# Leveraging Surprise: What Standards of Proof Imply that We Want from Jurors, and What We Should Say to Them to Get It

*D. Michael Risinger*[*]

## I. PROLOGOMENON

> The world is stranger than we imagine;
> the world is stranger than we <u>can</u> imagine.[1]
>
> Essentially, all models are wrong,
> but some models are useful.[2]

The relationship between mathematics and the world exterior to ourselves is mysterious even though it often seems obvious. Does

---

[*] John J. Gibbons Professor of Law, Seton Hall University School of Law.

[1] Adapted from J.B.S. HALDANE, POSSIBLE WORLDS AND OTHER PAPERS 286 (1927) ("I have no doubt that in reality the future will be vastly more surprising than anything I can imagine. Now my own suspicion is that the Universe is not only queerer than we suppose, but queerer than we *can* suppose.").

[2] George Box, in GEORGE E.P. BOX & NORMAN DRAPER, EMPIRICAL MODEL-BUILDING AND RESPONSE SURFACES 424 (1987). Box repeated the idea in various forms at various times in his writings, beginning in 1976. *See All Models Are Wrong*, WIKIPEDIA, https://en.wikipedia.org/wiki/All_models_are_wrong (last visited Mar. 18, 2018). This is why the canonical version in Box & Draper is usually, and properly in my view, attributed to Box alone. *See generally* BOX & DRAPER. For an interesting examination of the relationship of facts to models (or perhaps we should say the facticity of models), see Rebecca Haw Allensworth, *Law and the Art of Modelling: Are Models Facts*, 103 GEO. L. J. 825 (2015).

mathematics inhere in the world inhabited by humans such that it forms the actual substrate of external reality (or corresponds to it perfectly or infinitely closely)?  Or is mathematics a modelling tool for representing the world only in part, more or less accurately, but leaving out important aspects of even physical reality beyond the mathematical description deployed?  I believe many who find mathematics both inherently fascinating in its own right, and also who find that they have elite aptitude for following and understanding mathematical relationships—particularly those expressed in the specialized language of formal symbolization—often tend to think the former to be the case, without much reflecting upon the issue.  I suppose, not being one of those, it is only natural that I should regard mathematics, and all formal symbolization, only as tools for modelling aspects of reality.

Because of our various limitations, dimensionally and cognitively, the true nature of the fundamental reality of the world is forever beyond our grasp.  All we can hope for is some well-warranted reason to believe we are improving our grasp, without knowing for sure if so, and by how much.

When is formal probability theory a fit-for-purpose model of important aspects of our world?

First, it is useful to remember the origins of formal probability theory.  Apparently there have always been many people who liked to gamble.  The urge to wager on the unpredictable outcomes of games of chance is ancient.  In Italy and France in the 16th and 17th centuries the wealthy wagered large sums on such games, including dice-throwing games.  Obviously, some knowledge of heuristic probabilities for many such games had been worked out by players[3], but no one had worked out a system for fully analyzing such games even though they were, more or less by the definition of their rules, fully specifiable, and even though such a system would have been of great value to anyone wielding it in actual play.  However, it was not until Pascal and Fermat worked out formal techniques for analyzing dice games that the foundation was laid for a formalized system of probability analysis.  The techniques of analysis Pascal and Fermat worked out for dice were extended over the next centuries to the analysis of all manner of fully-specified formal systems.  The main point here is that, taking dice rolls as the origin, Pascal

---

[3]  Heuristics such as "never draw to an inside straight" in five-card draw poker, for instance.  Such heuristics are apparently the result of "chunking" a particular recurring problem into more tractable sub-problems and arriving at an informal rule of thumb concerning one or more of the chunks.  They are often accurate, but not necessarily so, and it was one such inaccurate heuristic (the then-common wisdom that betting that at least one double six would come up in twenty four rolls of two dice was a good bet) that precipitated part of Antoine Gonbaud's 1654 query to the Mersenne Salon that resulted in the reactions to the problem by Pascal and Fermat, whose correspondence then laid out the foundation for a fully formalized theory of probability. *See Antoine Gombaud*, WIKIPEDIA, https://en.wikipedia.org/wiki/Antoine_Gombaud (last visited Mar. 14, 2018).

did not work out probabilities empirically by rolling real dice and observing the outcomes. Rather, he assumed the conditions of the game were ideal, and worked out the probabilities analytically under this ideal assumption. No real dice and no real rolls of dice required.

This is a fine way of proceeding if one's goal is to arrive at analytic (mathematical or logical) truths in an ideal formal system that is fully specified. And the results will be exceedingly useful in the real world if, and only to the extent that, the ideal conditions correspond to the real-world conditions of physical play. With the dice available in the 17th century this assumption about correspondence was apparently accurate enough that the best way to play was to treat the real dice *as if* they were ideal dice. But that would not, or at least might not, have been the case if the dice available were the knuckle bones of animals that apparently were the first cast or rolled versions of dice. Those knuckle bones might have made certain bets, suggested by the analysis of ideal dice (or knuckle bones) to be winning bets, into losing bets in the real world.[4]

All models have this "as if" relationship to the real world. All are wrong (to some degree) but some are useful. The big question is, how do we judge their usefulness—their sufficient "as if" correspondence to the part of the world for which we are interested in using them.[5]

---

[4]  The story of knucklebones as rolled game pieces is a bit complicated. "Knucklebones" (also known in antiquity as tali) were originally the pastern bones of a sheep or goat. These bones were elongated, with two ends that were rounded and upon which the bones could not stand. There were therefore only four surfaces that a pastern bone could display when cast or rolled, not six as with dice. In addition, the irregularities of pastern bones made it the case that each surface had a different probability of coming up, and with real pastern bones this could of course vary some from bone to bone. Later "knucklebones" reduced but did not solve this problem by being carved from bone, ivory or other material in imitation of real pastern bones, but more uniform in shape and size. In any event, the probabilities of knucklebones throws were (and are) hard to model theoretically, and need empirical input. An attempt to put forth some theorized knowledge to be used by knucklebones players formed part of what is regarded in certain aspects to be one of the earliest proto-treatises on probability. *See* GERALAMO CARDANO, LIBER DE LUDO ALEA (THE BOOK ON GAMES OF CHANCE) § 31, 53–56 (Sydney Henry Gould trans., Dover ed. 2016).

[5]  "It is important . . . to keep in mind that the use of mathematical language is not a goal for its own sake in the natural sciences, but is a highly specialized technique for arriving at a very complete consensus. Any attempt to apply the same intellectual technique to social and behavioral phenomena must have the same purpose, and must therefore maintain comparable criteria of categorical precision and operational fidelity to the realities. By these criteria, unfortunately, most such attempts appear implausible and pretentious, and probably induce obfuscation and mental opacity where intended to clarify and persuade." JOHN ZIMAN, RELIABLE KNOWLEDGE: AN EXPLORATION OF THE GROUNDS FOR BELIEF IN SCIENCE 166 (1977).

Formal probability theory in various garbs[6] has long been put forth as something that should be regarded as a proper model for all or part of the jury trial process. I have long been suspicious of such claims. While formal probability theory does have practical lessons to teach that may help to either explain or to criticize and improve the trial process in various ways,[7] my view has always been that those lessons come by way of analogy, and that absent the availability of hard data, one must be suspicious of models that need to be supplied with cardinal numbers to work.[8] Given the nature of the evidence and inference involved in trials, attempts at non-metaphorical use of formal probability theory is generally not useful, and often affirmatively misleading. I am hardly alone in this view.[9]

As I was struggling to put together a more expanded account of this position, what should pop up on my SSRN evidence journal feed but a new article by Ron Allen which summarizes most of the reasons I was going to set out, probably better than I could.[10] In addition, Kevin Clermont has identified other circumstances which reinforce the conclusion that formal probability theory is not a useful model for the trial process.[11] I will not attempt to summarize all of Professor Clermont's points or Professor Allen's points here, beyond saying that Professor Clermont's emphasis on how little

---

[6]   I am restricting the notion of "formal probability theory" here to various schools of thought about the notion of probability and how best to formalize it that meet the so-called "Kolmogorov axioms," which specify inter alia that a probability is a positive real number between zero and one with the property of additivity. This might also be called "standard probability theory." There are other approaches to probability which do not satisfy the Kolmogorov axioms. *See* Alan Hajek, *Probability, Logic, and Probability Logic*, *in* THE BLACKWELL GUIDE TO PHILOSOPHICAL LOGIC 362–84 (Lou Goble ed., 2001).

[7]   I am quite drawn to David Schum's position in regard to various approaches to modelling uncertainty: "The different formal systems of probabilistic reasoning each resonate to different attributes of the very rich process of probabilistic reasoning. In my analysis of evidence, I have viewed these different formal systems not as normative guides but as heuristics for examining evidence." DAVID A. SCHUM, EVIDENTIAL FOUNDATIONS OF PROBABILISTIC REASONING 284 (1994).

[8]   *See* D. Michael Risinger, *Against Symbolization*, 11 LAW, PROBABILITY & RISK 247 (2012); D. Michael Risinger, *Reservations about Likelihood Ratios (And Some Aspects of Forensic 'Bayesianism')*, 12 LAW, PROBABILITY & RISK 63 (2013).

[9]   Reservations of this kind have been expressed ever since the emergence of versions of Bayesian probability analysis in the 1920s. *See, e.g.*, FRANK H. KNIGHT, RISK, UNCERTAINTY AND PROFIT 1–20, 197–232 (1921). For an excellent review of the problems facing formalized modeling of human inference, see James Franklin*, How Much of Commonsense and Legal Reasoning is Formalizable: A Review of Conceptual Obstacles*, 11 LAW, PROBABILITY & RISK 225 (2012).

[10]   Ronald J. Allen, *The Nature of Juridical Proof: Probability as a Tool in Plausible Reasoning*, 21 INT'L. J. EVIDENCE & PROOF 133 (2017). Not a new theme for Professor Allen. *See, e.g.*, Ronald J. Allen & Michael S. Pardo, *The Problematic Value of Mathematical Models of Evidence*, 36 J. LEGAL STUD. 107 (2007). Nor for me. *See, e.g.*, Risinger, *Against Symbolization*, *supra* note 8.

[11]   *See generally* KEVIN M. CLERMONT, STANDARDS OF DECISION IN LAW (2013).

the law relies on crisp categories of fact will return at the end of this piece. For now, however, I will only address the single characteristic of the trial where probability theory has been most invoked as providing the proper model—standards of proof.

## II. BURDENS OF PERSUASION, STANDARDS OF PROOF, JURY INSTRUCTIONS, AND EXPECTED JURY PERFORMANCE

In their most fundamental aspect, burdens of persuasion (risks of non-persuasion) cannot shift. What I mean by this is that, in regard to the individual case, the burden of persuasion is imposed on one party only, and only at a single point, in the court's instruction to the jury. Prior to that point, as to any issue, one can argue about where the burden *ought* to be placed under applicable legal doctrines, but it is only by examining the actual instruction that one can see where it was in fact placed, and at that point it will virtually always be placed on one and only one party as to any given issue.[12] Most "shifting" talk is actually a shorthand way of saying either (1) that something occurred during the unfolding of the case that put the case into a subcategory of cases where the imposition of the burden in the instructions is different as to an issue from what it would normally be but for that juridical event, or else (2) that the burden was erroneously placed in the instruction under applicable law.

I raise this point here because it goes some way toward explaining the emphasis in this paper on figuring out what we want jurors to do, and then operationalizing it as best we can through what they are told in the instructions (and perhaps what we allow them to be told in argument). Not only is the instruction to the jury the point at which the burden is placed, it is the vehicle for explaining to jurors what we expect of them pursuant to the applicable standard of proof, and for doing so (ideally) in the way most likely to obtain from them the performance we desire of them.[13]

---

[12]  One could theoretically have a "contingent instruction" in which the jury was told to address an initial basic fact, and if their finding as to that was A, then they were to proceed as if the burden of persuasion as to issue X were on Party 1, but if they did not find A, then they were to proceed as if the burden of persuasion as to issue X were on Party 2. This might be seen as working an actual post-instruction shift in the burden from Party 2 to Party 1 upon an affirmative finding of A. But I know of no real-world examples of such a mechanism being deployed.

[13]  Not exactly a new position for me: "The term 'beyond a reasonable doubt' . . . forces the individual juror to ponder the question of how the law wants him to decide, and to examine his own soul, so to speak, to answer the question of whether or not he is as sure as the law requires him to be. It is intended to make even the juror who thinks that the defendant 'did it,' in everyday terms, think twice." Harold A. Ashford & D. Michael Risinger *Presumptions, Assumptions & Due Process in Criminal Cases: A Theoretical Overview*, 79 YALE L. J. 165, 199 (1969).

### III. HOW THE PHENOMENON OF SURPRISE CAN CLARIFY STANDARDS OF PROOF

> Tell me that the great Mr. Gladstone, in his last hours, was haunted by the ghost of Parnell, and I will be agnostic about it. But tell me that Mr. Gladstone when first presented to Queen Victoria, wore his hat in her drawing room, slapped her on the back, and offered her a cigar, and I am not agnostic at all. That is not impossible; it's only incredible. But I'm much more certain it didn't happen than that Parnell's ghost didn't appear.
>
> —G.K. Chesterton[14]

> The most exciting phrase to hear in science, the one that heralds new discoveries, isn't "Eureka!" but rather "Hmm. . .that's funny."
>
> —Isaac Asimov[15]

> I could easily produce 20 likelihoods, but the question is, would you take any of them seriously?
>
> —David Schum[16]

In this part of the paper, I will develop and defend the proposition that when humans evaluate evidence and determine what they believe in regard to facts, the primary, though usually implicit, operator in those determinations is, or at least ought to be, the fundamental emotion of surprise.[17] I will further assert that this relationship between belief and surprise, at least in regard to anything with a defensible claim to the label "fact," is one of entailment, or very nearly so. Finally, I will argue that

---

[14] Speaking through Father Brown in *The Curse of the Golden Cross*, *in* THE INCREDULITY OF FATHER BROWN 122 (1926).

[15] This is the usual attribution, but there is no printed source for it, and it well may be wrong.

[16] SCHUM, *supra* note 7, at 233.

[17] Surprise is a primary category in modern taxonomies of emotion. *See, e.g.*, Paul Ekman, *Facial Expression and Emotion*, 48 AM. PSYCHOL. 384 (1993) (anger, disgust, fear, happiness, sadness and surprise); ROBERT PLUTCHIK, EMOTION: THEORY, RESEARCH, AND EXPERIENCE, VOL. 1 (1980) (joy, sadness, anger, fear, anticipation, surprise, trust and disgust); W. Gerrod Parrott, *Emotions in Social Psychology: Volume Overview*, *in* EMOTIONS IN SOCIAL PSYCHOLOGY: ESSENTIAL READINGS (W. G. Parrott ed., 2001) (love, joy, surprise, anger, sadness, fear). Only anger, fear, sadness, and surprise are on all three lists. Of course, there are names for intensified or diluted forms related to the primary forms. For surprise these might include intensified forms such as amazement and dilute forms such as unexpected or curious (in the sense of "that's curious"). The opposite of surprise is expectation—when you expect something you are not surprised by it. When something is unexpected it induces surprise.

asking questions about potential surprise is a much more fruitful way of determining and norming degrees of human belief in regard to facts than the hypothetical betting experiments favored by some, most notably by neo-Bayesians seeking a way to derive cardinal numbers to incorporate into applications of Bayes's rule.

I began developing this approach a few years ago, and in 2013 I presented a rough precis of the idea in the context of defining standards of proof at a colloquium at the University of Texas at the behest of Larry Laudan. As sometimes happens, the press of other obligations caused the subject to be put on the back burner. But I have now, in the shadow of pending retirement, returned to it with an eye to doing the extended reflection and analysis that might make such an apparently radical proposal seem at least tenable. During this process, I discovered, almost by accident, that a somewhat similar approach to the role of surprise had been put forth over a half-century ago by a prominent English economist, G. L. S. Shackle.[18] I say

---

[18]    Shackle was a respected although not dominant voice in economics through the 1970s. He even has a Wikipedia entry, and has enjoyed something of a renaissance among behavioral economists in since the late 1990s. In THE BLACK SWAN, Nassim Nicholas Taleb says: "Hayek is one of the rare celebrated members of his 'profession' (along with J. M. Keynes and G.L.S. Shackle) to focus on true uncertainty, on the limitations of knowledge, on the unread books in Eco's library . . . . Tragically, before the proliferation of empirically blind idiot savants, interesting work had been begun by true thinkers, the likes of J. M. Keynes, Friedrich Hayek, and the great Benoit Mandelbrot, all of whom were displaced because they moved economics away from the precision of second-rate physics. Very sad. One great underestimated thinker is G.L.S. Shackle, now almost completely obscure, who introduced the notion of 'unknowledge,' that is, the unread books in Umberto Eco's library. It is unusual to see Shackle's work mentioned at all, and I had to buy his books from secondhand dealers in London." NASSIM NICHOLAS TALEB, THE BLACK SWAN 185 (2007). I feel the need to make some observations concerning Nassim Nicholas Talem and his ideas, mainly as put forth in THE BLACK SWAN. I did not read THE BLACK SWAN until the manuscript for this piece was nearly finished (perhaps put off by an off-putting review in the New York Times Book Review). I did a general web search for references to G.L.S. Shackle, and discovered a reference to Shackle in THE BLACK SWAN, together with the text quoted in this footnote. I then felt constrained to read the book, since I was going to quote from it.

Those who have read THE BLACK SWAN will perhaps see that I found many of the substantive positions taken by Taleb to be congenial to my own thought. However, they must be extracted from one of the most idiosyncratic expositions I have ever encountered (the subject of the off-putting part of the Times review). His substantive analysis is embedded in semi-autobiographical essays that are, by turns, discursive, defensive, arrogant, insecure, dismissive, mean-spirited and just plain nasty (his recurrent fantasies of putting a rat down the shirt of such as he considers pompous know-nothings—usually prominent academics or CEOs—illustrate the latter point). But embedded in this screed (which is often entertaining when Taleb's prejudices match one's own) are points of both brilliance and great importance. Take for example, his point that many economists use Gaussian bell curve distributions as good models of economic phenomena when they clearly do not fit. He does a good job of showing why, and there is plenty more where that came from.

So do not let the rhetorical and emotional excesses of this aging *enfant terrible* put you off. Daniel Kahnemann, one of Taleb's heros (and mine), when proposing Taleb's name for inclusion on a list of the world's leading intellectuals in 2008, said that Taleb is "original and

almost by accident because, even though he was a significant voice in economics from the 1940s through the early 1980s, Shackle's work has had almost no penetration into the legal literature in general and into discussion of inference in particular.[19]  I discovered a reference to Shackle's theory of surprise in the last couple of years while undertaking a close reading of David Schum's monumental EVIDENTIAL FOUNDATIONS OF PROBABILISTIC REASONING.[20]  In this work, Schum embraces a big tent approach to the potential informative value of various approaches to probabilistic reasoning from evidence,[21] and on page 269, in explaining which approaches he would make central to his exposition (that is, Bayes, L. Jonathan Cohen, Glen Shafer and Lofti Zadah), he writes: "Before I begin my task, I must mention that my coverage of alternative systems of probabilistic reasoning is not exhaustive.  There are other systems I might have mentioned that also have useful properties.  Examples include the *evidentiary value model* and the theory of *potential surprise*."

That sounded right on point.  And so I obtained and read Shackle's DECISION, ORDER, AND TIME IN HUMAN AFFAIRS.[22]

As it turned out, Shackle's approach to surprise was somewhat limited by the context in which he deployed it.  As an economist who was mainly concerned with corporate economic planning decisions, he developed his approach around notions of potential surprise at time one, the point of a decision to take an action.  The object of that potential surprise was always

---

audacious," and has "changed the way many people think about uncertainty."  High praise indeed, and considering the source, warrants the effort to extract the important points (which are pretty clearly expressed when Taleb gets down to them) from the distractions that surround them.

[19]    Although Shackle's ideas on the centrality of surprise were the subject of significant discussion in the economics literature of the 1950s through the early 1970s, they have been mentioned only a handful of times in the periodical legal literature, and then generally only in passing.  *See* Fienberg, Stephen E., *Misunderstanding, Beyond a Reasonable Doubt*, 66 B.U. L. REV 651, 654 (1986); Adam J. Hirsch, *Cognitive Jurisprudence*, 76 S. CAL. L. REV. 1331, 1335 (2003); Tracy Lightcap, *Crucial and Routine Decisions: A New Explanation of Why Ideology Affects U.S. Supreme Court Decision Making the Way It Does*, 84 TULANE L. REV. 1491, 1498 (2010); and Alex Stein, *The Flawed Probabilistic Foundation of Law and Economics*, 105 N.W. U. L. REV. 199, 224 (2011).  (A search of the Heinonline law journal library data base using the terms "G. L. S. Shackle," "George L.S. Shackle," "George Shackle" and "G. Shackle" turns up 83 citations to Shackle or Shackle's work, beginning in 1960.  Only the listed sources mention his theory of surprise at all.  The first two mention it in a single line, and the third in a paragraph.  Only Professor Stein goes further.)  On the monograph side, Professor Ho devotes eight pages to a discussion of Shackle in connection with Professor Ho's own theory of categorical belief and burdens of persuasion.  *See* HO HOCK LAI, A PHILOSOPHY OF EVIDENCE LAW: JUSTICE IN SEARCH OF TRUTH 143–51 (2008). I thank Dale Nance for drawing my attention to this.

[20]    *See* SCHUM, *supra* note 7.

[21]    *See* quotation from SCHUM, *supra* note 7.

[22]    GEORGE L.S. SHACKLE, DECISION, ORDER, AND TIME IN HUMAN AFFAIRS (2d ed. 1969).

a future state of affairs that the decision maker would in fact encounter at some time in the future (assuming he lived that long).[23] Because of this, his reasoning does not always map on comfortably to a situation where the decision involved will not in itself lead to a future state when the object of the surprise estimate will be revealed. The latter is the normal case in regard to decisions by juries or judges made pursuant to legal standards of proof. Barring cognizable new evidence, the decision-maker will never know more about the "true state of the world" and how it mapped on, or maps on, or will map on, to his decision, than is known at the point where the decision is rendered.[24] Without going too deeply into Shackle's position, suffice it to say that I regard surprise as having applications beyond the context in which he developed his analysis, and do not find other parts of his decision theory particularly helpful.[25] But the general attraction to surprise remains, especially given the main alternative on offer, to which I will now turn.

It should be obvious by now that I am of the school that thinks that attempts at formalizing the meaning of burdens through the use of numerical expressions of probability is probably a bad idea in theory and definitely a bad idea in practice. Formal mathematization can have all sorts of unintended consequences. Take, for instance, the various attempts to leverage the so-called "Blackstone ratio" into a formal expression of the numerically expressed probability lying behind the notion of proof beyond a reasonable doubt. As Professor Allen and Professor Laudan have shown to a fare-thee-well, the ratio can be as well realized by raising the number of the guilty defendants acquitted as by lowering the number of innocents convicted.[26] And beyond that, the ratio in its Blackstone form can apparently easily be taken to imply the acceptability of a nearly ten percent rate of factually false convictions.[27] Put that way, I suspect it would seem wrong to a majority of people. The trip from a ratio expression to a percentage to a probability expression to a normative expression about system performance

---

[23] Shackle did provide for a revision of surprise if intervening information became available before the predicted future event resolved itself into a specific reality.

[24] I have used all the tenses in this sentence because legal "factfinders" render decisions about past states of the world in regard to some elements (actus reus in criminal cases, for instance), present states of the world in regard to others (the identity of a substance as contraband, for instance), and future states in regard to yet others (in regard to many remedial issues in tort, for instance).

[25] I thank Ron Allen for his reference to Marcello Basili & Carlo Zappia, *Ambiguity and Uncertainty in Ellsberg and Shackle*, 34 CAMBRIDGE J. ECON. 449 (2010); *see also* Carlo Zappia, *Non-Bayesian Decision Theory Ante Litteram: The Case of G.L.S. Shackle* (DEPFID Working Paper No. 4, 2008), http://repec.deps.unisi.it/depfid/text408.pdf.

[26] Ronald J. Allen & Larry Laudan, *Deadly Dilemmas*, 41 TEX. TECH. L. REV. 65, 75–79 (2008).

[27] I say "apparently" because I have often seen this done, even though Blackstone's ratio on its face deals with a ratio of the convicted innocent to the acquitted guilty, not the convicted innocent to the convicted guilty.

is a long journey over a path fraught with pitfalls.[28]

Nor can any other approach to mathematization currently supply credible numbers to mathematize what is meant by standards of proof. The main such approach often proposed seems to be subjective expected utility decision theory or some other aspect of modern Bayesianism, but in fact I believe such approaches are not fit for this purpose, for a number of reasons. To explain my position best, I will start with a short quotation from a previous writing[29]:

> There is no doubt that the term "Bayesianism" now covers schools of thought that would have surprised the Reverend Bayes himself,[30] and which embody a set of doctrines at least some of

---

[28]    As I have previously written: "As for the 'Blackstone ratio,' that is the label now generally given to Blackstone's version of the moral assertion 'it is better that ____ guilty go free than that one innocent be convicted.' I believe that all we can say about the intendment of this expression was that it was meant as a general declaration that, for any given crime, an error that convicts an innocent person is much worse morally than an error that acquits a guilty person. The number Blackstone chose to make this point was ten, although Alexander Volokh has rather amusingly shown that the notion is ancient, and that various thinkers over the centuries have put the number in the blank at various values between one and a thousand. The trope first appeared in English in Sir John Fortescue's *De Laudibus Legum Angliae*, composed in the late 1460s, where the number employed was 20. Sir Matthew Hale used the formula in the 1670s, but his number was only 5. What is clear is that when Blackstone used the ratio image, like Hale and Fortescue before him, he was not speaking as a mathematician or probability theorist or formal logician, but as what he was—a practical lawyer/judge commenting on the subjective certainty that should be required of juries making findings of guilt in individual criminal trials. It was a way to justify the principle of calling on jurors to apply great caution in convicting individual defendants in particular cases when the specific evidence of their guilt was relatively weak. It seems unlikely that the idea of actually determining the number of true acquittals, false acquittals, true convictions and false convictions produced by the criminal justice system as a whole, and actually comparing their ratios in any combination as a measure of system performance, would have struck him as possible or useful, had he ever formulated the question of system performance in that way, which so far as I know he never did.

So the Blackstone Ratio as analyzed by Allen and Laudan is in a central sense not really the Blackstone Ratio. Nevertheless, for those who have been tempted to convert the Blackstone Ratio to a canonical measure of system performance, Allen and Laudan establish how unconstraining the Blackstone Ratio used as a measure of system performance really is, since, taken as a mathematical expression, its conditions can formally be met by holding the number of convicted innocents steady and raising the number of acquittals of the guilty." D. Michael Risinger, *Tragic Consequences of Deadly Dilemmas: A Response to Allen and Laudan*, 40 SETON HALL L. REV. 991, 1002–03 (2010) (footnotes omitted. The main source is Alexander Volakh, *Guilty Men*, 146 U. PA. L. REV. 173 (1997)).

[29]    Risinger, *Reservations*, supra note 8, at 64–65 (original footnotes renumbered consequetively with current footnotes).

[30]    The exposition in the single article on the subject attributed to Thomas Bayes, published posthumously in 1763 (after editing by Richard Price, and with an introduction and commentary by him) in the Philosophical Transactions of the Royal Society, set out certain important aspects of what is now known as Bayes's Theorem, without any specific algebraic formulae. *See* (as the title quaintly puts it) F.R.S. Bayes, *An Essay Towards Solving a*

which are not implied by Bayes's Theorem itself.

But let me not put forth such a perhaps controversial and contentious proposition on my own authority.  Let me quote the eminent statistician, the late Jack Good, an early adherent, who said only a few years ago that the term Bayesian is now usually used to refer to "a whole philosophy or methodology in which 'subjective' or logical probabilities are used."[31]  So Bayes's Theorem does not define modern Bayesianism (or "neo-Bayesianism" as it is sometimes called) so much as the mysterious concept of "subjective" probabilities[32] does (that is, mysterious at least in some ways to me[33]).

---

*Problem in the Doctrine of Chances.  By the Late Rev. Mr. Bayes, F.R.S.  Communicated by Mr. Price, in a Letter to John Canton, A.M.F.R.S.*, 53 PHIL. TRANSACTIONS 370 (1763), http://rstl.royalsocietypublishing.org/content/53/370.  But it was clearly Laplace, almost certainly independently, whose work gave the more generalized notion its impetus and content.  *See generally* STEPHEN M. STIGLER, THE HISTORY OF STATISTICS: MEASUREMENT OF UNCERTAINTY BEFORE 1900 99–138 (1986).  As Stephen Fienberg has noted (noting others who have noted the same thing), "Bayes did not actually give us a statement of Bayes' Theorem, either in its discrete form (this came with Laplace), or in its continuous form with integration, although he solved a special case of the latter."  Stephen E. Fienberg, *When Did Bayesian Inference Become "Bayesian?,"* 1 BAYESIAN ANALYSIS 1, 3 (2006) (formula and note omitted).

[31]  Quoted from Good's personal correspondence with Stephen Fienberg in Fienberg, *supra* note 30, at 15.

[32]  Professor Fienberg traces the genesis of the explicit notion of "subjective" probabilities as degrees of belief to John Maynard Keynes's 1921 TREATISE ON PROBABILITY, although he notes that Keynes "allowed for the possibility that degrees of belief might not be numerically measurable."  Fienberg, *supra* note 30, at 9.  Fienberg also notes the use of the notion of "degrees of reasonable belief" (the qualifier seems later to have fallen out) by Harold Jeffreys and Dorothy Wrinch in a 1919 paper.  Fienberg, *supra* note 30, at 10–11.  Later, beginning with Borel and Ramsay, approaches to quantifying subjective probability through mind experiments involving betting under various assumptions and constraints were put forth as adequate mechanisms for quantification.  Fienberg, *supra* note 30, at 9–10.

[33]  I am aware of one sense in which all probabilities must be regarded as, in that sense, subjective, in the same way that language, being an abstraction, is in the same sense subjective.  Both are products of mind, and are therefore mind-dependent.  Something like this seems to have been the starting point for Bruno De Finetti.  *See* Robert F. Nau, *De Finetti Was Right: Probability Does Not Exist*, 51 THEORY AND DECISION 89 (2001) (quoting BRUNO DE FINETTI, THEORY OF PROBABILITY VOL. I, 1 (English trans., 1974)) ("Probability does not exist.").  Perhaps the most mysterious thing, then, is why anyone claims any probabilities are <u>not</u> subjective.  However, I think I get the main point of the frequentists:  despite various levels of mystery (including the ontological status of mathematics in general, see MARY LENG, MATHEMATICS AND REALITY (2010)), objective or "frequentist" probability is the appropriate model when one has reason to believe that abstract notions of exact mathematical probability specification (represented in their simplest form by specified "mental experiment" urn-drawing problems or similarly specified mental experiments regarding gambling games) actually represent a reasonable model of the empirical world in which a problem is faced and must be answered (real poker games, etc.).  When such circumstances do not exist (and more often than not they don't), best guesses (subjective probabilities) must be resorted to, or else

There seems to be a consensus that neo-Bayesianism, or simply "Bayesianism" full stop, as it is now known, as a "whole philosophy or methodology" began to emerge as a movement in the 1950s, and came to full flood in the 60s with, among other things, the emergence of influential rational choice models of expected utility decision theory.[34] These models, perhaps best represented by the work of Howard Raiffa and Robert Schlaifer,[35] have plenty of helpful applications, but they may not be as universalizable as their proponents believe.[36] But nevertheless, Bayesianism based on freewheeling notions of subjective probability became a constellation of beliefs and investments that attracted many adherents.[37]

While the formal expression of Bayes's Theorem (which was actually generated by Laplace) is no doubt an analytic truth when placed in an abstracted formal system, and it can be a powerful model of rational inference whenever well-warranted quantification is available in regard to information being added to a system, it is of little use without such numbers, since both the initial characterization of the starting point (the "prior probability") and the updating by new information, only work in any helpful way with actual cardinal numbers. But when there is no fully specified formal system (like the artificial worlds of gambling games or urn-drawing experiments) to provide numbers, and when there is no real empirical data to provide numbers, what to do? The Bayesian inference machine will not work without numbers. The modern Bayesian's answer is, we will just make them up.

---

we must defer decision.

    But in both language and in notions of probability, there is a necessary contribution from an exterior physical reality, and more so in regard to warranted probability statements about factual conditions than in regard to some kinds of natural language expressions such as value statements.

[34]   Fienberg, *supra* note 30, at 14–20.

[35]   The wellspring work is HOWARD RAIFFA & ROBERT SCHLAIFER, APPLIED STATISTICAL DECISION THEORY (1961).

[36]   *See generally* Matt Jones & Bradley C. Love, *Bayesian Fundamentalism or Enlightenment? On the Explanatory Status and Theoretical Contributions of the Bayesian Model of Cognition*, 34 BEHAV. AND BRAIN SCIENCES 169 (2011).

[37]   In this, it reminds me somewhat of another movement of a hundred years earlier, which I have given some study to (and written about a bit), that is, August Comte's Positivism, which was driven to a similar secular desire to identify and embody pure rationalism, and which ironically turned into a kind of secular religion. *See* D. Michael Risinger, *Boxes in Boxes: Julian Barnes, Conan Doyle, Sherlock Holmes and the Edalji Case*, 4 INT'L COMMENT ON EVIDENCE 1, 7, n.13 (2006).

Well, not exactly, but almost.  In the 1920s Frank Ramsay of Cambridge, perhaps influenced by Emile Borel, proposed a way of generating numbers that he claimed would then work for probability and Bayesian inference purposes every bit as well as the frequency numbers that were a normal product of the analysis of formally specified systems like games.  He said that a person could conduct a mental experiment involving betting on a proposition of fact, and that the honest answer to what you would wager represented a number that then could be used in Bayes's Theorem properly.[38]

Once this process is reflected upon, objections abound.  First, the resulting numbers are not really the specific probability statements that the numerical values resulting from the forced choice of the mental experiment would suggest (they have illusory precision).  Second, they are grossly unreliable between persons in many applications—different people will give significantly different numbers.  Third, when used for deriving probability statements about facts in the world, even the relationship of even the average of multiple bets by multiple people may have a weak bearing on the underlying empirical reality.[39]  And even this operation may in some sense be out of bounds, because many Bayesians will not necessarily go past saying that an *individual's* actions based on their beliefs (as measured by their bets) should obey the probability calculus, including Bayses's Theorem, even if those beliefs are wrong or ill-founded.  All this may be fine for personal uses in defining personal utilities in decisions where the costs of being wrong will fall on the person making the subjective bets, but not otherwise.  Additionally, using a betting model assumes people are familiar with betting and skilled at making real wagers.  This is hardly a social given.  And there are other available objections.[40]

---

[38]   *See supra* note 32.

[39]   That is, crowd-sourced estimates don't always get close to truth, even in regard to simple current countable facts in the world.  They may only work when the crowd has fairly specific actual knowledge, providing some legitimate bounds of estimate, as was the case in regard to the weight of the butchered meat that would be yielded by a particular ox (Francis Galton's the original example of a crowd-sourced average converging on a value virtually the same as the true value). *See* Francis Galton, Letter to the Editor, "*Vox Populi,*" 75 NATURE 450, 450–51 (1907), http://wisdomofcrowds.blogspot.com/2009/12/vox-populi-sir-francis-galton.html.  If you asked the same crowd to estimate how many beans were in a container that might be as small as a thimble or as large as a dump truck, with the volume of the container being generated at random between these two bounds, it is doubtful that you would get a magic convergence.

[40]   For a catalogue of objections to the betting model for generating numbers, see Lina Eriksson & Alan Hajek, *What Are Degrees of Belief?*, 86 STUDIA LOGICA 183, 186–90 (2007).  One should note that Ramsay and later followers regarded the bet as a proper number only if one were willing to act on it, but what that would mean in many contexts where subjective probabilities of this sort are used is difficult to say.

Let me be clear here.  I am not saying that there are no lessons to be drawn for practical reasoning from an understanding of Bayesian principles.  One does not have to be a full-tilt Bayesian to see the desirability of alternative hypothesis testing.  One only has to be a fan of Bacon's method of disconfirmation.[41]  Formal consideration of two competing hypotheses guards against tunnel vision, and can even be the basis for informal expressions of a likelihood ratio, or likelihood ratio-like opinion expressing one's expectancy of seeing the evidence given hypothesis one, versus one's expectancy of seeing the evidence given hypothesis two, without resort to unwarranted cardinality.[42]  And similar observations could be made concerning the necessity of having some informal understanding of the role of prior "probabilities"[43] so as not to misinterpret various pieces of evidence that might be proffered at trial.

But none of this requires, or justifies, the generation of cardinal number expressions by anyone for actual use in the inferential process of the trial.  So I am not so much anti-Bayes as anti-unwarranted cardinality.  My emphasis is not on the subjective nature of probabilities *vel non* in a general

---

[41]   This "disconfirmation" approach is more commonly associated with Karl Popper, but Bacon was first: "[I]t is the peculiar and perpetual error of the human understanding to be more moved and excited by affirmatives than negatives, whereas it ought duly to be impartial; nay, in establishing any true axiom, the negative instance is the most powerful." LORD FRANCIS BACON, NOVUM ORGANUM 24 (Joseph Devy ed., 1902) (1620). This is also central to L. Jonathan Cohen's system of "Baconian (non-Pascalian) probability. *See generally* L. JONATHAN COHEN, THE PROBABLE AND THE PROVABLE (1977).

[42]   Although I see the basic idea of the likelihood ratio as an important part of practical reasoning, I must part company at least to some degree with Ed Cheng regarding his suggestion that standards of proof can helpfully be defined in terms of required likelihood ratios. *See generally* Edward K. Cheng, *Reconceptualizing the Burden of Proof*, 122 YALE L. J. 154 (2012). While I understand that his claim is descriptive, and that he does not necessarily suggest that juries be instructed in accordance with his account, either way, it seems much too dependent on unwarranted cardinality for my taste.  On the other hand, properly operationalized levels of surprise (the non-cardinal flip side of the expectancy that drives probability-based likelihood ratios) might achieve an approximate correspondence to what he envisions for likelihood ratios.

[43]   I often use the term "probability" as a more general and less technical term than would satisfy a formal probability theorist.  Shafer and Vovk see one of Kolmogorov's most interesting moves to have been declaring probability itself to be, not something represented by a number, but the number itself, and not just any number, but one conforming to his axioms, referring to this as "the bold rhetorical move that Kolmogorov made in the Grundbegriffe, giving the abstract theory the name probability," and further explaining: "Kolmogorov's appropriation of the name *probability* was an important rhetorical achievement, with enduring implications. Slutsky in 1922 and Kolmogorov himself in 1927 had proposed a general theory of additive set functions but had relied on the classical theory to say that probability should be a special case of this general theory.  Now Kolmogorov proposed axioms for probability.  The numbers in his abstract theory were probabilities, not merely valences . . . ."  Glenn Schafer & Vladimir Vovk, *The Origins and Legacy of Kolmogorov's* Grundbegriffe, 35, 40 (Game Theoretic Probability and Finance Project, Working Paper No. 4, 2013), http://www.probabilityandfinance.com/articles/04.pdf.

sense, but on the empirical warrant which is, or ought to be, necessary to derive and use any probability statement in different contexts. It may be that numbers derived from subjective individual mental betting experiments are perfectly sufficient for some kinds of problems (contexts in which no iterative process is available to refine prior assumptions out of results, a decision must be made, and all the costs will fall on the betting party or parties, for instance), but not in others. However, appropriate criteria for acceptable best guesses for different classes of problems seem hardly ever to be discussed, at least so far as I have been able to discover, by Bayesians claiming to model all or part of the legal process. So again, perhaps I am not so much anti-Bayes as anti-unwarranted cardinality.

So what does this have to do with standards of proof? One reason some Bayesians think standards of proof are a good target for Bayesian mathematization is that standards of proof, whatever else they are for, are intended to speak to jurors about their level of certainty concerning the material issues of the case they are deciding after they have seen the evidence produced at the trial, and it is easy to view such levels of subjective certainty as "probability statements." In a general sense of the word "probability," this is true, but like all such individual personal level of certainty about specific factual issues (and, again, I will limit myself for the time being to actual binary "facts in the world," such as perpetrator identity) these subjective states which do not necessarily reflect a mathematized (or perhaps even mathematizable) probability directly. Unlike intelligent evaluations of the state of play in a well-specified and well-understood gambling game like dice or roulette, such subjective states are best described by the term "degrees of belief," a usage first put forth, apparently, by Jeremy Bentham,[44] but which re-entered philosophical and mathematical consideration of such things in the late 1910s and early 1920s. In fact, degree of belief "is *the* central notion of subjective probability theory, the core concept around which the entire theory is built."[45] Bayesians commonly turn degrees of belief into cardinal probability numbers by hypothetical betting exercises. What the betting model does for modern Bayesians is provide numbers which they claim measure degrees of belief in a way that can then be fully mathematically operationalized according to the normal rules of classical probability theory.

---

[44]   "Nobody can be ignorant, that belief is susceptible of different degrees of strength, or intensity." JEREMY BENTHAM, A TREATISE ON JUDICIAL EVIDENCE 40 (M. Dumont ed., 1825). This is from the one-volume work edited by M. Dumont containing Bentham's views on evidence, not the five-volume work on the same subject edited by Bentham's protégé John Stuart Mill and published two years later. The statement holds true as much for normative or metaphysical statements as for statements concerning things properly characterized as factual, but we will come to that in due course.

[45]   Eriksson & Hayek, *supra* note 40, at 184.

I am skeptical of this approach for a couple of reasons.  First, while I accept that standards of proof are the law's attempt to impose a requirement of different degrees of belief for different kinds of issues, this does not mean that any statement about appropriate system performance over large numbers of cases was the purpose for the creation of those standards.  Rather, it seems that the purpose of such standards at their creation was to define the level of subjective certainty necessary for such a decision to be a morally justified decision,[46] a purpose for which formalized modern probability notions are ill-suited functionally and expressively.[47]  This does not mean that we are wrong in common parlance to say that standards of proof involve judgments of how likely, or how probable, the existence of a fact is, but merely that those terms (like Blackstone's use of his ratio) are not properly seen as statements entailing or implying any formal version of modern

---

[46]   Jim Whitman  has very effectively shown this in JAMES  Q. WHITMAN, ORIGINS OF REASONABLE DOUBT: THE THEOLOGICAL ROOTS OF THE CRIMINAL TRIAL (2008), and none of the other historical analyses bearing on the development of legal standards of proof cited in note 46 *infra* contradict that conclusion.  The literature is reviewed and evaluated in Miller W. Shealy, Jr*., A Reasonable Doubt about "Reasonable Doubt,"* 65 OKLA. L. R. 225, 268–90 (2013).  Professor Shealy sees more tension between the analysis of Professor Whitman and those that speak more of emerging sensitivity to notions of probability, particularly Barbara J. Shapiro, than I do.  *See id.*

[47]   Until fairly recently, I also thought of standards of proof primarily as devices for distributing error.  However, I now think that although such error distribution is an inevitable result of the operation of such standards, the purpose, as opposed to the consequence, of such standards, is to get the jurors to understand the moral burden they bear in the individual case, not to reflect upon or speculate about how this ties in to system error and its distribution.  Kevin Clermont's ruminations on standards of proof over a long period of time seem generally consistent with this position, and Clermont establishes it to be the view of much of the rest of the world.  *See generally* Kevin M. Clermont, *Procedure's Magical Number Three: Psychological Bases for Standards of Decision*, 72 CORNELL L. REV. 1115 (1987); Kevin M. Clermont & Emily Sherwin, *A Comparative View of Standards of Proof*, 50 AM. J. COMP. L. 243 (2002); Kevin M. Clermont, *Standards of Proof in Japan and the United States*, 37 CORNELL INT'L L. J. 263 (2004); and Kevin M. Clermont, *Emotions in Context: Exploring the Interaction Between Emotions and Legal Institutions: Standards of Proof Revisited*, 33 VT. L. REV. 469 (2009).  *See generally* KEVIN M. CLERMONT, STANDARDS OF DECISION IN LAW: PSYCHOLOGICAL AND LOGICAL BASES FOR THE STANDARD OF PROOF, HERE AND ABROAD (2013).  Clermont labels the individual degree of belief required in such settings "confidence" as opposed to probability.  This is potentially a bit confusing, since "confidence" is also a standard concept in formal probability theory applied to implications of data derived from sampling techniques.  For an attempt to deploy that formal concept into a metaphorical explanation for why it is wrong to conceive of the preponderance standard as a knife-edge point specification of probability which is the least bit greater than 0.5, see Neil B. Cohen, *Confidence in Probability: Burdens of Persuasion in a World of Imperfect Knowledge*, 60 N.Y.U L. REV. 385, 397 (1985).  There have been other approaches to the problem of the knife edge probability interpretation of the preponderance standard.  *See* the discussions of Cohen and others in CLERMONT, *supra* note 11, at 16–25 (sub nom "equipoise as a range"), and the extended skeptical discussion examining such positions through the lens of total system error and optimized Keynsesian weight, and viewing such approaches as calling for an unjustified "tenacity of belief" in preponderance decisions in DALE A. NANCE, THE BURDENS OF PROOF: DISCRIMINATORY POWER, WEIGHT OF EVIDENCE, AND TENACITY OF BELIEF 251–78 (2016).

mathematized probability theory or statistical theory.    The notions of "probable" and "likely" predate Pascal by millennia.[48]

But my skepticism is driven by something deeper.  I think that betting exercises are poor proxy measures of the thing that, to my mind, truly reflects degrees of belief, at least in regard to facts in the world—the intensity of the surprise that would be experienced if it were established that the proposition believed to be true is in fact false.  My central claim is that people believe something to be true to the extent that they would be surprised to find out it was false.[49]

Another thesis is that the extent of surprise one would feel upon discovering a belief to be false is the best measure of subjective certainty, of the degree of belief,  I propose that a person's own degree of belief is best revealed to his or her self not by the betting exercises usually invoked in Bayesian approaches such as decision theory, but by a different mental experiment—not by asking what one would bet on the truth of a belief, but asking directly how surprised one would be to find out that the thing believed was false.  Finally, I believe that standards of proof in regard to facts are addressed to making jurors determine this degree of belief, and then norming them to understand the degree of belief which the law requires for an affirmative finding.

But how do we measure these levels of surprise, or at any rate induce jurors to measure them in themselves?[50]  Not by artificially generated cardinal numbers to put into a full probability calculus, but by a well-ordered system of categories designated by words like what are now called (rather

---

[48]    There is a rich literature on pre-Pascal probability notions, the most complete of which is JAMES FRANKLIN, THE SCIENCE OF CONJECTURE: EVIDENCE AND PROBABILITY BEFORE PASCAL (2001); *see also* WHITMAN, *supra* note 46; BARBARA J. SHAPIRO, BEYOND REASONABLE DOUBT AND PROBABLE CAUSE:  HISTORICAL PERSPECTIVES ON THE ANGLO-AMERICAN LAW OF EVIDENCE (1991); IAN HACKING, THE EMERGENCE OF PROBABILITY: A PHILOSOPHICAL STUDY OF THE EARLY IDEAS ABOUT PROBABILITY (1975); LORRAINE DASTON, CLASSICAL PROBABILITY IN THE ENLIGHTENMENT (1971).   All seem to agree that legal proceedings and theological controversies generated much of the non-mathematized thought on certainty and uncertainty.  Nor was it all informal, as illustrated by the Romano-canonical system of proof through categories amounting to proof fractions, with the requirement that the evidence mount up to "full proof." This even used numbers, in a way, though not in a way we would easily call "mathematization," or would find very congenial today.

[49]    The second epigram set out for this part of the paper ("that's funny") was to show surprise accompanying disturbed belief.  Apparently my willingness to define belief in terms of such a subjective emotion as surprise makes me a "liberal dispositionalist" in regard to philosophical accounts of belief. *See Belief*, STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Mar. 24, 2015), http://plato.stanford.edu/entries/belief/.  I am not surprised that some others have embraced the centrality of surprise, but I have not encountered it in any of the legal literature on standards of proof or the jury function.

[50]    This is a major issue—how to get some sort of relative weight out of what starts out as an ordinal system.

unfortunately, I now believe) "words of estimative probability."[51]  In fact, the traditional three standards of proof of our litigation system are a rank-ordered, three-category system defined by various formulas which can be thought of as words (or word formulas) of estimative probability.  But what I propose is a system which uses what I would prefer to call "words of estimative surprise," such as mildly surprised, surprised, quite surprised, greatly surprised, astonished, shocked, etc.

One interesting thing about such scales is that they can be easily understood as rank-ordered, and therefore having the transitive property of numbers, but as not having any of the other properties of mathematization such as additivity.  On the other hand, they are inherently imprecise, not merely because they are not as fine-grained as numbers, but because they do not represent equal quantities.  They are subjective, not merely in the sense that they are mind-dependent, but in the sense that there will be variability, very likely significant variability, between individuals in the strengths of surprise their designations represent.  What is needed is a way of norming jurors into reasonably consistent designations in the use of the scale.[52]  In various proposed alternatives to ordinary Bayesian approaches, there have been various attempts to supply what is regarded as needed or desirable cardinality.[53]  But I do not think that strict cardinality is needed at all.  Here

---

[51]    *See e.g.*, Sherman Kent, *Words of Estimative Probability*, originally published (though classified) *in* STUDIES IN INTELLIGENCE (1964), https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/sherman-kent-and-the-board-of-national-estimates-collected-essays/6words.html; Frederick Mosteller & Cleo Yountz, *Quantifying Probabilistic Expressions*, 6 STAT. SCI. 2 (1990).

[52]    This is, I suppose, an attempt to supply what Wigmore said was wanting:  "[N]o one has yet invented or discovered a mode of measurement for the intensity of human belief," JOHN H. WIGMORE, 9 A TREATISE ON THE ANGLO-AMERICAN SYSTEM OF EVIDENCE IN TRIALS AT COMMON LAW § 2497 (3d ed., 1940).  This is a major issue—how to get some sort of relative weight out of what starts out as an ordinal system. Thus it is not only a problem for degrees of belief, but a problem inherent in pure rank ordering.  Consider the following (the subject of the comment is something called "possibility theory" and it is given to set up a transition to something called "ranking theory," which tries, not altogether successfully, in my opinion, to put cardinal numbers on such ranks.  But the main point holds, I believe, for all purely ordinal systems): "Note, though, that an ordered partition . . . contains no more than ordinal information.  While your possibility distribution enables you to say how possible you take a particular possibility to be, an ordered partition only allows you to say that one possibility $w_1$ is more plausible than another possibility $w_2$.  In fact, an ordered partition does not even enable you to express that the difference between your plausibility for $w_1$ (say, tomorrow the temperature in Vienna will be between 70°F and 75°F) and for $w_2$ (say, tomorrow the temperature in Vienna will be between 75°F and 80°F) is smaller than the difference between your plausibility for $w_1$ and for the far-fetched $w_3$ (say, tomorrow the temperature in Vienna will be between 120°F and 125°F)."  Franz Huber, *Formal Representations of Belief, in* THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY ONLINE 33 (Spring 2016 ed.), https://plato.stanford.edu/entries/formal-belief/.

[53]    For instance, in Dempster/Schaffer Belief Functions, some amount of belief can be reserved because of incompleteness of information.  The belief given to the truth of a

is where Gladstone and Victoria can come to the rescue.

The "story model" of human information processing suggests that story form facilitates human understanding of the meaning of information in many contexts. [54]   What Chesterton has done in the little example/anecdote reflected in the initial epigram that starts this piece is tell a small story that illustrates how one can be virtually certain of some facts in the world without recourse to mathematical expression.  Indeed, the level of certainty reflected in the Gladstone-and-Victoria example goes well beyond what most people would require as proof beyond a reasonable doubt.  My notion is that various degrees of belief might best be illustrated to juries by such illustrative mini-stories which would capture the law's notion of the required level of surprise one should feel were things not as one must find them to be in order to give judgment.[55]

Here I must make a confession.  I do not have a system of such preferred and acceptable illustrative anecdotes to offer.  Gladstone and Victoria won't work very well in the modern world, because the social knowledge that makes it so powerful is simply not part of the average juror's social knowledge today.  Such anecdotes would have to be carefully crafted to be consistent with such conditions, and given the splintering of such knowledge

---

proposition can thus be, say, represented as .5 and the belief in the negative as 0.2, with 0.3 reserved or uninvested belief because of the weakness of the evidence.  So the probability or belief numbers do not add up to 1, but the sum of all three categories will be 1.  In circumstances not based on formal data, the theory thus has the same "numbers from nowhere" problem as modern subjective Bayesianism.  Dempster/Schaffer Belief Functions are sometimes referred to as an expansion of Bayesian reasoning.

[54]   The "story model" is a model of human cognition and inference that suggests that most ordinary humans process the inferential meaning of ordinary circumstantial information best when it is cast into story form, and that the generation of such plausible stories is central to both accurate human inference and persuasion.  It may be seen as an extension of schema theory in psychology, but in the legal setting it is based on the work of  Nancy Pennington and Reid Hastie.  *See* Nancy Pennington & Reid Hastie, *Evidence Evaluation in Complex Decision Making*, 51 J. PERSONAL & SOC. PSYCH. 242, 242–46 (1986); Nancy Pennington & Reid Hastie, *The Story Model of Juror Decision Making*, *in* INSIDE THE JUROR: THE PSYCHOLOGY OF JUROR DECISION MAKING (Reid Hastie ed., 1993).  Professor Allen was an early adopter.  *See, e.g.*, Ronald J. Allen, *The Nature of Juridical Proof*, 13 CARDOZO L. REV. 373, 410 (1991).

[55]   And it probably goes without saying that I reject any attempt to give expressions of mathematical modeling to juries in order to define standards of proof.  For recent arguments (more or less) in favor of such an instruction, see Peter Tillers & Jonathan Gottfried, United States v. Copeland: *A Collateral Attack on the Legal Maxim that Proof Beyond a Reasonable Doubt Is Unquantifiable?*, 5 LAW, PROBABILITY & RISK 135 (2006); Jack B. Weinstein & Ian Dewsbury, *Comment on the Meaning of "Proof Beyond a Reasonable Doubt,"* 5 LAW, PROBABILITY & RISK 167 (2006); Jon O. Newman, *Quantifying the Standard of Proof Beyond a Reasonable Doubt: A Comment on Three Comments*, 5 LAW, PROBABILITY & RISK 267 (2006).  For a critique with which I largely agree, see James Franklin, *Case Comment—*United States v. Copeland*, 369 F. Supp. 2d 175 (E.D.N.Y. 2005): Quantification of the "Proof Beyond Reasonable Doubt,"* 5 LAW, PROBABILITY & RISK 267 (2006).

into various subcultures, it seems a daunting task. In addition, I do not necessarily think that the only means to norm the levels of surprise representing required degrees of belief is the story form anecdote or parable. Other images, even ones using numbers in a metaphorical and informal way, such as the expression "one in a million," might be serviceable. Even simply the use of the terms of estimative surprise, like "shocked" for a description of reasonable doubt, might do. As a beginning, I am attaching a schematic which I find helpful in organizing my current thinking, if nothing else. It is perhaps best viewed as a discussion starter.[56]

Finally, the title of this piece asked in part, "What should we say to jurors to get what we want?" Up to now I have proceeded as if this "saying" is all to be done by the court through instructions. However, lawyers also address these issues in closing, and it may be that harnessing the story model and the notion of surprise to the end of alerting and norming the jury to the appropriate degree of belief required by the law's standards of proof might best be done in closing argument. At least it might be found a useful approach by those who do not have the burden of persuasion as a way to effectively argue these points.

---

[56] In viewing the schematic, one should keep in mind that it represents an ordinal system, and that the size of the boxes used to define the ordinal categories are not meant to be proportional to any cardinal system, or the number of cases in real world that might fall within each box.

**Degrees of belief for a factual proposition or complex of propositions (story, account; part of story) measured by surprise**

Note – this does not map on to degrees of belief about normative judgments very directly

Direction of increasing surprise if proposition under consideration were false

Direction of increasing surprise if proposition under consideration were true

Limit: Absolute Truth

Limit: Absolute Falsity

Vanishingly unlikely barely imaginable (alien manipulation)

Virtually certain not. Imaginable but extremely unlikely (Gladstone & Queen Victoria)

I'm sure not. Shocking if true. Possible but very unlikely (bad fiction)

Extremely improbable (I'd be very surprised if true)

Very probably not (I'd be surprised If true)

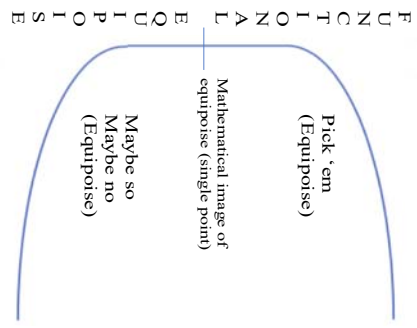Probably not (I'd bet against but not be surprised to lose)

A definite maybe

Pick 'em (Equipoise)

Mathematical image of equipoise (single point)

Maybe so Maybe no (Equipoise)

Maybe but maybe not

Probably (I'd bet but not surprised to lose)

Very probably (I'd be surprised if not) (alternatives unlikely)

Extremely probable (I'd be very surprised if not)

I'm sure (shocking if not) (alternatives extremely unlikely)

Virtually certain (alternatives imaginable but VERY extremely unlikely)

Certainty (alternatives barely imaginable, vanishingly unlikely)

Beyond reasonable doubt

Almost certain

Clear and convincing

Preponderance

FUNCTION OF ACQUIESCE

IV. THE LIMITS OF SURPRISE AS A KEY TO UNDERSTANDING STANDARDS
OF PROOF

When I first got excited about the potentials for surprise in describing and norming standards of proof, I envisioned a systematic treatment that could be deployed to rationalize any legal standard of proof in any context whatsoever. Then, on reflection, I began to suspect that surprise was not quite the universal can opener I had hoped.

I suppose I am someone who should have been least subject to being caught up by such a generalized enthusiasm. The "surprise" approach works well, and better than a formal probability approach, when decision thresholds concerning binary facts in the world are at issue. This by itself makes it exciting to me, since the main focus of my concern in other parts of my professional life is actual factual innocence in criminal cases, and the surprise approach (coupled with an associated word of estimative surprise such as "shocked") can be deployed there to great effect.

But the issues labeled "facts" in the legal process are often not facts at all, but often other kinds of issues which have been labeled issues of fact to justify or account for our committing them for resolution to "factfinders" such as juries.[57] And I have taken some effort to try to survey the kinds of issues that we do submit to the jury under the label "fact," and to analyze the actual "facticity" of each of the common recurring kinds of issue in both civil and criminal cases.[58]

---

[57] The truth of this general proposition has been long accepted. For an excellent treatment, see Ronald J. Allen & Michael S. Pardo, *The Myth of the Law-Fact Distinction*, 97 NW. U. L. REV. 1769, 1769–70 (2003). The locus classicus is Stephen Weiner, *The Civil Jury Trial and the Law-Fact Distinction*, 54 CAL. L. REV. 1867, 1876 (1966): "Clarity of thought is not advanced by debating whether law application is law-making or fact-finding, as commentators have done. Moreover, it is meaningless to assign the task in a specific case to judge or jury simply by use of the law and fact jargon, as courts have done. A far preferable solution would be to recognize a third category, and to deal with it as such, not on the basis of terminology, but on the basis of policy." *Id.*

Weiner then goes on to say that juries are often asked to determine "reasonableness." *Id.* "Such a flexible norm inevitably invests the decision-maker with considerable discretion to draw upon his own notions of how people should conduct themselves." *Id.* Here Weiner recognizes that the ultimate issues we commit to juries often grant to them what I have labelled a "normative warrant," which both authorizes and requires them to make a particularized value judgment to determine the element of the claim or defense under consideration. *See* D. Michael Risinger, *Unsafe Verdicts: The Need for Reformed Standards for the Trial and Review of Factual of Innocence Claims*, 41 HOUSTON L. REV. 1281, 1282, 1288–95 (2004).

[58] *See* D. Michael Risinger, *Searching for Truth in the American Law of Evidence and Proof*, 47 GA. L. REV. 801, 818–34 (2013), *sub nom* "II. Facts, Facticity and Facts of another Color." In this piece I used the simple binary fact-in-the-world ("Was A present on the northwest corner of Raymond Boulevard and McCarter Highway at 1:00 p.m. on June 1, 2013?" for example) as the model of full facticity, and analyzed other recurring legal issues such a negligence, market value, and various states of mind, in regard to how and how far they departed from the simple binary fact-in-the-world, in order to determine their claim to

And why is this?  Because there are beliefs that the mental experiment of potential surprise does not apply to, at least not sufficiently to do the desired work of norming degrees of belief, and unfortunately the law uses issues involving such beliefs as elements of claims and defenses (under the inappropriate label "fact").  It is only of marginal comfort to observe that notions of probability work no better for such issues.

Belief covers a multitude of creedal states in regard to a multitude of propositions.  Take the existence of God, for instance.  Many people believe in the existence of God.[59]  I am not going to attempt to rehearse all the ways in which this concept does not connote a crisply defined idea, or the many different and personal understandings that people have about the meaning of the term.  They say they believe, and I take them at their word.  In fact, many of them have a special kind of belief that is especially difficult to revise, called faith.  But that is not the real problem.  Whatever the concept of God represents, it would be difficult for someone with that belief to even conceive of knock-down proof of the contrary that they might come to encounter (in this lifetime), in the same way that they could conceive of a newly-found video of an event establishing beyond reasonable doubt that the person they were trying was actually, factually innocent of the charged crime.

I have used this example not just to show that belief covers metaphysical creedal investments not easily subject to being tested through thought experiments involving surprise, but as a bridge to other types of belief with related problems which are closer to the mark when it comes to substantive elements of claims or defenses.

As I have previously indicated, I have long been interested in the various kinds of issues that are submitted to juries under the label "fact."  In my attempts at a taxonomy of ultimate legal elements, I have developed a tripartite classification of such issues—binary facts, magnitude judgment issues, and normative issues.[60]  Binary facts were further divided in simple binary facts (or "brute facts"), such as whether the defendant in a criminal case discharged the fatal bullet into the back of the victim's head, and "facts of another color," that is, binaries that depended on counterfactual judgment, such as causation.  Magnitude judgment issues have no single right answer, and are best illustrated by determinations of market value; normative warrant issues are things like negligence, where the final determination is a particularized value judgment based on factual context, but going beyond it (was the action of the defendant careful enough, or not).  I later expanded upon and refined this approach in an attempt to describe the relative

---

one or another degree of "facticity." *See id.*

[59]  For these purposes I use the singular to denote a monotheistic belief.

[60]  *See* Risinger, *Unsafe Verdicts*, *supra* note 57, at 1290–95.

"facticity" of each such element.[61]

Without going into too much detail, suffice it to say that invoking a juror's potential surprise in discovering later that current belief was wrong makes more sense in regard to some issues than others. It works best as to binary facts in the world, less well but perhaps serviceably in regard to some states of mind (such as acting "knowingly") but not others, (such as acting "recklessly," which is actually a normative judgment). In addition, it asks a near meaningless question in regard to magnitude judgement issues (how surprised would you be to find out that the number you have assigned as the market value of these goods was not truly the market value), and a truly meaningless question in regard to normative issues such as negligence, where there is no fact of the matter, and the factfinder is the ultimate arbiter.

So surprise has problems of application to a range of non-fact "facts" that constitute elemental decisions submitted to juries. But formal probability approaches have similar problems of equal or greater magnitude when deployed in regard to such issues. Where surprise enjoys the greatest application, and the greatest advantage, as an approach to defining and norming standards of proof is in regard to actual facts-in-the-world issues such as whether the person charged with a crime was in fact the perpetrator of that crime. Maybe this is not sufficiently general an application to remake our views of standards of proof, but given my other investments concerning factual innocence, it is enough for me.

### AFTERWORD

Other participants in the symposium panel for which this paper was written wrote many kind and generous things regarding it, and made a number of cogent suggestions and observations which I want to acknowledge, and occasionally give some response to, here.

First, in regard to Professor Allen's piece, I have picked out four passages which I think make most of his critical points. I will set them out below, following each with a response which will not necessarily put his concerns to rest, but will merely outline the contours of my current thinking, which will have to be further developed in future work.

> 1. *Asking people if they would be really surprised sounds similar to asking them if they would make a really big bet on the truth of some proposition. One loses much of the ability to compute easily without cardinal answers, but you get the same basic output. Or so it seems to me at the moment. I look forward to further*

---

[61]    *See supra* note 55.

*explication from Professor Risinger.*[62]

I concede that mental experiments involving betting and mental experiments involving the intensity of surprise are related, since both are pointed toward trying to establish a degree of belief. However, I think they are different in important ways. My notion is that betting experiments pose the wrong set of mental operations to the factfinder, suggesting that the focus is on the factfinders' personal utilities and opening the door to "floating reasonable doubt" effects in the presence of a grisly crime or devastated victims in the courtroom. This is not made better when the factfinder is not an experienced bettor, although it is questionable whether most experienced bettors are actually good at rational bets. Most often they would be more rational to refuse to bet at all. Be that as it may, the "surprise" question asks the pertinent question more directly, in terms more naturally accessible to the entire range of factfinders, and concentrates on the issue directly on certainty as to guilt. Or so it seems to me at the moment.

> 2. *The second point is perhaps more fundamental, and it is one that Professor Risinger alludes to but does not develop, and that is the ambiguous relationship between ex ante and ex post analysis. Ex ante, the question fact finders face is whether a burden of persuasion has been satisfied, however that is accomplished. In asking fact finders how surprised they would be to find out they were wrong, what ex-post counter-factual worlds may be considered? Whether some new evidence might appear? Whether some probative evidence was excluded? Whether witnesses lied? Whether the government planted inculpatory evidence? Whether the defendant murdered a potentially important witness? I suspect, ex post, I would be very surprised by some outcomes, but not others. How the surprise theory handle such matters needs to be developed, I think.*[63]

I am sensitive to the fact that ex ante and ex post judgments must be viewed differently when general notions of error are at stake.[64] However, the mental experiment involved here is not really a true ex post judgment (which would be made after the information contradicting guilt was actually in hand), but a hypothesized ex post judgment. And, assuming a criminal

---

[62]  Ronald J. Allen, *The Declining Utility of Analyzing Burdens of Persuasion*, 48 SETON HALL L. REV 995, 1002 (2018). Note that in what follows I have omitted the footnotes from block quotations unless necessary for clarity.

[63]  *Id.*

[64]  *See* D. Michael Risinger, *Whose Fault: Daubert, the NAS Report, and the Concept of Error in Forensic Science*, 38 FORDHAM URB. L.J. 519 (2010).

case, my approach does not require the factfinder to imagine the specific way in which a provisional commitment to guilt was wrong, or the means of establishing it, but only to assume that by some means it was established, and to gauge the level of resulting surprise. Of course, that *level* of surprise would be influenced by how likely the factfinder felt the possibility of innocence to be by virtue of the strength of the case for guilt—the safety of the conclusion of guilt, as Professor Pardo would have it, about which more later. This is the very question a factfinder should be answering.

> 3. *The nature of the case matters as well. Suppose I am a juror in a criminal case and conclude (using relative plausibility) that there is a plausible case of guilt and none of innocence, although another juror thinks to the contrary. Whatever the outcome of the case (however the initial disagreement was resolved), how surprised would I be that another reasonable person's view of the evidence turned out to be a better appraisal of the situation than mine? Not very, I would think, for these are ambiguous problems. Map this onto Professor Risinger's proposal directly. Now my view is that I just plain would be very surprised if the defendant were innocent, and yours is that you would be very surprised that the defendant is guilty. Having deliberated on those views, I would not be very surprised at all if your view instead of mine were true, again because a reasonable person considering the same evidence as I did reached a different conclusion. How does the surprise theory handle this?*[65]

I assume the conflict was resolved by a hung jury. Any other resolution would seem to raise many other problems beyond the hypothetical. However, if you really wouldn't be surprised that your fellow juror was right in his assessment, then I think you would be obliged to vote to acquit.

> 4. *I encourage Professor Risinger to go more deeply into this interesting theory. Like all good ideas first propounded, Prof. Risinger's is provocative but not necessarily complete, and I raise these points as hopefully helpful suggestions to developing it further. I should also emphasize that Prof. Risinger is not attempting to explain juridical proof, but, I believe, to improve on what he sees. Part of the future development of the idea should include reasons to think fact finders would do better than they presently do by having this structure imposed upon them.*

---

[65]    Allen, *supra* note 62, at 1003.

> *I suspect a bigger problem would be well-intentioned but fundamentally unjustified modifications of methods that turn out to have unintended consequences. Indeed, you might even think that captures a fair amount of the rules of evidence. My advice—were I forced to give it—would be to look for ways to enhance the practical reasoning of fact finders rather than obstruct it. A good example of this is the gradual evolution of the hearsay rule from a rule of exclusion to a rule of admission. A terrible example is the Crawford line of cases that retards reliable decision making for the flimsiest of reasons.[66] It seems like the professionals—the judges, lawyers, and law professors—just cannot stop trying to get fact finders to do what we want them to do rather than just letting them do their job.*

I accept that my proposal is not complete, and would benefit from further development. And indeed, my project has a normative component—I believe that the function of the reasonable doubt standard in criminal cases, for instance, was historically, and ought still to be, a call on the conscience of the individual juror to examine his or her conscience in the most effective way to determine if they are sufficiently sure of guilt to meet our societal commitment to those charged with crime that they shall not be punished if there is any reasonable doubt as to their factual guilt. I also believe that the most effective means to accomplish this is through the mental experiment centered on surprise previously discussed. In putting this forth, I am attempting to get them to do what we want (or should want), but I am not interfering with them doing their job, because I believe that this is exactly what their job is.

In his contribution to this symposium, Professor Pardo deploys notions of modal argumentation and the justifications for modal arguments to supplement, and in some ways reinforce, the points I have put forth about the uses of surprise.[67]

---

[66]  *See* Crawford v. Washington, 541 U.S. 36 (2004).

[67]  Before proceeding further, I should define "modal" in the sense Professor Pardo uses it. Here is a handy explanation from the Stanford Encyclopedia of Philosophy: "Whereas facts about what is *actual* are facts about how things are, facts about *modality* (i.e., what is possible, necessary, or impossible) are facts about how things could, must, or could not have been. For example, while there are in fact eleven players on a soccer team, there could have been thirteen, though there couldn't have been zero. The first of these is a fact about what is actual; the second is a fact about what was possible, and the third is a fact about what is impossible. Humans are often disposed to consider, make, and evaluate judgments about what is possible and necessary, such as when we are motivated to make things better and imagine how things might be. We judge that things could have been different than they actually are, while other things could not have been. These modal judgments and modal claims therefore play a central role in human decision-making and in philosophical argumentation." Anand Vaidya, *The Epistemology of Modality*,

Professor Pardo relates this to notions of modal justification and the "safety" of beliefs.  At the end of this discussion he says:

> *Under [the modal] account, the key issue concerns what would have to have gone wrong, and in what ways, for the jury's verdict to be erroneous?  Answering this question will tell us something about how surprised jurors ought to be if their findings turned out to be false.  We are, in a sense, "surprised" if a particular set of lottery numbers comes up, given the unlikely odds.  But, in another sense, we aren't nearly as "surprised" as when an incredible series of unlikely events all occur.  Apart from the probabilities, it does seem to be the case that events that could not easily occur (unlike the lottery) are more surprising, and are more surprising for that reason.*[68]

I believe I agree with all of this, and thank Professor Pardo for the insight and exposition.  One thing I will say is that the last observation suggests to me that there is a legitimate use in juror deliberation for the evaluation of "Keynsian weight," or lack thereof, that is, how thin the proof proffered was, as compared to a more desirable level of completeness.[69]

Professor Pardo has many other helpful and generally supportive observations in his paper, and I will continue to chew on them as part of the program of further development.

Dale Nance's contribution to the symposium concentrates on the uses of formal expression, and on Shackle's formalization of his possibility theory.  At the conclusion he says thus:

> *I think I have gone far enough in suggesting both the problems and the intriguing lines of inquiry surrounding possibility/ surprise measures as tools of interpreting legal standards of proof and what we ask of fact-finders in their application.  My main point has been to suggest how formal theory can be useful to*

STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Apr. 13, 2015), https://plato.stanford.edu/entries/modality-epistemology/.  Many modal arguments are semantic, but not all.  Arguments utilizing counterfactual reasoning in inferences concerning causation ("but for" arguments, for example) are modal arguments, which depend on the possibility of something happening that concededly did not happen.  The mental experiment concerning surprise which I embrace involves a kind of modal question—how possible is it that the defendant I am I am inclined to convict is factually innocent.

[68]  Michael J. Pardo, *Epistemology, Psychology, and Standards of Proof: An Essay on Risinger's "Surprise" Theory*, 48 SETON HALL L. REV. 1039 (2018).

[69]  Professor Pardo reinforces this when later in the essay he says: "A number of general criteria affect the strength or quality of an explanation. These criteria include considerations such as: consistency, coherence, fit with background knowledge, simplicity, absence of gaps, and the number of unlikely assumptions that need to be made." *Id.* at 1047.

> *someone like Michael in his effort to articulate a better way to understand adjudicative fact-finding. Put simply, my message to Michael is this: formalism can be your friend, too. Personally, I have found it very instructive to pursue Michael's suggestion by taking a few steps toward the articulation of a coherent conception of the standards of proof using possibility theory.*[70]

Over the years I have had a great many lessons from Dale's lucid and careful expositions in any number of areas, from his seminal work on the "better evidence" principle to his fine recent book on the burden of proof. But he has always found more beauty in formal systems than I, being third-eye blind in this regard, can perceive. His careful examination of the details of Shackle's theory of possibility, its ambiguities and its implications, while insightful in its own right, I fear merely reinforces me in the reasons why I eschewed its direct adoption in the first place. I am sure we will continue to till our own fields, and that fecund cross-pollination will occur through the winds of time as a result.[71]

I also want to thank Professor Clermont for his symposium piece.[72] Although his project is concentrated on other approaches to accounting for standards of proof and their operation, he notes that we share a common rejection of formal probability theory as a proper model for these processes. I have learned much for Professor Claremont's writings on this subject, and also, I must say, from Professor Allen's responses to it. I believe that in the end they have more in common than is immediately apparent from their interchanges.

---

[70] Dale A. Nance, *Formalism and Potential Surprise: Theorizing about Standards of Proof*, 48 SETON HALL L. REV. 1017 (2018).

[71] Perhaps a short response to one observation in the Article would not be amiss. The Article declares:
"[W]e need some explanation about why the essentially epistemic enterprise of fact-finding in adjudication should be tied to what Michael regards as an *emotion*. The objection can be made that adjudication is supposed to be, to the extent possible, a rational enterprise, not an emotional one. This is not a trivial problem, as much of the law of evidence is about channeling fact-finders toward rational and away from emotional responses." *Id*. at 1025.

I believe this misconstrues the role of emotion, at least as to some emotions, in rationality. This is a theme I have had for a very long time, and it seems to be supported by much psychological literature on the subject. *See* D. Michael Risinger, *John Henry Wigmore, Johnny Lynn Old Chief and "Legitimate Moral Force": Keeping the Courtroom Safe for Heartstrings and Gore,* 49 HASTING L.J. 403, 440, n.95 (1998) (citing DANIEL GOLEMAN, EMOTIONAL INTELLIGENCE 52–54 (1995)): "Emotions, in all their shades and variety, are value-charged responses, positive or negative. The role of emotion in rational tasks is easy to underestimate, since without some emotional response to the task, there would be no reason to spend the effort to do it at all. There is even some empirical confirmation for this." *Id.*

[72] Kevin M. Clermont, *Common Sense on Standards of Proof*, 48 SETON HALL L. REV. 1057 (2018).

Last but certainly not least, I am indebted to Professor Edward Cheng and his young co-author Matthew Ginther (who is given principle author credit) for their report on the results of an empirical study designed to begin to compare the results that obtain when people are asked to estimate the probability of guilt as opposed to their level of surprise were the defendant not guilty under various hypothetical scenarios.[73]  While they are clear that the results must be viewed as preliminary and only the beginning of an exploration of the relative strengths and weaknesses of the two approaches (and the results which would obtain by virtue of an ordinary instruction were not tested), I am heartened to find that surprise estimates may display certain advantages, like avoidance of the base-rate fallacy, over estimates based on probability.  I hope there will be further empirical exploration of these issues by Ginther and/or Cheng, and by others, in the future.

And in conclusion, I want to thank all of those who participated in this symposium, and whose work made it the intellectual feast that it was, and that it will be for those who read the contents of these volumes.

---

[73]  Matthew Ginther & Edward K. Cheng, *Surprise v. Probability as a Metric for Proof*, 48 SETON HALL L. REV. 1081 (2018).