs about membership in fuzzy sets).

70. Dubois & Prade, *Possibility Theory*, *supra* note 69.

71. *See supra* notes 62–63 and accompanying text.

72. *See* Jeffrey A. Barnett, *Computational Methods for A Mathematical Theory of Evidence*, *in* CLASSIC WORKS OF THE DEMPSTER-SHAFER THEORY OF BELIEF FUNCTIONS, *supra* note 64, at 197, 200–01 (providing a neat mental image for these bounds); A.P. Dempster, *Upper and Lower Probabilities Induced by a Multivalued Mapping*, 38 ANNALS MATHEMATICAL STAT. 325, 325 (1967) (providing the mathematical proof for upper and lower bounds). In belief function terminology, "possibility" is often phrased as "plausibility." *See* SCHUM, *supra* note 69, at 236 (using the phrase "plausibility" in place of "possibility").

B. Way Forward

To summarize, the finding of facts firstly breaks down into three stages: the fact-gathering stage, the evidence-presenting stage, and the decisionmaking stage. The present concern is the third, mental stage. That stage, in turn, comprises a processing phase and an evaluating phase. The present focus is the processing phase, rather than the application of the standard of proof.

This Article will now contend that there are three steps in the logic of the factfinder's processing of evidence: (1) reasoning inferentially from a piece of evidence to a degree of belief and of disbelief in the element, where "element" means a fact necessary for a claim or defense to succeed under the substantive law;[73] (2) aggregating pieces of evidence that all bear to some degree on one element in order to form a composite degree of belief and of disbelief in the element; and (3) considering the series of elemental beliefs and disbeliefs to prepare for decision. The next three parts of this Article will treat those three steps.

II. INFERENCE FROM PIECE OF EVIDENCE TO ELEMENT

A. Mental Process

The place for anyone to begin understanding the logical processing of evidence by the factfinder is the marvelous book *Analysis of Evidence*.[74] In it, Professors Anderson, Schum, and Twining break down the mental process in finding facts.[75] They explain that "inferences" are the mental steps in connecting a piece of evidence to the fact to be proved, that is, the element or their so-called probandum.[76] Each inference progresses toward proof of the element by invoking an inductively derived "generalization" that implies the next step deductively.[77] "Ancillary considerations" are the evidence and understandings that refine each generalization.[78]

---

73. If what some other legal source calls an element actually entails separate and necessary facts, each of those facts should be treated as an element for our purposes.

74. ANDERSON, SCHUM & TWINING, *supra* note 3; *see* WILLIAM TWINING, JURIST IN CONTEXT: A MEMOIR 175–89, 214–16, 265–66 (2019) (discussing *Analysis of Evidence*).

75. *See* ANDERSON, SCHUM & TWINING, *supra* note 3, at 78–122 (treating "principles of proof" and "methods of analysis").

76. *See id.*

77. *See id.*

78. *See id.*

The authors then inject these concepts into Wigmore's celebrated charts for evidential analysis.[79] But they run into two problems. First, their expanded charts soon become incredibly complicated—way too complicated to represent the factfinder's actual reasoning.[80] The authors, however, were trying to represent, through argument visualization, the theoretical analysis required of a lawyer in preparing a case, a lawyerly task that benefits from almost any increase in rigor. Second, by adhering to traditional probability theory,[81] the

---

79. *See, e.g.,* JOHN HENRY WIGMORE, THE SCIENCE OF JUDICIAL PROOF, AS GIVEN BY LOGIC, PSYCHOLOGY, AND GENERAL EXPERIENCE, AND ILLUSTRATED IN JUDICIAL TRIALS § 178 (3d ed. 1937).

80. *See, e.g.,* ANDERSON, SCHUM & TWINING, *supra* note 3, at 139 (using a palette of symbols to signify all sorts of different evidential influences, while showing just a tiny fragment of a case's proof involving proof of opportunity (6) on the identity element (4) in a case of murder (1)):
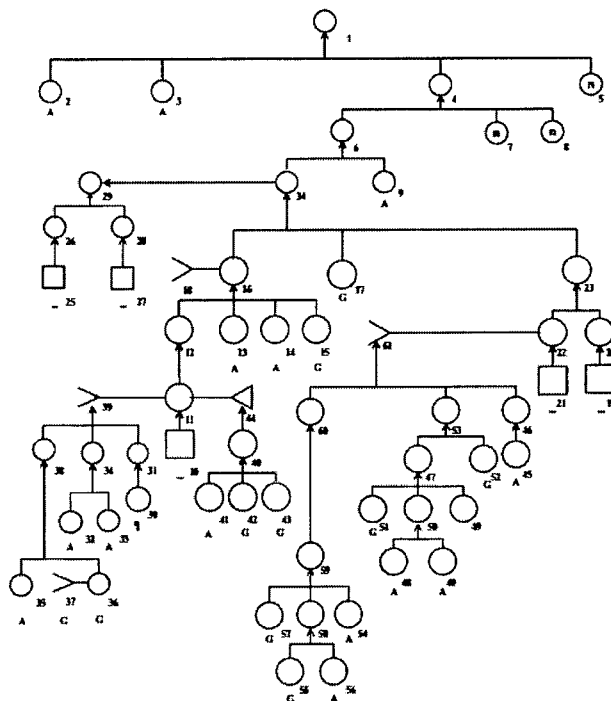
81. *See, e.g., id.* at 107 & n.38 (accepting that compound propositions must be believed as less certain than the weakest component); *id.* at 251, 256–57 (seeming to accept Bayes' theorem for combining evidence); *id.* at 103–04, 381 (leaving the conjunction paradox unsolved). Although by their second edition the authors claimed to be agnostic as to the appropriate logic system, they never abandoned traditional probability. *See id.* at 250–61; *see also id.* at xx–xxi (discussing the book's appendix on probability theory).

authors make their analysis much more complicated and opaque than need be. Shifting from their inapt focus on odds to the insight of multivalent beliefs would have made everything simpler and clearer.[82]

So, I can simplify their Wigmorean charts by representing the processing of evidence only from the factfinder's perspective. I diagram how the factfinder should logically link, in an accurate but feasible way, each piece of evidence to the particular element through a series of inferred multivalent beliefs. I do so in Figure 3. A real-life example will come in Figure 4.
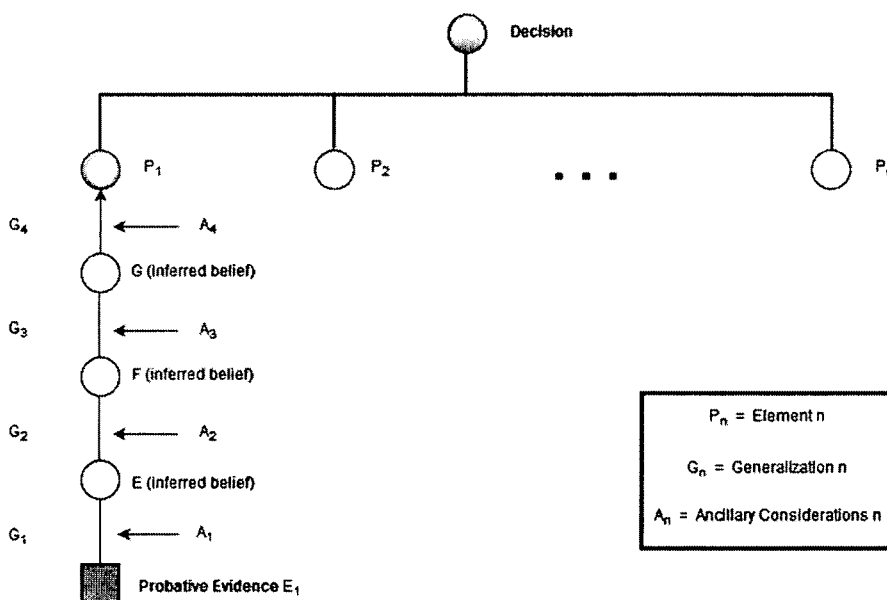
Figure 3: Representation of Inferential Reasoning

Figure 3 starts at the bottom with a piece of "probative evidence," so called because it tends to prove or disprove an element in the claim or defense with

---

82. The authors' allegiance to traditional probability prevents them from ever resolving the logically correct way to combine chained inferences, pieces of evidence, or separate elements—leaving these within the cognitive black box. *See id.* at 106 ("Again putting aside questions concerning the degree of additional strength provided when two propositions converge . . . ."). *Compare id.* at 103 ("These are questions that Wigmore did not address in any detail."), *with id.* at 260 (suggesting that Wigmore foresaw fuzzy logic).

some weight and credibility. It could represent part or all of any introduced evidence, testimonial or real.[83] The piece of probative evidence, E_1, connects up to the element or probandum, P_1, through a chain of inferences of varying force. The same evidence, if it bears on more than one element, will produce multiple chains.

Instead of probative evidence, the book's authors call E_1 "directly relevant evidence,"[84] but I think "relevant" adds nothing because irrelevant evidence is inadmissible, while "direct" gets us unnecessarily into the thicket of direct/circumstantial evidence.[85]   The   accepted   distinction   of direct/circumstantial is this: "Direct evidence is evidence which, if believed, resolves a matter in issue," while for circumstantial evidence, even if "accepted as true, additional reasoning is required to reach the desired conclusion."[86] But others have observed: *"Rule No. 1.* All evidence is either direct or circumstantial. *Rule No. 2.* There is no such thing as direct evidence."[87] The thinking behind *Rule No. 2* is that any argument can be further broken down, so that "in reality all evidence is subject to the frailties of circumstantial evidence."[88] Thus, it is best to use direct and circumstantial only as loose terms suggestive of few or many levels of inference, respectively.

The piece of probative evidence can be positive or negative, building to belief or disbelief, respectively, in the element. Or usually one piece of evidence would add both to the belief and the disbelief. The goal of evidential processing is to use E_1 in producing degrees of conviction, ranging from 0 to 1, in both a belief and a disbelief in P_1. Belief and disbelief will undergo comparison at the eventual standard-of-proof phase.

---

83. The evidence could even be statistical. Any such evidence would have to be linked up to the element by a chain of inferences, thereby converting into a belief. *See* FIELD, KAPLAN & CLERMONT, *supra* note 27, at 1487 (explaining how a factfinder converts statistical evidence: "After all, rationally converting the statistical evidence into a final degree of belief represents a substantial task, at least in all but the most fanciful cases. The evidence will have to be connected up with the issue in the individualized case by a series of permissible but uncertain inferences; also, the evidence may have to be discounted for defects in credibility; the belief may have to be adjusted in light of the probative value of the absence of other proof, an effect most often cutting against the proponent."); Clermont, *supra* note 31, at 46–48 (discussing combination of statistical and nonstatistical evidence).

84. *See* ANDERSON, SCHUM & TWINING, *supra* note 3, at 62–63 (inventing that term).

85. *See id.* at 76–77 (acknowledging that thicket).

86. 1 McCORMICK ON EVIDENCE, *supra* note 14, at 1000–01.

87. DAVID A. BINDER & PAUL BERGMAN, FACT INVESTIGATION: FROM HYPOTHESIS TO PROOF 77 (1984).

88. *Id.* at 79; Giovanni Tuzet, On Probatory Ostension and Inference 10 (Oct. 2019) (unpublished manuscript) (on file with the Bocconi Legal Studies Research Paper Series) ("By itself, evidence proves nothing. Nor does it yield verdicts. It must be 'inferentialized.'").

.

The opponent's *contradictory* evidence will come in as an ancillary consideration on the proponent's chain of inferences. By contrast, if a piece of *conflicting* evidence (that is, one of two items of evidence that could both be true, but that lead to different conclusions, by way of rival explanation) is offered by the opponent, then a new chain of inferences builds from the new evidence to a belief in not-P_1; the opponent must support each inference in the new chain, while the other party tries to defeat some of those inferences to create a disbelief in not-P_1 (that is, a belief in P_1).[89]

In other regards, the end-point concepts of the probative evidence, E_1, and the element, P_1, are apparent. But the middle steps in the chain of inferences require more explanation, which follows.

i.   Inferences

In Figure 3, evidence E_1 is linked to the element by a reasoning chain indicated by inferred propositions E, F, G, and P_1. The factfinder would plot the chain's route by the logical process of abduction,[90] with the lawyers working hard to prod the factfinder into following the route most favorable to their side.[91]

The logical step from proposition to proposition is inference.[92] The chain of reasoning can be long, that is, many inferences may lie between evidence

---

89. On conflicting and contradictory evidence, see ANDERSON, SCHUM & TWINING, *supra* note 3, at 69–70.

90. *See* 5 CHARLES SANDERS PEIRCE, COLLECTED PAPERS OF CHARLES SANDERS PEIRCE: PRAGMATISM AND PRAGMATICISM 106 (Charles Hartshorne & Paul Weiss eds., 3d ed. 1935) ("Abduction is the process of forming an explanatory hypothesis."). Very similar representation of thought has gone into creating software for intelligence analysis. *See, e.g.*, T. Alan Keahey & Stephen G. Eick, A Visualization Architecture for Intelligence Analysis (June 4, 2004) (unpublished manuscript), https://doi.org/10.1117/12.539241 [https://perma.cc/G5QX-CD6U]. The role of abduction therein is discussed by Diane Cluxton, Stephen G. Eick & Jie Yun, Hypothesis Visualization (2004) (unpublished manscript), https://ieeexplore.ieee.org/document/1382915 [https://perma.cc/VL5J-NAKQ]. *Cf.* ANDERSON, SCHUM & TWINING, *supra* note 3, at 56–58 (discussing abduction in fact investigation).

91. On the lawyer's processing of facts, see BINDER & BERGMAN, *supra* note 87, at 78; Charles J. Faruki, *Using Critical Thinking to Analyze Facts*, LITIGATION, Summer 2019, at 52, 52.

92. *See* 1A O'MALLEY, GRENIG & LEE, *supra* note 21, § 12:05 (instructing that "[i]nferences are simply deductions or conclusions which reason and common sense lead the jury to draw from the evidence received in the case").

and element.[93] Although chains can be more complex than drawn, they are all subject to the same analysis.[94]

Each of the inferences may be to some degree false and thus create doubt interposed between the evidence E_1 and the element P_1.[95] The doubt could arise because the inference rests on a wobbly generalization, which has been further undermined by ancillary considerations that question the generalization's quality.[96] Accordingly, the factfinder must proceed by formulating a belief function for each inference, as explained next.

ii. Generalizations

Generalizations supply justification for each reasoning step upward, so that every inference is founded upon a generalization. Generalizations commonly are inductive if-then statements that are beliefs about the world.[97] They often rest on a probabilistic assertion, but they always are subject to some question as to their quality.[98] The factfinder uses the generalization to infer from the prior proposition in the chain of reasoning to a posterior proposition.[99] The generalization acts as a major premise for deductive, or syllogistic-like, reasoning along these lines (the minor premise is the prior proposition, taken as true, and the conclusion is the posterior proposition):[100]

> Most [many, some, etc.] a's are b's
> X is an a
> X is likely [might well be, could be, etc.] a b.

---

93. *See* Morgan, *supra* note 18, at 943–45 (giving an example). On combining these so-called catenate inferences, see ANDERSON, SCHUM & TWINING, *supra* note 3, at 107–08.

94. Alternative generalizations supporting one inference seem complex. *See* SCHUM, *supra* note 69, at 85 ("These additional linkages involving the elements of argument are extremely important in our attempts to capture a wide array of important and interesting subtleties in evidence."). If freed from the misleading implications of traditional probability, however, we can collapse the alternatives into a single chain by the MAX rule. The resultant single chain will be the strongest route from evidence to element. *See infra* text accompanying note 142. Likewise, alternative or disjunctive elements are subject to same collapsing treatment.

95. ANDERSON, SCHUM & TWINING, *supra* note 3, at 61 (making precisely this point).

96. *See* SCHUM, *supra* note 69, at xiii ("In any inference task our evidence is always incomplete, rarely conclusive, and often imprecise or vague; it comes from sources having any gradation of credibility.").

97. *See* ANDERSON, SCHUM & TWINING, supra note 3, at 62, 383.

98. *See supra* note 83 (discussing statistical evidence).

99. See ANDERSON, SCHUM & TWINING, supra note 3, at 78–122.

100. *See* Henry E. Kyburg, Jr., *Uncertain Inferences and Uncertain Conclusions, in* PROCEEDINGS OF THE TWELFTH CONFERENCE ON UNCERTAINTY IN ARTIFICIAL INTELLIGENCE 365, 366–67 (1996) (showing how uncertainty of major premise carries forward to conclusion).

These generalizations might have been entered into evidence, or been judicially noticed, to the limits of the rules of evidence. Most are implicit, however, with their induction often occurring intuitively; the factfinder is permitted to bring to bear common knowledge, although not personal knowledge, in formulating generalizations.[101]

The inference is only as strong as the generalization's strength and accuracy.[102] It is critical to identify expressly the generalization upon which an inference depends in order to determine the force of the inference and thereby identify the weak points in the reasoning. Obviously, generalizations can be dangerous, especially if implicit and unexpressed.[103] They might be speculative in origin or vague in phrasing. They are apt to be value-laden stereotypes drawing on myths and prejudices. In any event, use of generalizations is endemic to inferential reasoning.[104]

In Figure 3, generalizations (labeled G_1 through G_4) are associated with each of the four links in the chain of reasoning from evidence E_1 to element P_1, running through propositions E, F, and G. Thus, generalization G_2 justifies the inference of proposition F from proposition E. This generalization might say: If an event like E occurs, then usually [frequently, often, etc.] a result of F will follow.[105] That is, if E were true, the factfinder would believe F to a degree x that reflects the quality of the generalization. The belief in F would thus

---

101. *See* J. Alexander Tanford, *An Introduction to Trial Law*, 51 MO. L. REV. 623, 700 (1986) (distinguishing common experience from personal experience).

102. *See* ANDERSON, SCHUM & TWINING, *supra* note 3, at 264–65 (distinguishing between strength and accuracy of a generalization). A qualified or amorphous generalization might give only weak support to the inference. Indeed, the more cautious and precise and hence accurate the generalization, the weaker the support for the inference is likely to be.

103. *See id.* at 276 (using "Generalizations are dangerous" as section heading); Elizabeth Thornburg, *(Un)Conscious Judging*, 76 WASH. & LEE L. REV. 1567, 1571 (2019) (arguing that factfinders are "influenced in their thinking by factors such as heuristics, implicit biases, and cultural cognition"). Jurors' generalizations are part of the reason that the Constitution requires jury trials. They are also part of the reason that we fear juries. Beyond imposing a limit on irrational factfinding, judges normally do not check the jurors' generalizations. *But see* Peña-Rodriguez v. Colorado, 137 S. Ct. 855 (2017) (inquiring into racial prejudice).

104. *See* BINDER & BERGMAN, *supra* note 87, at 82–89 (discussing inferential reasoning); ALEX STEIN, FOUNDATIONS OF EVIDENCE LAW 96 (2005) ("The generalization factor is not intrinsically problematic . . . . Weight of generalizations that fact-finders use derives from the empirical instances systematically exhibiting the factual pattern that purports to be a generalization. The word 'systematically' embraces two criteria: that of number and that of variety. A recurrent factual pattern acquires the generalization status when both the number of its individual instances and their variety increase . . . .").

105. ANDERSON, SCHUM & TWINING, *supra* note 3, at 62 (making precisely this point).

measure the certainty of inferring from E to F. It is a conditional belief, which assumes E to be true. In other words, $x = Bel(F|E)$, which may be read as the degree of belief in F if E is fully believed.

iii. Ancillary Considerations

Ancillary evidence and understandings comprise reactions to a generalization. Like a generalization, an ancillary consideration can derive from evidence actually introduced by any party, or it can spring from an interjection of judicial notice or of the factfinder's critical common sense.[106] The ancillary consideration can undermine or support the generalization as a major premise.[107] It will entail its own chain of questionable inferences in linking up to that premise. The various ancillary considerations bearing on the generalization must be combined with the original data underlying the premise, by the weighted-arithmetic-averaging method described in Part III, to produce a degree of belief in the next proposition up the inferential chain.[108]

In Figure 3, ancillary considerations A_1 lie between E_1 and E. Just because evidence E_1 says that event E occurred does not establish that E did occur. A generalization G_1 would be that such evidence is usually trustworthy. But obviously, there is here a matter of "credibility." Credibility means the extent to which we believe what the probative evidence says.[109] Credible *testimonial* evidence should be on personal knowledge and have (1) veracity (testimony in accordance with witness's beliefs), (2) objectivity (testimony not based on expectations or desires), and (3) sensitivity (testimony resting on good sensory evidence); credible *real* evidence should be (1) authentic (in that it is what it purports to be), (2) accurate (and based on a sufficiently sensitive sensing device), and (3) reliable (or repeatable).[110] Ancillary considerations would raise any of these features of credibility to undermine or support the generalization.

Ancillary considerations can affect the higher-level inferences too. These considerations would attack or reinforce the corresponding generalization. In a sense, they thereby address the credibility of the generalization. Thus, although A_1 determines what might be called the primary credibility of E_1, A_2 through A_4 can also bear on credibility. They do so by addressing the accuracy of their corresponding generalization.

---

106. *See id.* at 63, 380.

107. *See id.* at 62–63, 380 (using the term "ancillary evidence" or "indirectly relevant evidence").

108. *See* SCHUM, *supra* note 69, at 83–85 (discussing complex combinations of evidence).

109. ANDERSON, SCHUM & TWINING, *supra* note 3, at 381.

110. *See id.* at 63–70 (discussing credibility generally).

iv. Example: Sacco & Vanzetti

A well-known legal example of inferential reasoning comes from the infamous case of Sacco and Vanzetti.[111] Those Italian immigrants and anarchists were convicted of and executed for the robbery and fatal shooting of two payroll guards in South Braintree, Massachusetts, on April 15, 1920.[112] A few weeks after the shooting, they were arrested on unrelated suspicions.[113] Sacco was then carrying a pistol. Much later they were charged with the murders.[114]

The big issue at the murder trial was identity.[115] One of the arresting officers, Connolly, testified about Sacco's behavior upon arrest.[116] The factfinder had to connect that testimony with the identity element. The path of reasoning proceeded through Sacco's supposed consciousness of guilt, on which the prosecution had to rely heavily in order to prove his identity as one of the murderers.[117] But that path revealed many sources of uncertainty.[118]

---

111. On this case, see *id.* at 21–23, 124–25, 160, 251–52, 285–86, 335–36; HERBERT B. EHRMANN, THE CASE THAT WILL NOT DIE: COMMONWEALTH VS. SACCO AND VANZETTI (1969); OSMOND K. FRAENKEL, THE SACCO-VANZETTI CASE (1931); FELIX FRANKFURTER, THE CASE OF SACCO AND VANZETTI: A CRITICAL ANALYSIS FOR LAWYERS AND LAYMEN (1927); G. LOUIS JOUGHIN & EDMUND M. MORGAN, THE LEGACY OF SACCO AND VANZETTI (1948); JOSEPH B. KADANE & DAVID A. SCHUM, A PROBABILISTIC ANALYSIS OF THE SACCO AND VANZETTI EVIDENCE (1996); ROBERT H. MONTGOMERY, SACCO-VANZETTI: THE MURDER AND THE MYTH (1960); FRANCIS RUSSELL, TRAGEDY IN DEDHAM: THE STORY OF THE SACCO-VANZETTI CASE (1962); THE SACCO-VANZETTI CASE: TRANSCRIPT OF THE RECORD OF THE TRIAL OF NICOLA SACCO AND BARTOLOMEO VANZETTI IN THE COURTS OF MASSACHUSETTS AND SUBSEQUENT PROCEEDINGS 1920–7 (1928); WILLIAM YOUNG & DAVID E. KAISER, POSTMORTEM: NEW EVIDENCE IN THE CASE OF SACCO AND VANZETTI (1985).

112. ANDERSON, SCHUM & TWINING, *supra* note 3, at 21–22.

113. *Id.* at 22.

114. *Id.*

115. RUSSELL, *supra* note 111, at 7.

116. *See* 1 THE SACCO-VANZETTI CASE, *supra* note 111, at 751–53 (reprinting transcript of introduction of this evidence).

117. *See* 2 *id.* at 2257 (charging the jury: "Therefore, the mind, being conscious of every bodily act theretofore committed, it knows whether or not such act is one of innocence or guilt. If it indicates guilt, that is evidence of consciousness of a guilty act, and evidence of a consciousness of a guilty act is evidence tending to prove commission of such guilty act, and evidence of the commission of a criminal act tends to prove the identity of the author of such criminal act.").

118. *See* SCHUM, *supra* note 69, at 75–92 (composing the inferential chain adapted here).

Decision

P₁ = Accused's Identity   P₂ = Victim's Death   P₃ = Unlawful Force   P₄ = Malice Aforethought

G₈   A₈

K (Inference) = That serious crime was the crime charged

G₇   A₇

J (inference) = Sacco was conscious of having been involved in a serious crime

G₆   A₆

I (inference) = Sacco was conscious of having committed a criminal act

G₅   A₅

H (inference) = Sacco intended to escape

G₄   A₄

G (inference) = Sacco intended to use the pistol to threaten or harm the arresting officers

G₃   A₃

F (inference) = Sacco intended to draw the pistol

G₂   A₂

Evidence F* = Officer Connolly testified that his later search revealed a pistol concealed at Sacco's belt

E (inference) = Sacco did try several times to put his hand under his overcoat despite being warned not to do so

G₁   A₁ = Credibility

Probative Evidence E₁ = Officer Connolly testified in effect "Sacco attempted on several occasions to put his hand under his overcoat in spite of being warned by me not to do so."
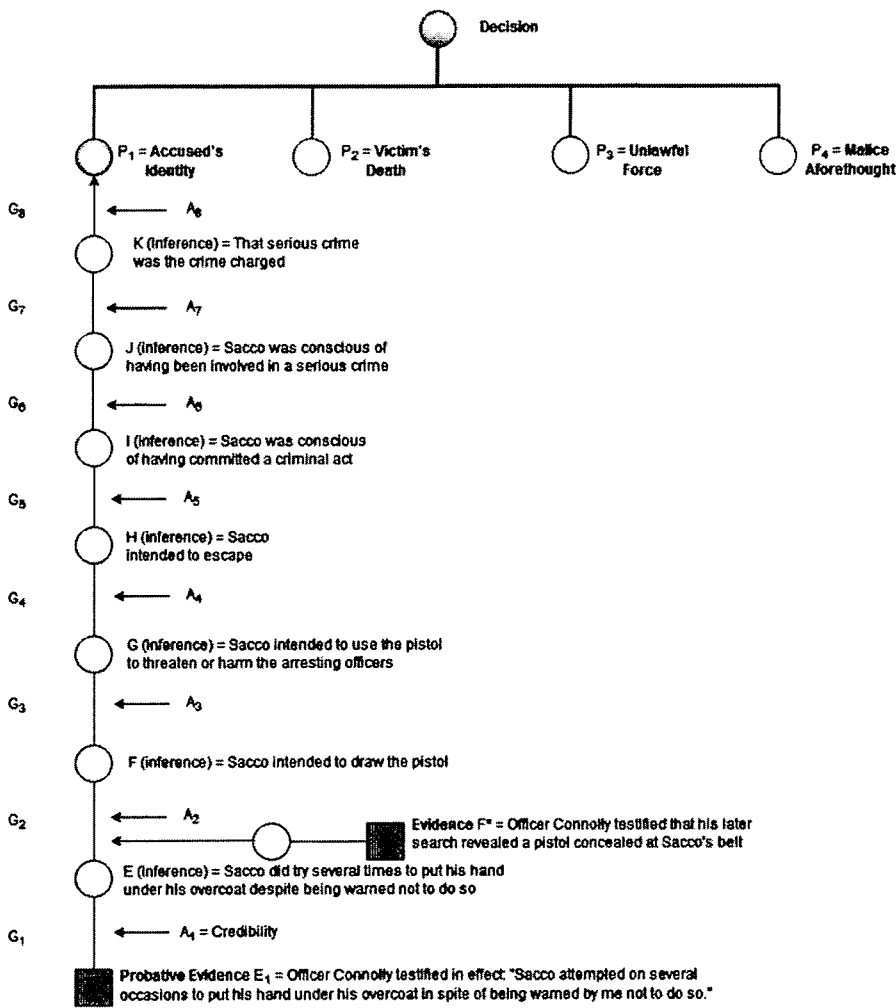
Figure 4: Inferential Reasoning in Sacco & Vanzetti

Figure 4 is admittedly complicated. And it is the tip of the iceberg, because a chain or chains had to be constructed for every piece of probative evidence in the long trial. But it is not unrealistic. The diagrammed chain is a necessary minimum of reasoning. The factfinder could not properly avoid the task of inferring from each piece of evidence to the element or elements that the evidence tended to prove or disprove.

To begin somewhere, the generalizations involved in Figure 4 would run something like this:

G_1 = Any events to which an officer testified under oath usually occurred

G_2 = Arrestees with a concealed weapon sometimes attempt to draw it

G_3 = Arrestees who draw a weapon usually intend to use it

G_4 = Arrestees who use a weapon often intend to escape

G_5 = Arrestees who intend to escape usually are conscious of having committed a criminal act

G_6 = Arrestees who intend to escape and are conscious of having committed a criminal act frequently will be conscious of having committed a serious crime, like a robbery and shooting

G_7 = Arrestees who are conscious of having committed a serious crime often will be conscious of having committed the charged crime

G_8 = Arrestees with consciousness of guilt as to the charged crime usually committed it

As to each generalization, ancillary considerations are relevant. For a prime example, A_1 bears on the primary credibility of Connolly's testimony.[119] Ancillary chains of inferences as to his veracity, objectivity, and sensitivity would feed into the G_1 generalization. Also, ancillary evidence, such as the facts that Connolly never mentioned Sacco's hand movements until Connolly testified[120] and that pressures existed for him to testify a certain way, would undercut G_1. Similarly, A_2 considerations contain multiple strains, including evidence F*. For other examples, the A_6 pieces of evidence that Sacco was carrying a pistol at arrest and that Sacco was a despised radical who had been involved in distributing anarchist literature seriously undercut generalization G_6 that he was likely conscious of committing some serious offense, while the A_7 ancillary consideration that the prosecutors' evidence had suggested no other serious crime committed by Sacco would support generalization G_7 that the serious crime in consciousness was the charged crime.

Other pieces of evidence E_2, etc., such as lies Sacco told police officers, would also show Sacco's consciousness of guilt (K). And, independently of his consciousness of guilt, still other evidence E_{n+1}, etc., such as bullet forensics or the carrying of a pistol itself, would help support or undercut identity (P_1).

---

119. *See id.* at 109–12 (discussing the testimony's credibility).

120. *See* YOUNG & KAISER, *supra* note 111, at 67–70, 162–63 (making this observation).

B. Probative Force

The "probative force" of evidence means how powerful the piece of evidence is in favoring or disfavoring some element in the case at hand.[121] In Figure 3, G_1, G_2, G_3, and G_4, as well as A_1, A_2, A_3, and A_4, all contribute to determining the probative force of E_1 on P_1. The force will depend on (1) the weight and credibility of the evidence, which turns on analysis of (2) the force of each necessary inference from E_1, through E, F, and G, to P_1, as determined by the strength of its generalization and as ancillarily affected by its accuracy, and (3) the force of the whole chain of inferences taken together.[122] Yet, in trying to measure probative force, one will encounter great disagreement among legal theorists about how grading of probative force should be done. Here, this Article moves into contested territory.

First, to begin with definitions, evidence is relevant if, taken as true, it allows us to revise our belief or disbelief in some element.[123] The "likelihood ratio" expresses the relative likelihood of the existence of the given item of evidence upon the alternative assumptions that the fact to be proved exists and that the fact does not exist, which reduces to $Bel(P_1|E_1) \div Bel(P_1|\text{not-}E_1)$.[124] The likelihood ratio thus conveys relevance by showing whether the evidence allows the factfinder to revise a prior belief in the element (which is zero before the item of evidence) to form a different posterior belief or disbelief (after taking the new evidence into account). To be more precise, the "weight" of the evidence measures the extent of that revision, once matters of primary credibility are put to the side.[125] The factfinder's fixing the weight of the evidence will inevitably be a bit of a stab at judgment. But after we inject the primary credibility, we have the raw materials for estimating probative force.

---

121. *See* ANDERSON, SCHUM & TWINING, *supra* note 3, at 71, 226–29, 384–85 (defining "probative force").

122. *See id.* at 65–71, 226.

123. *See* FED. R. EVID. 401(a) (defining "relevant").

124. *See* 1 MCCORMICK ON EVIDENCE, *supra* note 14, at 996–98 (discussing "likelihood ratio"); *infra* text accompanying note 169. To be less Bayesian, perhaps the likelihood ratio should be defined in terms of beliefs and disbeliefs, as $[Bel(P_1|E_1) \div Bel(\text{not-}P_1|E_1)] \div [Bel(P_1|\text{not-}E_1) \div Bel(\text{not-}P_1|\text{not-}E_1)]$. Moreover, speaking in terms of beliefs and disbeliefs, rather than Bayes' theorem, avoids the "analytic gap between epistemic relevance and probability" that troubled Professor Pardo. Michael S. Pardo, *The Nature and Purpose of Evidence Theory*, 66 VAND. L. REV. 547, 583 (2013); *see* L. Jonathan Cohen, *Some Steps Towards a General Theory of Relevance*, 101 SYNTHESE 171, 181 (1994) ("Anything that can sanction a reason, even if an incomplete or inconclusive reason, for accepting a particular type of proposition as a correct answer, or for rejecting it as an incorrect answer, to an askable type of question can count as a criterion of relevance.").

125. On "credibility," see *supra* text accompanying note 110.

Second, assessment of belief in a specific inference presents another challenge. As already explained, the factfinder will believe the step in going, say, from E to F to a degree x, which depends on the strength of the generalization G_2 as adjusted for its accuracy revealed by ancillary considerations A_2.[126]   The probative forces of the multiple ancillary considerations need to be combined, with each other and with the generalization's underlying data, by the weighted-arithmetic-averaging method to be described in Part III. So, measuring the conditional belief x in that step will likely be a rough process.

Third, the biggest stumbling block, or source of disagreement, concerns how to combine the inferential steps into an overall belief function for the element generated by the piece of evidence. Conditional beliefs in the steps to E, F, G, and P_1 have to be accepted in order to support the element. The process is conjunction because each step in the chain must be accepted. The evidence's affirmative probative force so reduces to a problem of conjoining a string of degrees of belief, which is resolved next.[127]

i.   Product Rule

Classical logic would not even hesitate at the conjunction problem.[128] The probability operation for *a* AND *b* is multiplication of the probabilities of independent events, and multiplication of Prob(*a*) by Prob(*b*|*a*) for interdependent events.[129] These simple calculations, which many people call collectively the product rule (or the general product rule or the chain rule), are the right way to compute odds for a future conjoint event.[130]

In trying to perform the legal task of finding facts, however, the logical and practical problems of applying probability and its product rule become legion.[131] The most obvious is that the product rule would start producing nonsensical results for law cases as the number of conjoined facts starts

---

126. *See supra* note 105 and accompanying text.

127. Combining disbeliefs will involve the MAX rule, as later explained. *See infra* text accompanying note 142.

128. *See* Morgan, *supra* note 18, at 944 ("Now it must be obvious that the value of item A as probative of F varies inversely with the number of inferences between A and F.").

129. JOHN ALLEN PAULOS, BEYOND NUMERACY: RUMINATIONS OF A NUMBERS MAN 149–50 (1991).

130. *See* COHEN, *supra* note 30, at 58–59.

131. *See* Clermont, *supra* note 22, at 1459–62, 1484–86 (contrasting traditional probability unfavorably with multivalent logic for purposes of factfinding).

increasing.[132] Proof would become nearly impossible. Indeed, you would end up believing almost nothing in the world, as just about any belief rests on a chain of conjoined inferences:

> Wigmorean analysis provides a technique for identifying and making explicit the generalizations involved at each step of an argument. At the same time, that analysis may cumulatively seem like an invitation to extreme skepticism. It regularly provides strong ammunition for attacking an opponent's argument and for questioning one's own. So many generalizations seem so vulnerable in so many respects that one may be led to the conclusion that all arguments about evidence are built on shifting sands.[133]

The book *Analysis of Evidence* suggests that the practicing lawyer need not worry about this philosophical point.[134] But I disagree. The point raises major theoretical[135] and practical[136] difficulties. Conjoining propositions is a universal and constant task in factfinding, as this discussion of combining inferential steps illustrates. Conjunction is necessary for any fact resting on multiple inferences, and every found fact rests on multiple inferences because every finding begins with an inference about the credibility of the evidence.[137] The reality is thus that conjunction is a necessary act in *every* fact found by a

---

132. *See* Branion v. Gramly, 855 F.2d 1256, 1264 (7th Cir. 1988) (rejecting multiplication of odds in a legal case by saying: "Every event, if specified in detail, is extremely improbable; indeed, with *enough* detail it is unique in the history of the universe. It is always possible to take some probabilities, small to start with, and multiply them for effect.").

133. ANDERSON, SCHUM & TWINING, *supra* note 3, at 102.

134. *See id.* (saying such problems are "not an immediate concern for the practicing lawyer"); *see also* Risinger, *supra* note 20, at 811 ("[A]rguing either radical skepticism or the primacy of some form of philosophical idealism will not cut any ice in a courtroom.").

135. *See* Michael S. Pardo, *The Paradoxes of Legal Proof: A Critical Guide*, 99 B.U. L. REV. 233, 280–82, 288–89 (2019) (arguing that paradoxes, like conjunction, force one to grapple with the fundamental issues). 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) (stating the same). These 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) (resolving the aggregation paradox).

136. Many dismiss the conjunction paradox as a mere theoretical wrinkle without practical worry, arguing for example that most cases involve a single disputed issue and that multiple issues are seldom independent. Clermont, *supra* note 22, at 1493 n.117. *But see id.* at 1493–95 (rebutting a number of such theorists).

137. *See supra* text accompanying note 87 ("There is no such thing as direct evidence.").

legal factfinder! Moreover, the task of conjunction extends well beyond trial, into pretrial devices and other legal-factfinding settings inside and outside courts, as well as into settlement negotiations and other law-office applications of law that depend on expected factfinding. Finally, we all intuitively find and combine facts in countless settings of daily life.

The point here is so important that it bears repeating. Conjunction is not an arcane oddity. Any thought worthy of being called a thought entails reasoning by conjoined inferences. René Descartes would never have been able to get beyond, or even to, "I think, therefore I am."[138] No belief worthy of being called a belief would survive the application of the product rule. Descartes would never have been able to infer his way to a belief in God.[139] Therefore, the law too must get conjunction right, for the sake of accuracy, fairness, and efficiency.

ii. MIN Rule

Multivalent logic recognizes that combining beliefs is a mathematical task different from combining odds.[140] The highly developed and widely accepted mathematics for combining degrees of beliefs instructs that the conjunction has a degree of belief equal to the weakest of the conjoined beliefs, in accordance with the so-called MIN rule that appears as a basic operator of the multivalent-belief logic system.[141] This conjunction operator is a more general replacement for the product rule, which appears as a special case where all values can only

---

138. *See* SIMON BLACKBURN, THINK: A COMPELLING INTRODUCTION TO PHILOSOPHY 19–20 (1999) (explaining that the premise of thinking rests on inference from alternative possibilities: a mind that doubts or a deception of a mind by the Evil One).

139. *See id.* at 34 (reconstructing the argument as: "I have the idea of a perfect being. This idea must have a cause. A cause must be at least as perfect as its effect. So something at least as perfect as my idea caused it. Therefore such a thing exists. But the thing must be perfect, that is, God.").

140. *See* Huber, *supra* note 44, at 10 (stating "that fair betting ratios should indeed obey the probability calculus, but that degrees of belief, being different from fair betting ratios, need not").

141. *See* Clermont, *supra* note 31, at 50–51 (discussing operators). Philosophers and logicians agree with the mathematicians. *See* COHEN, *supra* note 30, at 89–91, 220–22, 265–67 (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); Dubois & Prade, *A Set Theoretic View*, *supra* note 69, at 403 (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*, *in* CUTS AND CLOUDS, *supra* note 53, at 438, 438 (arguing against the product rule and in favor of the MIN rule). For a formal proof that the MIN and MAX rules make sense in multivalent logic, see Clermont, *supra* note 31, at 51 n.32, 67–68.

be 1 or 0. Moreover, the same MIN rule applies whether the beliefs are independent or interdependent, unlike the product rule. So, if a person believes *a* and believes *b*, then by the principle of conjunctive closure the person believes *a* and *b* together, although of course not more than the person believes *a* or *b* separately.[142]

The new degree of belief will be the minimum of the affirmative beliefs to be conjoined. The belief measure for conjunction of beliefs is *Bel(*E AND F*)* = MIN(*Bel(E)*, *Bel(F)*);[143] in Figure 5, the belief in the conjunction of the beliefs in propositions E and F is *Bel(E)*. However, the new degree of disbelief will be the maximum of the disjoined disbeliefs, following the so-called MAX rule. The belief measure for disjunction of disbeliefs is *Bel(*not-E OR not-F*)* = MAX(*Bel(*not-E*)*, *Bel(*not-F*)*);[144] in Figure 5, the belief in the disjunction is *Bel(*not-E*)*. Belief and disbelief will range from 0 to 1, and they should add to less than one. The uncommitted belief reflects epistemic uncertainty.

Bel(E)	uncommitted	*Bel(not-E)*

Bel(F)	uncommitted	*Bel(not-F)*

0   1
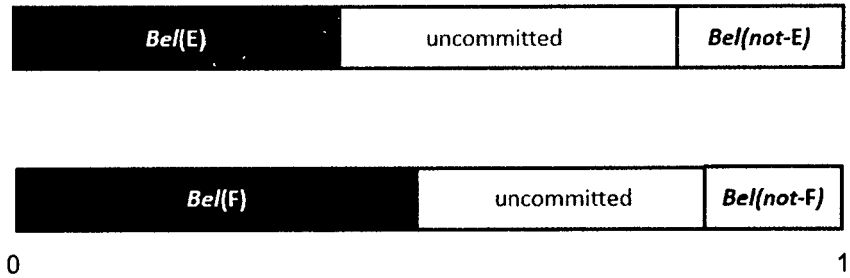
Figure 5: Application of MIN Operator

Why do beliefs combine differently from odds? Probability built on bivalent logic is a so-called additive system, meaning that *p* as the chance of

---

142. *See* Simon J. Evnine, *Believing Conjunctions*, 118 SYNTHESE 201, 201, 214, 222 (1999) (stating the principle as "If S is rational, then if S believes *A* and S believes *B*, then S believes *A and B*," and defending the principle as generally valid while recognizing the remaining difficulties of the lottery and preface paradoxes); Hannes Leitgeb, *The Review Paradox: On the Diachronic Costs of Not Closing Rational Belief Under Conjunction*, 48 NOÛS 781 (2014) (stating a similar idea). These theorists attempt the near-impossible task of deriving the intuitive principle of conjunction closure while still assuming bivalence. If they were to recognize that beliefs are multivalent, then the principle would appear as a logical operator (and those paradoxes would become manageable).

143. *See supra* notes 141–42.

144. *See id.*

truth and $1 - p$ as the chance of falsity add to one.[145] Multivalent belief and disbelief normally add to less than one, making belief function theory into a nonadditive system. This distinction proves critical, because the product rule prevails only in additive systems.[146] The reason is that the product rule depends on its assumption that to the degree anything is not proven true, it is false, rather than merely not proven.[147] Let me explain.

On the one hand, the probabilistic odds of a dictate under bivalence the odds of *not-a*. That is, the complement of the probabilistic 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 <u>or</u> b being revealed as false goes up. So, conjoining a <u>and</u> b means increased odds of *not-a* <u>or</u> *not-b*. Accordingly, the chance of a <u>and</u> b being true goes down, by a multiplicative amount in accordance with the product rule.

On the other hand, under multivalent logic, a belief is an evidence-based measure of sureness about the real world. The complement of a belief is not a disbelief, but it is instead the degree to which the belief was not proven. As the factfinder shifts from evidence of a to evidence of b, the degree of a's not being proven has no effect on or interaction with the degree of b's not being proven. Accordingly, the sureness of a and b does not slip below the smaller of the two. To put it in other words, if my sureness about a is adequate to form a belief as I navigate an uncertain world, the fact that I also believe b does nothing to lessen my belief in a, with the consequence that I believe a and b together.

Think of the MIN rule in action. Think first of separate beliefs. A typical legal test asks whether a violation of some procedural law occurred and whether

---

145. Additivity is one of probability's three basic Kolmogorov axioms: "If two events cannot happen jointly, the probability that one or the other occurs is equal to the sum of their separate probabilities." ANDERSON, SCHUM & TWINING, *supra* note 3, at 251; *see* Hájek, *supra* note 30, § 1 (listing the three axioms as non-negativity, normalization, and additivity). In an additive system, like bivalent logic, a set and its complement add to the universe, or one. Thus, the probability that an event will happen and the probability that it will not happen add to one. By contrast, multivalent logic, as employed in belief functions, rejects as an assumption the law of the excluded middle and its consequence of additivity. *See supra* note 31. Thus, degrees of belief and disbelief in a fact, where the factfinder retains some belief as uncommitted between true and false, do not add to one.

146. The product rule derives from the givens of probability, which are the bivalence assumption and Kolmogorov's three axioms including additivity, and so it is inoperative when those conditions do not hold. *See* Brian R. Gaines, *Fuzzy and Probability Uncertainty Logics*, 38 INFO. & CONTROL 154, 155–59, 161 (1978) (saying that "both multiplication/addition, and max/min, connectives may be seen to arise from constraints on an underlying probability logic").

147. *See infra* note 150.

it affected outcome. No one would ever think of applying the product rule, and so nobody would require an exaggerated showing on the two issues.

Next, the beliefs in inferences E and F are not totally separate beliefs, in that the belief in F is a conditional belief dependent on E. But they conjoin the same way. Let $\Rightarrow$ mean "implies." Theorists have derived an inference rule that says if α and $(\alpha \Rightarrow \beta)$ are proved as beliefs to degree λ and μ, respectively, then we can assert β at degree MIN(λ, μ).[148] So, if one views E as a belief and views the inference (E $\Rightarrow$ F) as a conditional belief, then the belief in F will be the minimum of those two beliefs. As one goes up the chain of inferences, one will believe the latest inference to the degree of the minimum of all the preceding beliefs. That is, one will believe P_1 to the extent of the weakest link in the chain.

In sum, conjoining beliefs is fundamentally different from figuring joint odds. The logical fact is that degree of belief and probability of truth are different measures. The mathematical fact is that degrees of belief follow the MIN rule, not the product rule.

iii. Correct Approach

Should the factfinding process treat the inferences E, F, G, and P_1 as bivalent probabilities or as multivalent beliefs? The proper approach is to view them not as odds of events all happening together to produce a result, but as degrees of belief that coexist. If one accepts that the legal factfinding process involves conjoining multivalent beliefs, then multivalent mathematics will give the right answer. That is, the MIN rule is the correct approach because it gives the *accurate* calculation.

Some commentators have reacted that switching from bivalence to multivalence cannot change reality and therefore cannot change computational results.[149] That view is the result of misunderstanding. One calculates to a different result not by some magical shift of perspective, but because one is combining a different kind of measure. One calculates to a different result not by some magical kind of logic, but by utilizing a more sensitive measure. Beliefs can convey the full output of factfinding. Beliefs include a measure of

---

148. *See* GIANGIACOMO GERLA, FUZZY LOGIC: MATHEMATICAL TOOLS FOR APPROXIMATE REASONING 113–16 (2001) (describing so-called necessity logic, which is an offspring of possibility theory); *cf.* DIDIER DUBOIS & HENRI PRADE, FUZZY SETS AND SYSTEMS: THEORY AND APPLICATIONS 167–68 (1980) (arguing, outside the context of necessity logic, that the multivalent belief in β is not less than MIN(λ, μ)).

149. *E.g.*, Ronald J. Allen & Michael S. Pardo, *Relative Plausibility and Its Critics*, 23 INT'L J. EVIDENCE & PROOF 5, 49–58 (2019).

epistemic uncertainty, while probabilities do not. To combine factfindings while still carrying forward all the uncertainties, the factfinder must employ a nonadditive system, where belief and disbelief do not necessarily add to one and where some belief can remain uncommitted.[150] The product rule would instead discard uncertainties before multiplication, and thereby give the wrong answer on which to base decision. Therefore, for accuracy's sake, the law should combine beliefs by multivalent logic's rules.

A simple example can demonstrate the accuracy of the MIN rule. Say that $Bel(a) = 0.90$ and $Bel(not\text{-}a) = 0.10$ on perfect evidence of two independent elements, and that $Bel(b) = 0.12$ and $Bel(not\text{-}b) = 0.03$ on very imperfect evidence. Then, $Bel(a$ AND $b) = 0.12$. If instead one were to use probabilities, $Prob(a) = 90\%$ and $Prob(b) = 80\%$.[151] Applying the product rule then yields a 72% probability for the conjunction. Which is the more accurate representation of the strength of the conjunction, a representation on which one will base future actions? Given airtight and thin evidence like this, has the proponent made a weak showing of the conjunction, say 0.12, or a fairly strong showing, say 72%? It is easy to see that probability theory introduced an error when it discarded uncertainties and converted the weak showing on *b* to 80%. An 80% chance on airtight evidence is very different from an 80% chance on thin evidence.

That example is a slam-dunk demonstration of the superior accuracy of the MIN rule. Still, some more light comes from a second example that uses vagueness, which is the other major kind of uncertainty captured by multivalent logic and not by traditional probability. Imagine you want to know if Tom is

---

150. A conjoined belief is proven only to the degree that no contradictory disbelief is entertainable—and a contradictory disbelief is entertainable all the way to the degree measured by the largest sum of disbelief and uncommitted belief on any component finding. *See* Huber, *supra* note 44, at 14 (discussing conjunction and disjunction). Huber argues that shifting from Shafer's mathematical formulas to compatible set theory, such as possibility theory, *see supra* note 69, makes the MIN and MAX rules easier to picture. Thus, if $Bel(E)$ expresses a degree of membership in the set of fully believed facts, $1-Bel(E)$ is the possibility or entertainability of not-E given uncertainty. If one then tries to picture $Bel(E \cap F)$, it will exist where neither not-E nor not-F is possible, which by the MAX rule is the maximum possibility of either not-E or not-F. *See* Huber, *supra* note 44, at 14 ("The idea is, roughly, that a proposition is at least as possible as all of the possibilities it comprises, and no more possible than the 'most possible' possibility either."). Alternatively put, E and F will be conjoinedly necessary where each of E and F is necessary, which is measurable by the MIN rule. That is, to the extent not-E or not-F is possibly true, the conjunction of E and F cannot be necessarily true. *See* Dubois & Prade, *A Set Theoretic View*, *supra* note 69, at 402–03 (discussing the relation of belief functions and possibility theory).

151. Normalizing the beliefs is one means to perform the pignistic transform. *See supra* note 53 and accompanying text. Normalizing 0.12 and 0.03 yields 80% and 20%.

tall and smart. Let X be the universe of men, and let A be the set of tall men and B be the assumedly independent set of smart men. Tom is a 0.50 member of A and a 0.40 member of B, which means something like, "Tom is of roughly average height" and "Tom is not that smart." The MIN rule would yield a 0.40 belief in that intersection of tall and smart traits, or "Because Tom is not really tall and Tom is not that smart, Tom is not such a tall, smart man." The probabilistic calculation, however, would yield a 20% chance, or "Because Tom is likely not tall and Tom is likely not smart, Tom is likely a short, dumb man." The key insight is that Tom's memberships do not mean that Tom has a 50% chance of being so tiny as to be completely outside the set of tall men and a 60% chance of being truly dumb; instead, they mean a 0.50 and 0.60 degree of the factfinder's not being convinced of tallness or smartness. Multiplication of probabilities gives the chance that Tom is both indubitably tall rather than short and indubitably smart rather than dumb, while what we want to know is the degree to which he belongs to the set of men classifiable as both tall and smart. The inappropriateness of using the product rule becomes much more obvious as one combines more and more elements in the calculation, such as tall, smart, rich, and bald men. The product will approach 0%, even if some of the values are very high, while the MIN rule will go no lower than the minimum value.

Inferential reasoning is thus rightly viewed as involving links in a chain whose strength together is the strength of its weakest link. But let me explain this correct approach in yet a different way. Uncertain propositions fall mainly into one of two piles: one of bivalent measures for which the product rule suffices, and another of multivalent measures for which rationality requires the more general MIN rule. The probability of revealed truth, with all views committed between true and false as for betting purposes, falls into the first pile. The belief in finding truth, with its accompanying nonbelief and disbelief, goes into the second pile.

The theorist could argue coherently for putting legal factfinding in either pile, a decision that then dictates the proper combination rule. However, based on extant doctrine expressed in cases and judicial instructions that mandate element-by-element decisionmaking, there is little doubt that the law has cast legal factfinding into the second pile.[152] I am further contending that the law's choice was a wise one. The law has no cogent interest in the odds for betting

---

152. *See, e.g., In re* Corrugated Container Antitrust Litig., 756 F.2d 411, 416–17 (5th Cir. 1985) (requiring proof of each element to a preponderance); *see infra* text accompanying note 196 (quoting instruction).

on unattainable truth, a measurement that ignores all kinds of meaningful uncertainty. Choosing the second pile avoids the many difficulties of dealing in probabilities. Moreover, as already explained, the second pile fits the image, and so clarifies the theory, of legal factfinding—a process that aims to measure the degree of belief established by the evidence, while leaving some belief uncommitted to reflect epistemic uncertainty. Finally, whatever the law wants, it might get beliefs from its factfinders. They are painfully conscious of uncertainty when they pronounce what they believe. They are not about to project the odds of a revelatory event that is difficult even to verbalize.

Defending this application of the MIN rule for conjoining beliefs is normative. Although the MIN rule could also be descriptive of actual legal practice, the safest thing to say about the practice of conjunction is that humans are not perfectly logical.[153] However, one might defensibly go further to say that nothing in practice suggests that legal factfinders apply the product rule in inferential reasoning, as shown by humans' ready willingness to form inferential beliefs that would never survive the product rule.[154] More rigorous

---

153. The conjunction fallacy, which is not to be confused with the conjunction paradox, supposedly posed by the law's element-by-element application of the standard of proof, emerged from the famous Linda problem: based on a description of Linda (*"Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations."*), a strong majority of people rank "Linda is a bank teller and is active in the feminist movement" as being more probable than "Linda is a bank teller." That result seems illogical, in that specific conditions do not combine to be more probable than a single condition. *Compare* DANIEL KAHNEMAN, THINKING, FAST AND SLOW 156–65 (2011) (discussing this conjunction fallacy), *with* Ralph Hertwig & Gerd Gigerenzer, *The "Conjunction Fallacy" Revisited: How Intelligent Inferences Look Like Reasoning Errors*, 12 J. BEHAV. DECISION MAKING 275, 275 (1999) ("We conclude that a failure to recognize the human capacity for semantic and pragmatic inference can lead rational responses to be misclassified as fallacies.").

154. The just-described Linda problem demonstrates that humans often neglect the product rule. But it may further indicate a neglect of the difference between aleatory uncertainty and epistemic uncertainty among the testers. If the categories of bank teller and active feminist are viewed as crisp categories, subject only to a frequency measure of randomness, then the product rule is appropriate. But if the categorization is subject to epistemic uncertainties, then the product rule does not apply—and the applicable MIN rule implies a much lower level of illogic among the subjects, because the conjunction would be as likely as the unlikely employment as a bank teller. Indeed, many experiments show that if the Linda problem is phrased in expressly frequentist terms, the subjects are much more inclined to employ the product rule. *See, e.g.*, Amos Tversky & Daniel Kahneman, *Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment*, 90 PSYCHOL. REV. 293, 309 (1983) (saying that phrasing the problem in terms of frequencies "markedly reduced the incidence of the conjunction fallacy"). "People do not think of the Linda problem in terms of frequencies, or as an exercise in probabilistic reasoning. They see the description of Linda as a story about a real person,

testing is sketchy, but likewise provides no evidence that humans use the product rule for combining degrees of belief.[155] In one experiment, subjects had to judge the degree to which objects such as cars and wine barrels were "metallic containers" and, three days later, the degree to which they were "metallic objects" and "containers;" the subjects tended to favor the MIN rule over the product rule.[156] In sum, even if it is not established that humans default to a strict MIN rule, it is quite conceivable that rough and ready factfinding emphasizes the weakest conjoined fact.[157]

In summary, when an inference rests on an inference from a piece of evidence, the conjoined strength of belief drops to the likelihood of the least likely inferential step. The conjoined belief in the element is as strong as the weakest link, fixing the affirmative probative force of the piece of evidence. If the piece of evidence supports a disbelief in the element, the disbelief is as strong as the greatest disbelief in the inferential reasoning. I thus argue that multivalent logic is the correct way to proceed in this first step of evidence processing, and I further argue it is feasible for human factfinders and may encapsulate what they actually do when inferring from an item of evidence to an element of the case.

* * *

Now, let me climactically generalize about the kinds of decisions for which beliefs are the appropriate measure. The distinction is certainly not legal versus nonlegal decisionmaking. In daily life, some decisions about future events rest

---

and the biography of Linda which ends only by identifying her as a bank teller is not, without more information, a satisfactory account." KAY & KING, *supra* note 46, at 91.

155. *See* Rami Zwick, David V. Budescu & Thomas S. Wallsten, *An Empirical Study of the Interpretation of Linguistic Probabilities*, *in* FUZZY SETS IN PSYCHOLOGY 91, 91 (Tamás Zétényi ed., 1988) (concluding that people do not naturally use the product rule).

156. *See* U. Thole, H.-J. Zimmermann & P. Zysno, *On the Suitability of Minimum and Product Operators for the Intersection of Fuzzy Sets*, 2 FUZZY SETS & SYS. 167, 168 (1979); *cf.* Evnine, *supra* note 142, at 214 ("Take any two propositions that are of no special logical or emotional significance to someone, say that grass is green and that snow is white. I contend (this is intended as an empirical observation about our practice) that if we have good grounds for attributing belief in each of these to a person, then, absent any special circumstances, that is all we need to attribute to that person a belief in their conjunction. Or if, for example, we are summing up a position someone has just explained at some length, we can do so by attributing to that person a large conjunctive belief. This conjunctive belief describes in a single proposition a number of different propositions that were expressed severally over a period of time."). Other experiments give some support to the MAX rule for disjunction. *See* Zwick, Budescu & Wallsten, *supra* note 155, at 98, 115 (discussing a previous study and reporting a new one).

157. *Cf.* Zwick, Budescu & Wallsten, *supra* note 155, at 114–16 (finding an averaging method employed for conjunction).

on traditional probabilities (should I carry an umbrella today?) and some should instead employ beliefs (should I buy a gun for self-protection?—a decision that involves combining contestable evidence with imprecise values and that thus prevents the conscientious decisionmaker from allocating all belief on each factor to either true or false, unlike the umbrella decision that is dominated by simple probabilities). Nor is the distinction between legal factfinding and other legal decisions. In law outside factfinding *sensu stricto*, some decisions rest on probabilities (should the court grant a preliminary injunction in light of expected costs?) and some involve beliefs (should we give a remedy that depends on whether a violation of some procedural law occurred and whether it affected outcome?). For legal factfinding alone, other distinctions—such as unsettlable or unknowable facts as opposed to facts that will be eventually revealed, or past versus future facts, or vague versus crisp facts—are suggestive but remain under- or over-inclusive. The real test is whether for the sake of accuracy, *the decisionmaker needs to keep track of epistemic uncertainties* in addition to any aleatory uncertainty. Legal factfinders must hold all the kinds of uncertainty in mind when combining their findings.[158] Traditional probability ignores epistemic ignorance resulting from imperfect evidence and ignores epistemic indeterminacy resulting from vagueness and the like, but multivalent beliefs retain uncommitted belief as a measure for epistemic uncertainties.

Thus, the nature of the assigned decisionmaking task determines the appropriate measure. Does the assigner want the decisionmaker to account for the weakness or absence of evidence or to measure any vague elements? That is the case for factfinding.[159] Or does the assigner want the decisionmaker to

---

158. *See* Martina Fedel, *Uncertainty, Indeterminacy and Fuzziness: A Probabilistic Approach, in* PROBABILITY, UNCERTAINTY AND RATIONALITY 219, 219 (Hykel Hosni & Franco Montagna eds., 2010) ("In practical reasoning, whether it is aimed at drawing inferences or making decisions, we need to give appropriate weight both to our uncertainty, about facts or events or consequences of our actions, and to the indeterminacy that arises from our ignorance about [these] matters.").

159. The analysis of finding fact extends to lawmaking. *Cf.* Kevin M. Clermont, *Rules, Standards, and Such*, 68 BUFF. L. REV. 751, 785–86 (2020) (using multivalent ideas to guide in formulating law). It also extends to scientific thinking, in tasks of both inferential reasoning and constructive theorizing. *See* Kevin M. Clermont, *The Logic of Uncertainty in Law and Life*, 19 LAW PROBABILITY & RISK 181, 181 (2020). Indeed, it extends to all areas of decisionmaking that involve epistemic uncertainty. As beautifully put as a general proposition by GARY LAWSON, EVIDENCE OF THE LAW: PROVING LEGAL CLAIMS 2 (2017):

> As it happens, law is an especially fruitful avenue for exploring these ideas precisely because, at least with respect to proof of facts, the law sometimes thinks self-consciously about some of the abstract features of proof and how they bear

commit all belief in order to announce odds? Outside of factfinding, the law often wants this latter approach, whereby the decisionmaker disregards epistemic uncertainties. An example would be when a court is reviewing (or restraining in advance) a jury's factfinding.[160] There we do not expect the reviewer to retain uncommitted belief in applying the standard of review. The "evidence" for applying the standard of review is complete, in that the question is whether the jury impermissibly erred in deciding the case *on the given proof*. We want from the reviewer a yes/no answer based on the odds of the jury's having unreasonably weighed its beliefs and disbeliefs to find for the winner, with the odds' complement being the probability of the jury's having had the authority to find for the winner. We do not want the reviewer's "belief."[161]

III. AGGREGATING PIECES OF EVIDENCE

The combined probative force of the pieces of evidence on an element can nudge down (as by conflict or contradiction) or up (as by convergence or corroboration)[162] with a new piece of evidence.[163] In the Sacco & Vanzetti example of evidence of consciousness of guilt, we have seen how other evidence E_{n+1}, etc., such as eyewitness testimony and bullet forensics, would help undercut or support the same element of identity, P_1.

---

on decision making. But I believe that the structure of proof that I discuss here is universalizable to any cognitive endeavor.

160. Proving law may be another example where epistemic uncertainties drop out. LAWSON, *supra* note 159, at 28–44 (showing powerfully that questions of law are epistemically equivalent to questions of fact). But that point does not imply that the rules for pronouncing law need be the same as for finding fact. The system does draw a law/fact line, so that adversary principles do not apply to law questions and burden of proof has no role to play there. Normally, the standard of proof seems to be "more likely than not," in the sense of the judge's adopted reading of the law being better than any alternative reading. *See* CLERMONT, *supra* note 23, at 95–99 (discussing standards of decision for issues of law); Clermont, *supra* note 22, at 1497–505 (discussing standards like inference to the best explanation). Given our system of stare decisis, there is no burden of proof in the sense of the risk of nonpersuasion. We would not want the law to turn on which party is named the proponent or on how good a job the proponent did.

161. *See* Clermont, *supra* note 22, at 1495 n.121 (discussing review as a two-step process).

162. On these types of evidence, see *supra* text accompanying note 89 and *infra* text accompanying note 187.

163. *See* Al-Adahi v. Obama, 613 F.3d 1102, 1105 (D.C. Cir. 2010) ("Those who do not take into account conditional probability are prone to making mistakes in judging evidence. They may think that if a particular fact does not itself prove the ultimate proposition (e.g., whether the detainee was part of al-Qaida), the fact may be tossed aside and the next fact may be evaluated as if the first did not exist.").

Aggregation of evidence takes place when inferring from a single piece of evidence to an element.[164] But now our focus shifts to aggregating separate pieces of evidence that bear on a single element. As a normative matter, then, how should the factfinder aggregate evidence to calculate the new composite probative force of that evidence taken together? Figure 6 represents the aggregation of all evidence on P_1.
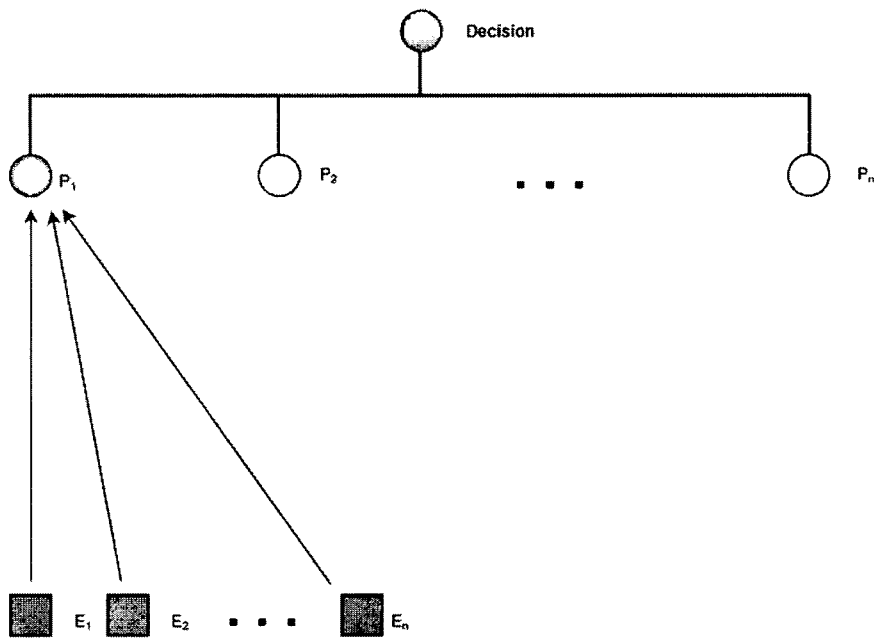

Figure 6: Representation of Aggregating Evidence

This mental process of aggregating evidence is completely distinguishable from the above-treated process of conjoining beliefs. Aggregating evidence does not centrally involve conjunction, where each belief must be accepted, but instead calculates how items of reinforcing or undercutting evidence add to or subtract from one another to form a composite belief (and a composite disbelief) in P_1.[165]

---

164. *See supra* text accompanying notes 108, 127 (combining generalization with ancillary considerations).

165. *See* SUSAN HAACK, *Proving Causation: The Weight of Combined Evidence, in* EVIDENCE MATTERS, *supra* note 44, at 208, 216–28 (discussing the task generally as an epistemological matter).

Aggregating evidence is perhaps easier than conjoining beliefs to picture, as it seems simpler and more intuitive, but it is much harder to formalize. On the one hand, conjoining (and disjoining) beliefs follows directly from the basic operators of the logic system. The product rule governs in bivalent systems because those systems are additive. For multivalent beliefs the operator would be the MIN (or MAX) rule, because belief function theory is nonadditive. On the other hand, aggregating evidence is a mathematical problem whose solution must be derived under the prevailing logic system, be it bivalent or multivalent. The solution is Bayes' theorem under classical logic but becomes an even more complicated problem under multivalent logic's belief function theory.

A. *Bayes' Theorem*

Among those theorists who view evidence probabilistically, the dominant answer for aggregating evidence invokes Bayes' theorem.[166] It links the perceived probability before and after observing new evidence.[167] By mathematically updating the initial probability with the new evidence, one gets an updated probability.

The starting point for Bayes' theorem is Prob(A), the prior probability of A. Then the posterior probability of A, after accounting for new evidence B, is the conditional probability Prob(A|B), which may be read as the probability that A will occur if B is known certainly to have occurred.[168] Prob(A|B) calculates to be Prob(A) multiplied by the support B provides for A, a support that Thomas Bayes (or really Pierre Simon Laplace) equated to Prob(B|A) ÷ Prob(B). The measures Prob(A) and Prob(B) are the probabilities of observing A and B independently of each other. So, here is the theorem:[169]

$$\text{Prob(A|B)} = [\text{Prob(B|A)} \cdot \text{Prob(A)}] \div \text{Prob(B)}$$

The Bayesian likelihood ratio is defined as the effect of the new evidence on the odds of A. The odds of A are Prob(A) ÷ Prob(not-A), so that the effect on the odds, or the likelihood ratio, is Prob(B|A) ÷ Prob(B|not-A).[170]

Despite its internal mathematical soundness, and despite the many insights it generates for law, most observers from many disciplines have voiced serious

---

166. *See* CLERMONT, *supra* note 23, at 120–21 (laying out the basics of Bayes' theorem).

167. *See* SHARON BERTSCH MCGRAYNE, THE THEORY THAT WOULD NOT DIE 8 (2011).

168. *See* JOHN ALLEN PAULOS, INNUMERACY: MATHEMATICAL ILLITERACY AND ITS CONSEQUENCES 63–67 (1988) (using examples to show the transition from conditional probability to Bayes' theorem).

169. *See* Richard O. Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021, 1023 n.12 (1977) (deriving the theorem).

170. *See id.* at 1022–25 (deriving the likelihood ratio).

doubts about whether Bayes' theorem should be seen to play a broad role in legal factfinding.[171] Like any probability-based equation dealing with legal evidence, the problems of Bayes' theorem are numerous. First, and most fundamental, is that the theorem leaves no place for epistemic uncertainties, thus painting the world as being black and white even though most of the world appears in shades of gray. It does not handle well the situation of incomplete, inconclusive, ambiguous, or dissonant information, and has to use an inadequate fudge factor to account for the state of the evidence.[172] That is, it is built on additive bivalent odds, and so does not work with nonadditive multivalent beliefs. Second, and most wonky, is that the theorem fails to give the factfinder an initial prior probability as a starting point. In the proper state of initial ignorance in a civil case, a popular starting point is to posit that the plaintiff's claim has a 50/50 chance. That starting point unfortunately comports with neither the actual probabilities nor the law's instructions; in pure probabilistic form, it introduces the logical problem of making a feather's weight of evidence sufficient to carry the burden of persuasion over a silent defendant; and it produces inconsistencies when there are more than two hypotheses in play.[173] Third, and most obvious, is that the theorem has no claim to be a realistic representation of how legal factfinders do or could aggregate evidence.

---

171. *See, e.g.*, McGRAYNE, *supra* note 167, at 34 (recounting the centuries of controversy generated by Bayes' theorem); Nancy Pennington & Reid Hastie, *Juror Decision-Making Models: The Generalization Gap*, 89 PSYCHOL. BULL. 246, 262–68 (1981) (discussing its problems); Shafer, *supra* note 49, at 809–16 (stating the same). *Compare* Paul Bergman & Al Moore, *Mistrial by Likelihood Ratio: Bayesian Analysis Meets the F-Word*, 13 CARDOZO L. REV. 589, 590 (1991) (attacking), *with* D.H. Kaye, Commentary, *Credal Probability*, 13 CARDOZO L. REV. 647, 648 (1991) (defending), *and* Lempert, *supra* note 4, at 440–50 (stating the same).

172. *See* Lea Brilmayer & Lewis Kornhauser, *Review: Quantitative Methods and Legal Decisions*, 46 U. CHI. L. REV. 116, 135–48 (1978) (observing, while laying out all the logical problems, that: "To handle these difficulties, other schools of statistical inference customarily supplement probability statements with higher-order probability statements or nonquantifiable qualifications—statements about the size of the sample, for instance—that bear on the weight of the evidence.").

173. *See* State v. Spann, 617 A.2d 247, 254 (N.J. 1993) (saying that ".5 assumed prior probability clearly is neither neutral nor objective"); Kevin M. Clermont, *Trial by Traditional Probability, Relative Plausibility, or Belief Function?*, 66 CASE W. RES. L. REV. 353, 368–69 (2015) (discussing the situation of a lack of proof); Lempert, *supra* note 4, at 462–67 (noting that employing 50/50 as the appropriate odds when ignorant of the true facts can cause many problems). On problems with choosing any different prior probability, see Jaffee, *supra* note 50, at 980–85.

B. *Dempster-Shafer Theory*

The main focus of belief function theorists involves this problem of how to aggregate items of evidence. Their stimulus is the realization that Bayes' theorem is a gross oversimplification for aggregating uncertain evidence, one intentionally built onto traditional probability to shear off most of the uncertainty and thereby make all the evidence easily commensurable.[174] The result of their work is an array of sophisticated mathematical tools for aggregating pieces of variously uncertain evidence to produce a composite degree of belief.

Many of those theorists rely on the prominent Dempster rule to govern the task.[175] That rule is complicated, because it abstractly addresses the problem in very general terms (Bayes' theorem turns out to be a special case of the Dempster rule).[176] I could reprint its central formula, but it would be meaningless without many obscure definitions. Suffice it to say, the Dempster rule aggregates the belief functions from different pieces of evidence into a new belief function by orthogonal sum, a process beyond the ken of the ordinary factfinder as a practical matter.[177] Moreover, the Dempster rule has proved to be theoretically inappropriate for many kinds of evidence.[178]

---

174. *See* SHAFER, *supra* note 49, at 26.

175. *See id.* at 6, 25, 57–67 ("[W]e construct a belief function to represent the new evidence and combine it with our 'prior' belief function—i.e., with the belief function that represents our prior opinions. This method deals symmetrically with the new evidence and the old evidence on which our prior opinions are based: both bodies of evidence are represented by belief functions, and the result of the combination does not depend on which evidence is the old and which is the new."); Barnett, *supra* note 72, at 197, 198–204 (explaining the Dempster rule).

176. For a comparison of Bayesian probability judgments and belief functions, see Glenn Shafer & Amos Tversky, *Languages and Designs for Probability Judgment, in* CLASSIC WORKS OF THE DEMPSTER-SHAFER THEORY OF BELIEF FUNCTIONS, *supra* note 64, at 345, 345.

177. *See* Kari Sentz & Scott Ferson, *Combination of Evidence in Dempster-Shafer Theory* 13–15 (Sandia Nat'l Labs. 2002), https://prod-ng.sandia.gov/techlib-noauth/access-control.cgi/2002/020835.pdf [https://perma.cc/BTT2-D2M8].

178. *See, e.g., id.* at 17 ("Suppose that a patient is seen by two physicians regarding the patient's neurological symptoms. The first doctor believes that the patient has either meningitis with a probability of 0.99 or a brain tumor, with a probability of 0.01. The second physician believes the patient actually suffers from a concussion with a probability of 0.99 but admits the possibility of a brain tumor with a probability of 0.01. Using . . . Dempster's rule, we find that . . . *Bel*(brain tumor) = 1. Clearly, this rule of combination yields a result that implies complete support for a diagnosis that both physicians considered to be very unlikely." (relying on Zadeh, *Reviews of Books*, *supra* note 69, at 82)).

Accordingly, the Dempster rule is quite contested,[179] generating many competitors.[180] In fact, there is no one correct way to aggregate evidence into a new belief function, as each alternative rule makes certain assumptions about the natures of the different pieces of evidence and their uncertainty. In particular, the competing rules differ in how they handle conflicting evidence. Almost all are even more complicated than the Dempster rule.[181]

Professor Shafer himself suggested one of the simplest alternatives.[182] It is called the discount-and-combine rule. It works when the evidence is highly conflicting, as in a legal case. The "discount" step diminishes each belief derived from an item of evidence in accordance with the belief's unreliability.[183] "The obvious way to use discounting with Dempster's rule is to discount belief functions at different rates before combining them—discounting at higher rates those belief functions one particularly distrusts and whose influence one wants to reduce."[184] The "combine" step averages the discounted beliefs. To illustrate, let $Bel^d_i(a)$ be the belief in the element a that is generated from an item of evidence, i, and derived by multiplying $Bel_i(a)$ by its discount factor, d_i, where $0 \leq d_i \leq 1$. Then, the average discounted belief based on n items of evidence is, by the discount-and-combine rule, $Bel^d(a) = [Bel^d_1(a) + Bel^d_2(a) + \ldots + Bel^d_n(a)] \div n$.

---

179. *See, e.g.*, Dubois & Prade, *A Set Theoretic View*, *supra* note 69, at 403 ("If subjective probability theory is acknowledged as being too restrictive to model uncertainty judgments, then Shafer's subjectivist interpretation of upper and lower probabilities can be questioned on the same grounds. From a mathematical point of view, [Shafer's] theory of evidence is nothing but the rules of probability theory applied to imprecise statements, while classical probability theory leaves no room to imprecision. As a consequence the rules of combination of bodies of evidence are given by the rules of probability theory, and what is behind the problem of validating Shafer's theory as a theory of measurement of subjective uncertainty is the validity of the rules of (subjective) probability theory (and especially the rule of additivity).").

180. *See* Sentz & Ferson, *supra* note 177, at 17–27 (describing thirteen alternatives, including possibility theory). Exciting new work by mathematicians in the Netherlands has added another complicated alternative that is a belief-function analogue of Bayes' theorem. *See* Kerkvliet & Meester, *supra* note 30, at 145–51.

181. *See* Sentz & Ferson, *supra* note 177, at 8–13 (generalizing).

182. *See* SHAFER, *supra* note 49, at 251–55 (laying out this alternative); Sentz & Ferson, *supra* note 177, at 17–18 (stating the same).

183. *See* SHAFER, *supra* note 49, at 252.

184. *Id.* at 252–53.

A similar alternative for conflicting evidence would be weighted arithmetic averaging.[185] It is the only other alternative that has any claim to feasibility. The "weighting" step credits each belief from an item of evidence in accordance with its significance. The "arithmetic averaging" step divides the sum of the weighted beliefs by the total of the weights assigned, which removes the distorting effect of the weighting. Let $Bel^w_i(a)$ be the belief in the element a that is generated from an item of evidence, i, and derived by multiplying $Bel_i(a)$ by its weight factor, w_i, where $0 \leq w_i$. Then, the average belief based on n items of weighted evidence can be meaningfully expressed either as $Bel^w(a) = [Bel^w_1(a) + Bel^w_2(a) + \ldots + Bel^w_n(a)] \div [w_1 + w_2 + \ldots + w_n]$ or as $Bel^w(a) = [w_1 \cdot Bel_1(a) + w_2 \cdot Bel_2(a) + \ldots + w_n \cdot Bel_n(a)] \div [w_1 + w_2 + \ldots + w_n]$.

In this weighted arithmetic averaging, what is meant by "significance?" It is not the probative force of the item of evidence, because the inferential reasoning process has already accounted for that factor. It is instead something revealed by the belief function itself. If a belief is much stronger or weaker than its corresponding disbelief, then the item of evidence is especially clear. That is, if the ratio of belief to nonbelief is either big or small, it should get a higher w.[186] Obviously, according a w will constitute another stab at judgment in the course of the reasoning process, but high accuracy in w is not critical, as the effect of the weights is moderated by the formula's denominator.

One insight generated by weighted arithmetic averaging is that a second, independent piece of evidence by itself will have no effect on belief if it generates the same level of belief as the first piece of evidence. That insight about redundant, duplicative, or cumulative evidence might trouble some readers as being nonintuitive. But note that the concern here is not interdependent *corroborating* evidence, which heightens credibility and will come in as an ancillary consideration on some chain of inferences. The concern is *converging* evidence, that is, separate items of evidence that support one

---

185. *See* Scott Ferson & Vladik Kreinovich, Representation, Elicitation, and Aggregation of Uncertainty in Risk Analysis—From Traditional Probabilistic Techniques to More General, More Realistic Approaches: A Survey 75–77 (Nov. 1, 2001) (unpublished manuscript), https://core.ac.uk/download/pdf/46729549.pdf [https://perma.cc/Z3BA-7B7J] (laying out this alternative); Sentz & Ferson, *supra* note 177, at 27 (mentioning this as one of the thirteen alternatives). This alternative does not rely on additivity, unlike the Dempster rule itself.

186. *Cf.* Ferson & Kreinovich, *supra* note 185, at 75 ("[W]ider intervals correspond to worse measurements, with larger systematic error. In this case, it makes sense to assign smaller weights to these bad measurements, thus decreasing their impact on the aggregation result."); 1 MCCORMICK ON EVIDENCE, *supra* note 14, at 996–98 (discussing "likelihood ratio").

element.[187] If one item leads to a 0.30 belief in P_1 and another item also induces a 0.30 belief, do we believe P_1 more than 0.30? No. Even under Bayes' theorem, the posterior probability remains the same after introduction of additional evidence of equal probability.[188] Still, such converging evidence is relevant, because it will result in counting the repetitive 0.30 belief more heavily when averaged with all the other evidence.

One worry generated by weighted arithmetic averaging is that combining small beliefs generated by weak evidence with powerful beliefs will lower the overall degree of belief, even if the combination involves weighting. Recall, however, that the relative sizes of belief and disbelief matter, and weak evidence will bring down both. Weighted arithmetic averaging, in this regard of mixing weak and strong evidence, seems clearly superior to probabilistic combination, whereby tangential evidence is converted into betting odds and hence rendered indistinguishable from strong evidence for the purpose of Bayes' theorem.

C. Correct Approach

The various pieces of evidence bearing on an element each produce a degree of belief and of disbelief. We should aggregate the pieces' probative force in a manner consistent with belief function theory. Weighted arithmetic averaging serves that function, while being able to handle conflicting evidence and being comprehensible enough to employ. So, logic says to take the belief function produced by each piece of evidence, and then aggregate all the beliefs by weighting their significance and by averaging. Do the same for disbeliefs, and we have created a new composite belief function for P_1.

Happily, this averaging method fits human capabilities and inclinations. Factfinders might intuitively use it already. Interestingly, it very much resembles the early psychology theory of how factfinders actually find facts: information integration theory.[189] Experiments showed that mock jurors'

---

187. On converging and corroborating evidence, see ANDERSON, SCHUM & TWINING, *supra* note 3, at 106–07; COHEN, *supra* note 30, at 94–95, 280–81.

188. For a sophisticated Bayesian discussion of cumulative evidence, see Lempert, *supra* note 169, at 1041–52.

189. *See* NORMAN H. ANDERSON, FOUNDATIONS OF INFORMATION INTEGRATION THEORY (1981) (providing a conceptual introduction to a theory that scales stimuli by evaluation, weights them by importance, and algebraically combines them to form an overall judgment); Norman H. Anderson, *Cognitive Algebra: Integration Theory Applied to Social Attribution*, 7 ADVANCES IN EXPERIMENTAL SOC. PSYCHOL. 1, 2 (1974) (providing experimental support). Anderson updated his theory and

output conformed with this approach: The human decisionmaker making a finding of fact would begin with an initial impression, or predisposition, and then would process additional units of information; each of these, including the predisposition, would receive a scale value as to evidential strength, S, which was seemingly a measure of the probability of the fact's existence if the informational unit were true; each would also receive a weighting factor, W, which was seemingly a measure of evidential importance that somehow took into account its credibility and tellingness; and the decisionmaker would then combine these into a weighted arithmetic mean that measured the fact's overall likelihood.[190] Now, the mock jurors may have been viewing evidence more holistically, and they likely were not doing this calculation consciously, but the experiments indicate that their output was consistent with having proceeded by weighted arithmetic averaging. The following thus represents the judged likelihood based on k units of information:

$$J = \frac{\sum W_k S_k}{\sum W_k}$$

This formula's differences from belief function theory's weighted arithmetic averaging are that belief function theory treats multivalent beliefs rather than probabilities, moves the predisposition into the proof process as generalizations and ancillary considerations, redefines S as the degree of belief or disbelief, and provides a better definition for W in terms of the belief-to-disbelief ratio.

What the information integration and belief function theories suggest is that it is both natural and correct for factfinders to look at the probative forces of all evidence bearing on an element and then average them: Each item of evidence would have produced a degree of belief and a degree of disbelief, that is, a belief function regarding the element; but obviously, the factfinders would not weight all the items of evidence the same, that is, the factfinders would weight each item according to its evidential significance before averaging; and so the

---

experimentally extended it to the legal context in NORMAN HENRY ANDERSON, MORAL SCIENCE 85 (2017), https://psychology.ucsd.edu/_files/norman-anderson-book/Chapter%204.pdf [https://perma.cc/GBX7-H2NM] ("Substantial support for averaging theory has been found in experiments on legal judgment.").

190. *See, e.g.*, Martin F. Kaplan, *Cognitive Processes in the Individual Juror, in* THE PSYCHOLOGY OF THE COURTROOM 197, 198 (Norbert L. Kerr & Robert M. Bray eds., 1982) ("This process of information evaluation, weighing, and integration is central to understanding the juror's cognition . . . ."); Martin F. Kaplan & Gwen DeArment Kemmerick, *Juror Judgment as Information Integration: Combining Evidential and Nonevidential Information*, 30 J. PERSONALITY & SOC. PSYCHOL. 493, 497 (1974) ("[T]he information provided to a juror . . . possesses both scale value and weight . . . .").

factfinders would roughly calculate a weighted arithmetic mean to create a composite belief function for the element.

Therefore, because weighted arithmetic averaging is a mathematically acceptable approach to the problem of aggregating pieces of evidence and because it seems suited to the uncertain beliefs and conflicting evidence encountered in legal cases, I choose it over its more complex competitors. I thus argue that weighted arithmetic averaging is the logical way to proceed in this second step of evidence processing, and that it is also feasible for human factfinders and may encapsulate what they actually do when aggregating all evidence bearing on an element of the case.

The formulae discussed above help model the logic involved in aggregating pieces of evidence, so theorists can better understand what should (and maybe intuitively does) go on. But of course, I am not suggesting that factfinders should be instructed in those terms.

IV. COMBINING ELEMENTS

A legal case will usually involve more than one element (a finding necessary for a claim or defense to succeed under the substantive law), as shown in Figure 7. In the Sacco & Vanzetti example, we focused on the element of the accused's identity (P_1).[191] But the prosecution also had to prove beyond a reasonable doubt the other elements: the victim's death (P_2) by unlawful force (P_3) with malice aforethought (P_4).
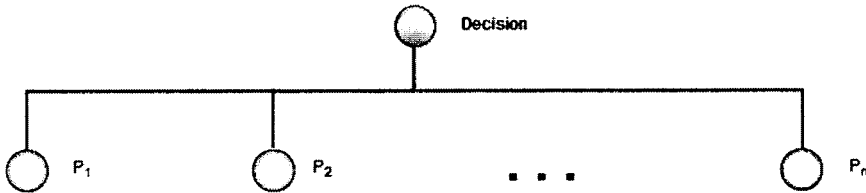

Figure 7: Representation of Combining Elements

This step of combining elements would proceed by the conjunction/disjunction methods of Part II rather than by the weighted-arithmetic-averaging method of Part III, because the elements are all necessary. If the elements were to be combined, they would be joined by the MIN and

---

191. *See* FRANKFURTER, *supra* note 111, at 9.

MAX rules.[192]   But does it matter if the factfinder performs any such combination before rather than after the application of the standard of proof?

A. Atomistic Processing

Once the factfinder has developed a belief function for each element, the factfinder could move from the processing phase to the evaluation phase. There the factfinder would apply the standard of proof to reach a decision on the element. And there, this Article on evidential processing meets up with all my prior work on evaluation.[193]

To apply the standard of proof, the factfinder compares belief to disbelief. For example, a civil case's preponderance standard asks the natural question that the law seems to pose by "more likely than not?": Is the fact shown to be more true than false, that is, do you believe the burdened party's allegation more than you disbelieve it?[194] This believed-more-than-disbelieved standard calls for constructing separate beliefs for a and *not-a* while leaving some belief uncommitted, and then comparing the sizes of the beliefs in a's truth and falsity while ignoring the uncommitted belief. A preponderance of the evidence therefore means that $Bel(a) > Bel(not\text{-}a)$, not that $Bel(a) > 0.50$. If $Bel(a) > Bel(not\text{-}a)$, the factfinder would say it believes a. Indeed, finding an element to exist will follow from a smallish belief's being found to exceed perceptibly an even smaller belief in its contradiction.[195] To continue with my running example, the factfinder should find a if $Bel(a) = 0.40$, when $Bel(not\text{-}a)$ appears as 0.20 and the uncommitted belief equals 0.40.

---

192. *See* Clermont, *supra* note 22, at 1485–95 (justifying use of MIN and MAX rules).

193. *See id.* at 1464 n.23, 1505 n.161 (citing my prior work on the evaluating phase).

194. Higher standards of proof would demand a greater predominance of belief in relation to disbelief. *See id.* at 1481–84 (discussing "clear and convincing evidence" and "beyond a reasonable doubt").

195. If the plaintiff has carried the burden of production and if the plaintiff's proof is perceptibly stronger than the defendant's after taking into account any failure to produce evidence, *see supra* note 67, decision must go for the plaintiff. The court cannot choose not to decide, and a decision for the plaintiff is less likely an error than decision for the defendant would be. *See* Larry Laudan, *Strange Bedfellows: Inference to the Best Explanation and the Criminal Standard of Proof*, 11 INT'L J. EVIDENCE & PROOF 292 (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.

Id. at 304–05.

The law finally intrudes here, specifying a course of decision rather than leaving factfinding to the factfinder's common sense. This is a pattern civil jury instruction:

> 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.[196]

That is, the law tells the factfinder to proceed element-by-element. Simply put, apply the standard of proof to each element; do not apply it to the whole claim or defense. If each element passes the standard of proof, then the whole case passes. So, if the factfinder believes P_1, believes P_2, believes P_3, and believes P_4, then the law believes the conjunction of those elements.

According to the law, then, the factfinder need not combine the elements at all. It compares belief and disbelief for each element. If belief prevails on each element, the burdened party wins. The law appears so perfectly consistent with multivalent logic that it must have been built on that logic.

B. Holistic Processing

By contrast, holistic theorists contend that the human factfinder ignores the law's instructions and evaluates the case as a whole.[197] If these theorists are correct, and they very well may be, then the factfinder has to combine the elements to get a sense of the whole case. To the extent the holistic factfinder proceeds with any accuracy, it would have to conjoin the beliefs and disjoin the disbeliefs in all the elements before applying the standard of proof.

How should the holistic factfinder conjoin and disjoin beliefs? As explained in Part II, the mathematics for combining beliefs instructs that the conjunction has a degree of belief equal to the weakest of the conjoined beliefs.[198] The mathematics also instructs "that a disjunction has a degree of belief equal to the strongest of the disjoined beliefs."[199] If the belief in each

---

196. 3 O'MALLEY, GRENIG & LEE, *supra* note 21, § 104:01; *see* Ronald J. Allen & Sarah A. Jehl, *Burdens of Persuasion in Civil Cases: Algorithms v. Explanations*, 2004 MICH. ST. L. REV. 893, 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)).

197. *See supra* note 22 (discussing the story model).

198. *See supra* note 141 and accompanying text.

199. *See* Clermont, *supra* note 22, at 1486.

element is stronger than its corresponding disbelief, the conjunction of all the elements' beliefs is normally stronger than the disjunction of all the disbeliefs. If each element is more likely than not, then the elements' conjunction normally should be more likely than not. Or, more generally, if each element passes the standard of proof, then the elements' conjunction should pass the standard of proof.

So, it makes no difference whether the factfinder proceeds atomistically or holistically. The two approaches should work out to the same outcome.[200]

C. Correct Approach

I thus argue that, in the main, element-by-element application of the standard of proof is a logical and feasible way to proceed. The factfinder need never combine the elements. The result is that there will be no step number 3. The factfinder can proceed directly from step number 2's belief-function-for-each-element to the standard-of-proof phase.

Although it would make little difference to the logical outcome if a human factfinder were to apply the standard of proof to the whole claim or defense only after conjoining the elements, the more systematic element-by-element approach has practical advantages. Telling factfinders to proceed element-by-element should make their path to decision more careful and diligent.[201] It would also make instructing them on how to proceed simpler and more comprehensible. Finally, the element-by-element instruction avoids the need to instruct on using the MIN rule and against using the product rule. Therefore, the judges' practice of stressing the element-by-element instruction is an excellent practice.

There is, however, at least one hiccup in adopting an element-by-element instruction. The MIN/MAX analysis, which is a little more complicated than

---

200. *See id.* at 1485, 1493, 1504–05 (elaborating this conclusion).

201. Particular care, or institutional change, is necessary where a legal rule requires an affirmative decision before proceeding to another legal rule. *See* Clermont, *Rules, Standards, and Such, supra* note 159, at 797–99 (describing this kind of decision). An example would be determining intellectual disability before the issue of the death penalty. *See* John H. Blume, Sheri Lynn Johnson, Paul Marcus & Emily Paavola, *A Tale of Two (and Possibly Three)* Atkins: *Intellectual Disability and Capital Punishment Twelve Years After the Supreme Court's Creation of a Categorical Bar*, 23 WM. & MARY BILL RTS. J. 393, 409–12 (2014) (finding that juries are more apt to find no-disability than judges, who are trained to go issue-by-issue). One explanation might be that juries, following the story model, answer the overall question of death worthiness, rather than focusing on intellectual disability alone. *See also* Jeffrey J. Rachlinski, Chris Guthrie & Andrew J. Wistrich, *Probable Cause, Probability, and Hindsight*, 8 J. EMPIRICAL LEGAL STUD. (SPECIAL ISSUE) 72, 72 (2011) (providing another illustration).

described so far,[202] produces a difficulty when a fairly strong disbelief, which is nonetheless insufficient under the standard of proof to overcome the belief in the same element, is bigger than a sufficient belief on another element.[203] For example, in a hypothesized civil case, if *Bel(a)* = 0.40 and *Bel(not-a)* = 0.20, and if *Bel(b)* = 0.20 and *Bel(not-b)* = 0.00, then *Bel(a* AND *b)* = 0.20 and *Bel(not-a* OR *not-b)* = 0.20. Thus, the element-by-element approach would produce a result (proponent wins) different from the strictly logical approach (proponent loses). The law rejects this logical outcome by insisting on element-by-element application of standard of proof. Still, there are a number of reasons to think the law's approach might still be optimal.

First, the logical outcome is arguably overcautious. The instinct of most observers is that the proponent should win after convincingly prevailing on each element. The proponent would prevail under a probability theory.[204] Given humans' naturally rough handling of the conjunction problem,[205] the proponent might win in real life, whether the factfinder took an atomistic or holistic approach to beliefs or to probabilities. Maybe the law should be wary of going against such strong intuition.

Second, the law would not want to, and does not, charge its factfinders to perform the difficult mental task of comparing conjunction and disjunction of elements. Combining a series of elements is difficult even to verbalize. Add to the difficulty that comparison of a series of elements might involve comparing a belief in one element to the disbelief of a different element, and

---

202. *See* Clermont, *supra* note 173, at 385–89 (discussing the MIN and MAX rules for belief functions). I have been using multivalent logic somewhat less for its mathematics than for its imagery that so nicely captures legal factfinding. For a defense of using a broad view of multivalent logic as imagery, see Liu & Yager, *supra* note 64, at 2–12 (formalizing an image of belief functions); Vilém Novák, *Modeling with Words,* SCHOLARPEDIA (2008), http://www.scholarpedia.org/article/Modeling_with_words [http://perma.cc/2JWT-YLZ8] ("Mathematical fuzzy logic has two branches: fuzzy logic *in narrow sense* (FLn) and fuzzy logic *in broader sense* (FLb). FLn is a formal fuzzy logic which is a special many-valued logic generalizing classical mathematical logic . . . . FLb is an extension of FLn which aims at developing a *formal theory of human reasoning.*").

203. The complication does not become a serious concern in connection with the chains of inferences of Part II. There, the factfinder is trying to fix the probative force of a piece of evidence, without instructions to evaluate inference-by-inference; anyway, the force of any one piece of evidence will usually be diluted by other evidence. Here in Part IV, the effect of one fairly large disbelief could be determinative under the standard of proof.

204. Normalizing the beliefs yields Prob(a) = 67% and Prob(b) = 100%. *See supra* note 53 and accompanying text (discussing pignistic transforms). Applying the product rule then yields a 67% probability of the conjunction.

205. *See supra* text accompanying note 153 (discussing the empirical evidence).

you have a process that is apt to stymie most factfinders. The law's element-by-element method is more comprehensible (and corralling) than any combination-of-elements method, and it works out to be largely equivalent.

Third, the exceptional situation of the hypothesized case would not be common. Under a set of coherent beliefs and disbeliefs, if *b* is less likely than *a*, then *not-a* should normally be less likely than *not-b*. Trying to impose strict logic on factfinders to pick up the rare exception might not be worth the candle. Thus, for conjoining beliefs in elements, the law apparently, and wisely, pursues workability with a simplifying assumption of well-behaved belief functions.[206]

Fourth, I note that the step of evaluating elements coincides with a switch in commonly employed legal images: Law and lawyers tend to shift unconsciously from a style of thinking that conforms with belief functions (used by factfinders when processing evidence, with acute awareness of uncommitted belief) to a form of speech that conforms with fuzzy logic (used by lawyers, when discussing the output of evidence processing as "more probable than not" or whatever, having set aside the uncommitted belief).[207] Fortunately, belief functions and fuzzy logic are compatible, being alternative versions of multivalent logic and each applying the MIN and MAX rules.[208] Belief functions differ from fuzzy logic in their treatment of uncertainty. Belief functions treat epistemic uncertainty front and center in terms of uncommitted belief, while fuzzy set theory moves any second-order imprecision, that is, uncertainty about the estimate of degree of membership, into the additional dimension of a so-called ultrafuzzy set. It is fuzzy logic's separation of first-order uncertainty from second-order imprecision that makes the operation of its MIN and MAX rules simpler in appearance. This switch in logical frameworks is not necessary, but the switch both conforms to lawyers' imagery and makes

206. *See* Rott, *supra* note 64, at 310–11 (assuming contraposition when conjoining belief functions, so that if *b* is less likely than *a*, then *not-a* should be less likely than *not-b*); *cf.* COHEN, *supra* note 30, at 114, 221, 256, 267 (assuming contraposition for his inductive probability, which leads to the conclusion that "the plaintiff proves his over-all case on the balance of probability if, and only if, he thus proves each of his component points").

207. *See* HAACK, *supra* note 50, at 57 (illustrating the legal system's usages).

208. *See supra* note 69 (supporting the compatibility of the logical versions). For a mathematical defense of a complete switch from belief functions to fuzzy logic or possibility theory and their straightforward MIN and MAX rules, and for the simplifying assumptions on which the switch rests, see Didier Dubois & Henri Prade, *Consonant Approximations of Belief Functions*, 4 INT'L J. APPROXIMATE REASONING 419, 419, 421 (1990) ("Viewing a fuzzy set as a consonant random set, it is shown how to construct fuzzy sets that may act as approximations of belief functions.").

it easier to picture and discuss what is going on.[209] So, by normalizing the degrees of belief at the end of evidential processing, we can speak of fuzzy beliefs being the likelihood that a fact is true and a complementary measure of being unproven, even though we remain in a nonadditive system. In terms of those fuzzy beliefs, we at least can feel more comfortable with the notion that if the belief in each element is stronger than its corresponding disbelief, the conjunction of all the elements' beliefs is stronger than the disjunction of all the disbeliefs.

V. CONCLUSION

Once one recognizes that degrees of belief capture the output of factfinding, and that traditional probability does not, the way to represent the logic of factfinding becomes fairly obvious. Abandoning probabilistic images—going from bivalent logic to multivalent logic—is mentally and emotionally challenging, however, so the proper reasoning has remained hidden behind theorists' acceptance of a cognitive black box. This Article is the first to provide a complete account of how to reason logically from evidence to a decision on facts.

First, the factfinder should connect each item of evidence to a fact to be proved by constructing a chain of inferences. Each inference rests on a generalization, which ancillary considerations can support or undercut. By multivalent logic's usual rules for conjunction and disjunction, the factfinder's degree of overall belief in the fact to be proved will be as strong as the weakest inference in the chain, while the degree of disbelief will be as strong as the strongest disbelief in the chain of reasoning. A belief function amalgamates the belief and the disbelief, along with uncommitted belief that reflects epistemic uncertainty, to give a final measurement of the probative force of the item of evidence on the element.

Second, the factfinder must aggregate the probative force of all the items of evidence that bear on any one fact to be proved. To do so, the factfinder should compute, albeit intuitively, a weighted arithmetic mean of the belief functions for all the items: Get the belief function for each item, weight beliefs and disbeliefs in accordance with evidential significance, and divide the separate

---

209. *See* MIRCEA REGHIŞ & EUGENE ROVENTA, CLASSICAL AND FUZZY CONCEPTS IN MATHEMATICAL LOGIC AND APPLICATIONS 354 (1998) (referencing belief functions and fuzzy logic, and observing: "In order to treat different aspects of the same problems, we must therefore apply various theories related to the imprecision of knowledge."); SCHUM, *supra* note 69, at 41, 200–01 (disbelieving that it is "possible to capture all of this behavioral richness within the confines of any single formal system of probabilities").

sums of beliefs and disbeliefs by the sum of the weights. The result is a composite belief function for the element.

Third, the factfinder may choose to combine the elements of claim or defense. But the better approach is to apply the appropriate standard of proof to each element. This means that the factfinder would proceed directly to the evaluation phase, which will involve a standard of proof that compares the degree of belief that appears in the element's composite belief function with the degree of disbelief. Evaluation does not require a combining of the elements.

To put it more compactly, the factfinder should construct a chain of inferences to produce a belief function for each item of evidence bearing on an element, and then weight-and-average them to produce for each element a composite belief function ready for applying the element-by-element standard of proof. Mapping this normative method for processing legal evidence has been a worthy undertaking. More significantly, the mapping provides further demonstration of how embedded the multivalent-belief model is in our law. Traditional probability has a tight hold on modern legal minds, although not on the law itself. Accepting multivalent beliefs in lieu of bivalence's probabilism would solve so many theoretical problems, disarm so many criticisms of the law, and just explain so much for us that one must wonder why it meets such resistance.