

## The Persistence of the Probabilistic Perspective

*Richard D. Friedman\**

I.INTRODUCTION .....	1589
II.OUTLINES OF A PROBABILISTIC ACCOUNT .....	1589
III.CALLEN’S CONCERNS .....	1592
IV.INFERENCE TO THE BEST EXPLANATION .....	1596

### I. INTRODUCTION

The publication now of an essay written by Craig Callen nearly a decade ago is cause for wistful celebration. Even while we are reminded how suddenly and prematurely Craig’s life ended, it is good to have one more academic contribution from him, especially because it is marked by the erudition, thoroughness, gentleness, and humor that characterized him.

Though Craig’s essay is a commentary on an approach to fact-finding presented by Michael Pardo and Ronald Allen, it discusses at some length problems that he perceived in what he called “mathematical models of inference from evidence.” I am one of those who believes that standard probability analysis can be useful in thinking about some problems in juridical fact-finding. I believe that speaking of “mathematical models” might tend to give a misleading view of the enterprise. At least for the most part, we do not advocate introducing numbers into the analysis, or calling for jurors to make complex calculations, or any calculations for that matter. But probabilistic reasoning can provide a useful heuristic for thinking of the fact-finding process. So first I will sketch the outlines of a probabilistic account of that process. Then I will respond to Craig’s concerns, some of which are expressed by Pardo and Allen as well. And then I will take the opportunity to join, belatedly, in the conversation about the Pardo and Allen approach.

### II. OUTLINES OF A PROBABILISTIC ACCOUNT

I will begin by addressing the standard of persuasion. This is the level of confidence the trier of fact must have that the facts favor claimant in order to find for that party. (For simplicity, I will assume that the trier is a jury and the claimant is either a plaintiff or a prosecutor.) Conceptually, I believe, the problem is actually rather simple. A trial requires a decision under

uncertainty, meaning that whichever decision one makes, there is a chance that the decision will be wrong. Let's focus on a simple, two-option decision—guilty or not guilty, liable or not liable. The choice of standard will depend on a relative weighing of the harm caused by one wrong decision or another. Suppose that Option One has far worse consequences if wrong than does Option Two. Then a sensible decision-maker will choose Option One rather than Option Two only if she has a high degree of confidence that Option One rather than Option Two is correct, or, put another way, only if she thinks Option One is far more probable than Option Two.

Criminal law clearly reflects this analysis. The received (and justified) wisdom over centuries, in our system and in many others, has been that an incorrect conviction is many times worse than an incorrect acquittal,<sup>1</sup> and so we tell the jurors not to enter a verdict of guilt unless they are virtually certain that the accused is in fact guilty; hence, the beyond-a-reasonable-doubt standard. In most civil cases, we say that an error favoring the defendant is about as bad as an error favoring the plaintiff. The usually accepted rule is that to find for the plaintiff, the jury need only find that the facts more likely than not favor the plaintiff; the case of exact equipoise is decided in favor of the defendant, giving some weight (which I think should probably be considerably greater) to recognition of the inertial value of maintaining the *status quo*. But in some civil cases, a pro-plaintiff error appears so much worse than a pro-defendant error that we impose an intermediate standard, usually articulated as “clear and convincing.” And, for that matter, there are some decisions in which courts will grant relief even if the balance of probabilities is *against* the party seeking it.<sup>2</sup>

Note that in describing these standards according to a probabilistic model,<sup>3</sup> I have avoided using any numbers, because they are not necessary and courts do not ordinarily use them either. But the concept of *magnitude* is essential. It is not enough to say that the harm caused by an incorrect choice of Option One is greater than that caused by an incorrect choice of Option Two; to set the standard of persuasion appropriately, we need to have a sense of *how much* worse one error is than the other.

---

\* Richard D. Friedman, Alene and Allan F. Smith Professor of Law, The University of Michigan Law School.

<sup>1</sup> See generally Alexander Volokh, *n Guilty Men*, 146 U. PA. L. REV. 173 (1997) (providing a whimsical look at the matter, gathering many statements of the principle from a wide array of sources).

<sup>2</sup> See, e.g., 18 U.S.C. § 2703(d) (2009) (“reasonable grounds to believe”).

<sup>3</sup> I do not mean to argue that other considerations do not enter into the determination of whether a case has been sufficiently proven to be presented to the jury; I am agnostic on that. See Richard D. Friedman, *Towards a (Bayesian) Convergence?*, 1 INT'L J. EVIDENCE & PROOF 348, 351 (1997); see generally DALE A. NANCE, *THE BURDENS OF PROOF: DISCRIMINATORY POWER, WEIGHT OF EVIDENCE, AND TENACITY OF BELIEF* (2016).

## 2018] PERSISTENCE OF THE PROBABILISTIC PERSPECTIVE 1591

Under this view, the job of the jury, once all the evidence is presented to it, may be thought of as asking a comparative question: “How probable is it that the evidence arose by a course of events that warrants a verdict for Party One as compared to a course of events that would warrant a verdict for Party Two?” This approach is perfectly consistent with what has come to be called the “story model” of litigation.<sup>4</sup>

So far I have spoken about what Pardo and Allen call the macro level. I’ll turn briefly, as they do, to what they call the micro level—the relevance and probative value of particular items of evidence. I think we agree that relevance and probative value are essentially the same phenomenon, but that relevance is a binary term and probative value is a matter of degree;<sup>5</sup> this is how Federal Rules of Evidence 401 and 403, respectively, use the terms. Note that Rule 401 defines relevance in explicitly probabilistic terms: Evidence is relevant if it has “any tendency to make . . . more or less probable than it would be without the evidence” a fact that is “of consequence in determining the action.”<sup>6</sup> This probabilistic orientation was not adventitious, and it reflected a deep-seated view, not the whimsy of latter-day modelers. In a 1966 comment accompanying his first draft of the Rule, which in substance has remained essentially unchanged through the drafting process and ever since, the Reporter, Edward Cleary, noted that the draft’s approach was based on a 1941 article by George F. James, *Relevancy, Probability and the Law*.<sup>7</sup> That article, Cleary said, “has been followed by most of the subsequent writers on the subject.”<sup>8</sup> Indeed, James, in explaining how A’s design to kill B is relevant to the question of A’s guilt of killing B, takes an explicitly probabilistic view: “We cannot now say that A is probably guilty, but we can say that *the apparent probability of his guilt is now greater than before the evidence of design was received*.”<sup>9</sup> And that, of course, ties in nicely with the overall structure of the trial, because the prosecutor is trying to show that, taking into account all the evidence and everything that we know material to the case, the probability of guilt is very high indeed.

---

<sup>4</sup> See Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 CARDOZO L. REV. 519 (1991). I have discussed the compatibility of probabilistic and story models at some length in *Infinite Strands, Infinitesimally Thin: Storytelling, Bayesianism, Hearsay and Other Evidence*, 14 CARDOZO L. REV. 79 (1992).

<sup>5</sup> See Michael S. Pardo & Ronald J. Allen, *Juridical Proof and the Best Explanation*, 27 LAW & PHIL. 223, 241–42 (2008).

<sup>6</sup> FED. R. EVID. 401.

<sup>7</sup> George F. James, *Relevancy, Probability and the Law*, 29 CALIF. L. REV. 689 (1941).

<sup>8</sup> RICHARD D. FRIEDMAN & JOSHUA DEAHL, FEDERAL RULES OF EVIDENCE: TEXT AND HISTORY 55 (2015).

<sup>9</sup> James, *supra* note 7, at 699 (emphasis in original).

Advocates of a probabilistic analysis of fact-finding recognize that assignment of a probability to a given proposition is a subjective matter. They also recognize that jurors do not, and should not, usually go through prescribed algorithms in analyzing evidence. Accordingly, the usefulness of probabilistic analysis depends only on the proposition that if a juror is doing her job well, she will act in accordance with certain fundamental principles. For example, all other things being equal, the probability of a proposition as assessed given a particular piece of evidence will be higher: (1) the higher is the prior probability of that proposition – that is, the probability of the proposition as assessed without the evidence; and (2) the greater is the likelihood ratio of the evidence with respect to the proposition – that is, the ratio yielded by dividing (a) the probability that the evidence would arise given the truth of the proposition, by (b) the probability that the evidence would arise given the falsity of the proposition.

### III. CALLEN'S CONCERNS

Callen raised several concerns about what he calls “mathematical models.” I believe that none of these concerns undermine the analytical usefulness of probabilistic models, at least if they are kept within the confines suggested above. I will address his points serially, though in a slightly different order than he did.

First, Callen properly emphasized the importance of soft variables.<sup>10</sup> That would indeed be a concern if one asked a jury to reach a particular numerical assessment of probability on the basis of numerical evidence. But there is no need that a probabilistic analysis asks for anything of the sort, and it can posit that jurors implicitly assign probabilities on the basis of non-numerical information just as easily as it can posit that they do so on the basis of hard data.

Second, Callen emphasized, as he has before, the problem of computational complexity.<sup>11</sup> I think there are several answers. First, any model of fact-finding *ought* to be able to handle, as a theoretical matter, all the complexity in the world. But that does not mean that jurors who are operating according to the model must use it in full force. Indeed, they need not, and various simplifications and approximations make the task considerably easier.<sup>12</sup> At the same time, commentators considering a

---

<sup>10</sup> Craig R. Callen, *Spotting A Preponderance Of The Evidence In The Wild: Inference To The Best Explanation And Sufficiency Of The Evidence*, 48 SETON HALL L. REV. 1517 (2018).

<sup>11</sup> *Id.* at 1527. I have responded at some length to this contention in *Answering the Bayesioskeptical Challenge*, 1 INT'L J. EVIDENCE & PROOF 276, 288–90 (1997).

<sup>12</sup> Peter Donnelly, *Approximation, Comparison, and Bayesian Reasoning in Juridical*

## 2018] PERSISTENCE OF THE PROBABILISTIC PERSPECTIVE 1593

problem of proof can use as much complexity in their analysis as seems helpful.<sup>13</sup>

Third, Callen pointed out that mathematical logic is reversible—that if, for example, one calculated a given probability but found that unappealing, one might reassign probabilities and recalculate “until satisfied that the result seems reasonable, and that the application of the theorem incorporates the relevant criteria.”<sup>14</sup> True enough. I am not sure that such a process of reflective equilibrium would be a problem even if one were suggesting that jurors followed an algorithm under which they assigned elemental probabilities and then, based on those assignments, assessed how probable the hypothesis of interest is in light of all the evidence; it might well make sense to take a second look and see if taken together all the assignments yield sensible results. And particularly if one only expects the jury to compare the probabilities of stories, or of groups of stories, I think the process Callen suggested poses no problem at all.

Fourth, Callen pointed to issues related to incompleteness of evidence.<sup>15</sup> I agree that jurors must take into account not only the evidence that is presented to them but also the fact that certain types of evidence are *not* presented to them. There is absolutely nothing in a probabilistic model that prevents jurors from inferring that if a given type of evidence favored the party with better access to it, that party would probably have presented it. Nor does a probabilistic model prevent jurors from making that inference part of an overall assessment of the factual issue at stake.

Fifth, Callen discussed the reference class problem—if one is trying to determine the probability that the bus that caused an accident belonged to a given company, is the most appropriate class to examine the buses in town, or those in town that got into accidents, or those that run on the street where the accident occurred, or something else?<sup>16</sup> I agree that this would be a serious problem if one expected jurors to treat a datum as a probability—for example, “Eighty percent of the buses in town are blue, and therefore I believe that the probability that this accident was caused by a blue bus is 80%.” But a datum is just a piece of information; it is not a probability

---

*Proof*, 1 INT’L J. EVIDENCE & PROOF 304, 305–07 (1997); *Answering the Bayesioskeptical Challenge*, *supra* note 11, at 288–89.

<sup>13</sup> See, e.g., Richard D. Friedman, *Route Analysis of Credibility and Hearsay*, 96 YALE L.J. 667, 707–19 (1987) (using probabilistic analysis to assess probative value of, *inter alia*, statements of intent in proving subsequent actions of another); Richard Friedman, *Character Impeachment Evidence: Psycho-bayesian [!?] Analysis and a Proposed Overhaul*, 38 UCLA L. REV. 637 (1991) (using probabilistic analysis to assess probative value of prior convictions with respect to credibility of witnesses).

<sup>14</sup> Callen, *supra* note 10, at 1527.

<sup>15</sup> Callen, *supra* note 10, at 1527.

<sup>16</sup> Callen, *supra* note 10, at 1530.

assessment, which is a statement of a level of confidence in a given proposition. There is no doubt that it can be a difficult and uncertain task to determine how to use a generalized datum in making a sensible assessment of the probability of a particularized proposition; one might well decide, for example, that even if most of the buses that run through town during the day are blue, that is not true of the particular time and place of the accident, and one might have reason to suspect that buses belonging to the Blue Bus Company are more (or less) likely than others to get into collisions. But this is a real problem that inevitably confronts anyone who must make a decision given incomplete information; it is not a conceptual problem of a probabilistic analysis of such a situation.

And finally, Callen raised the so-called problem of conjunction.<sup>17</sup> The problem is thought to be this: Suppose that an ordinary civil claim consists of multiple elements, and the jury finds both that each of them is more likely than not true and that it is more likely than not that they are not all true. What should happen in this circumstance? Those who, like Callen, Pardo and Allen, are skeptical of probabilistic models of fact-finding believe that this situation poses a conundrum for a probabilistic account. I believe the contention has been answered satisfactorily numerous times over the years.<sup>18</sup> The simple response is that triers of fact should be given a conjunctive instruction—that is, they should find for the plaintiff only if they find to the requisite level of confidence that all of the required elements are true. (A finding that the plaintiff fails to meet that standard with respect to any one of the elements would immediately warrant a finding for the defendant, because then the conjunction could not meet the standard.) Such an instruction would not conflict with any entrenched body of law; it may be that some jury instructions now given appear to call for a verdict for the plaintiff if each element meets the standard of persuasion, but there is no reason to believe that those who have crafted such instructions have carefully thought the matter through. In any event, if the current law and the prescription of the probabilistic model were in genuine conflict on this

---

<sup>17</sup> Callen, *supra* note 10, at 1529.

<sup>18</sup> The matter was addressed, completely satisfactorily in my view, by Dale Nance more than thirty years ago in *A Comment on the Supposed Paradoxes of a Mathematical Interpretation of the Logic of Trials*, 66 B.U. L. REV. 947 (1986). I have addressed it myself several times, including in *Answering the Bayesian Challenge*, 1 INT'L J. EVIDENCE & PROOF 276, 279–84 (1997). A very thorough recent treatment, which I believe is completely in accord with the arguments made by Nance and me, is David S. Schwartz & Elliott Sober, *The Conjunction Problem and the Logic of Jury Findings*, 59 WM. & MARY L. REV. 619 (2017). Another article that reaches essentially the same conclusion is Mark Spottswood, *Unraveling the Conjunction Paradox*, 15 LAW, PROBABILITY & RISK 259, 264 (2016) (arguing that factfinders should be instructed “that they should only find a defendant liable when the plaintiff has met the burden of persuasion for all of the elements *taken together as a whole*”) (emphasis in original).

## 2018] PERSISTENCE OF THE PROBABILISTIC PERSPECTIVE 1595

question, the proper conclusion would be to recognize that the law is in error and should be changed, and to cheer for an instance of theoretical analysis improving prevailing doctrine.

Callen responded by pointing out that a conjunctive instruction would raise the average probability of each element required for a plaintiff's verdict; he said that "[i]t is difficult to see why the number of elements should affect the burden of persuasion for each or why the burden should be as high as .7 given the typical instruction."<sup>19</sup> (Pardo and Allen make the same point.<sup>20</sup>) But the average required probability per element is a number of no particular significance. And note that how many elements a claim is deemed to have is simply a matter of word choice. One might think of "breach of duty" as a single element or as two elements, "existence of a duty" and "breach of that duty." The substance of the plaintiff's burden should not depend on which choice of articulation the law makes. So if we restate one element as two, it follows inevitably that the average probability per element that the plaintiff must demonstrate will increase—and that is of no concern whatsoever. Suppose, for example, a plaintiff has to show as an element of his claim a continuous period of disability over a given period, say a year. If instead we broke that element into 365 separate ones, one for each day, the substance of the claim would not change. The average probability per element required for the plaintiff to prevail would increase, but that does not matter; the jury would likely conclude that if the plaintiff was disabled on any given day then it was almost certain that he was disabled on any closely proximate day.<sup>21</sup>

---

<sup>19</sup> Callen, *supra* note 10, at 1529.

<sup>20</sup> Pardo & Allen, *supra* note 5, at 253–54.

<sup>21</sup> I discuss this point in considerable detail in *Infinite Strands*, *supra* note 4, at 97–98 n.50. Strangely, Pardo and Allen assert that under a conjunctive instruction, "[p]laintiffs' likelihood of success will depend on how the claim is defined." Pardo & Allen, *supra* note 5, at 255. They may be making two separate points. First, if the same substance is broken into more elements, the average probability per element required for the plaintiff will increase. As I have already indicated in the text, that is of no significance. (I am not certain, but they may recognize this.) Second, the average probability per element required for the plaintiff will also increase if an additional element is added to the legal requirements of the claim, so that the plaintiff's task is made substantively more difficult. But that is just as it should be. Compare a pair of simple card games. In Game 1, there is a single pile of shuffled cards (not necessarily a full deck), and your job is to call "Red" if and only if you believe that a card chosen at random from the pile is more likely than not to be red. You should make the call if and only if you believe that there are more red cards than black cards in the pile. In Game 2, there are two shuffled piles, drawn from separate decks, and you should call "Red" if and only if you believe that it is more likely than not that two cards, one drawn at random from each pile, will *both* be red. Obviously, now you should make the call only if you believe that the average probability of a red card being drawn from each pile is substantially greater than .5; to be more precise, the average must be greater than 2/2, or about .707. That could occur if you believed that each pile had a proportion of red cards greater than that, or if you knew that one pile was all red and the other a little more red than black, or any combination in between.

Ultimately, Callen did appear to recognize that an instruction might require, for the plaintiff to win a verdict, only that the jury find the probability of “all of the elements of the claim” to be greater than .5. But, he said, prescribing this means only that “mathematical models . . . cannot be used for anything more detailed than a global assessment of the evidence.”<sup>22</sup> The conclusion does not follow. All the jury needs to find, and all it should be instructed to find, is whether the probability of the conjunction of the elements is greater than the prescribed standard of persuasion. But that does not limit what the probabilistic analysis can be used for. One can, if it seems useful for any purpose, use the probabilistic approach to analyze the probability of a given element or the effect on the whole case, or on an element, of a given piece of evidence.

In short, I do not find that any of Callen’s arguments should make us hesitate to use a probabilistic model of the fact-finding process, or to use probabilistic reasoning in attempting to make careful analysis of particular problems of proof.

#### IV. INFERENCE TO THE BEST EXPLANATION

The approach presented by Pardo and Allen—labeled “PAT” by Callen—is an application to the juridical context of the practice known in philosophy as inference to the best explanation, or IBE. As applied by Pardo and Allen to proof at trial, IBE proceeds in two stages. First is generation of competing explanations for the evidence presented at trial; this is primarily the work of the parties, but Pardo and Allen acknowledge that fact-finders are not limited to the explanations put forward by the parties.<sup>23</sup> Second is determination of which party the best explanation favors.<sup>24</sup>

---

There is nothing strange about this.

Pardo and Allen assert that “counterintuitive results pose a serious challenge to the Bayesian approach.” Pardo & Allen, *supra* note 5, at 255. But they have failed to present any results that are actually counterintuitive. For a fuller discussion of what Allen apparently thinks are “weird” results, see *Answering the Bayesioskeptical Challenge*, *supra* note 11, at 281–82.

<sup>22</sup> Callen, *supra* note 10, at 1530.

<sup>23</sup> Pardo & Allen, *supra* note 5, at 234. This appears to me to be a welcome change of view from an earlier work of Allen’s, in which he said trials should be conceived of “as comparing the probability of the fully specified case of the plaintiff to the probability of the equally well specified case of the defendant.” Ronald J. Allen, *A Reconceptualization of Civil Trials*, 66 B.U. L. REV. 401, 425 (1986).

<sup>24</sup> There is ambiguity in speaking of *the* best explanation; if each of multiple accounts differs in some details but under the governing law each would lead to the same result, should they be considered to be separate explanations or just alternative parts of the same explanation? Allowing a party to benefit from the aggregation of such accounts makes sense, and squares with the probabilistic account I have outlined, but it seems in tension with the idea of identifying *the* best explanation or, as Allen has previously articulated the parties’ responsibility, see Allen, *supra* note 23, at 425, presenting “fully specified” accounts.

## 2018] PERSISTENCE OF THE PROBABILISTIC PERSPECTIVE 1597

What makes one explanation better than another? Pardo and Allen say there are “objective criteria” for evaluating explanations: “For example, of two explanations, all other things being equal, the one that explains more of the evidence will be better than the one that explains less; the one that is more in accord with what else we know will be better than the one that is not; and so on.”<sup>25</sup> I do not find this satisfying. Even if one had a good metric for determining what it means to say that one account explains more of the evidence than does another (or, what I think would be more comprehensible, to say that one account *better* explains the evidence than does another<sup>26</sup>) or that one is more in accord with what else we know than another,<sup>27</sup> and even if we knew how to put together these and other, unidentified, assertedly “objective criteria” for determining which of two explanations is better, we should ask what makes these particular criteria matter to the judicial system. It seems to me that the answer is that these factors tend to make one account *more probable* than the other.

Pardo and Allen might seem at one point to acknowledge this, because they suggest that in the ordinary civil case, the best explanation is “the most *plausible* version of the litigated events.”<sup>28</sup> Actually, though, they are determined to show that cardinal probabilities are not a useful basis for thinking about the juridical fact-finding process.<sup>29</sup> The standard of persuasion as usually articulated in the ordinary civil case shrouds a problem, because if one takes seriously the idea that the plaintiff should win if the balance of probabilities tilts at all in his favor, by no matter how small an amount, then one need not worry about magnitude; for a verdict in favor of the plaintiff to be justified, it is enough that the pro-plaintiff account be just

---

Accordingly, I have been perplexed in the past as to whether, or the extent to which, Allen would allow aggregation of such accounts. See, e.g., *Infinite Strands*, *supra* note 4, at 93–94 n.40. Allen has previously criticized “bunching” as part of a Bayesian explanation of fact-finding. E.g., Ronald J. Allen & Brian Leiter, *Naturalized Epistemology and the Law of Evidence*, 87 VA. L. REV. 1491, 1507 (2001). But now Pardo and Allen appear to be fully on board with the idea of aggregating factually different accounts among which “the substantive law is indifferent.” Pardo & Allen, *supra* note 5, at 236–37. So I am glad to see this apparent disagreement removed. And the removal is, I believe one factor among several that brings Pardo and Allen’s IBE approach closer to the probabilistic one.

<sup>25</sup> Pardo & Allen, *supra* note 5, at 245.

<sup>26</sup> As I have suggested above, it seems to me that to have a chance of persuading the jury, any given theory of the facts of the case must purport to account for *all* of the evidence. Does an asserted eyewitness testify that the accused committed the crime? Perhaps that testimony is perjured. Does DNA evidence appear to demonstrate that the accused was the source of a blood stain found at the scene of the crime? Perhaps samples were mixed up in the lab. Some explanations, of course, will be more persuasive than others.

<sup>27</sup> I suppose my standard for determining that one account is more in accord with what else we know than is another would be that, given everything we know, the first account is more probable than is the other. But I doubt that is what Pardo and Allen mean.

<sup>28</sup> Pardo & Allen, *supra* note 5, at 266 (emphasis added).

<sup>29</sup> Pardo & Allen, *supra* note 5, at 261.

slightly more probable than the pro-defendant account.<sup>30</sup>

But outside that context, the shroud does not hold. In the criminal context, Pardo and Allen contend that juries should be instructed according to such formulations as “whether there is a plausible explanation or version of events consistent with innocence.”<sup>31</sup> These formulations seem perfectly reasonable to me—but, as Pardo and Allen acknowledge, they do not reflect a process of inference to the best explanation.<sup>32</sup> So the IBE approach, it appears, suffers itself from a lack of parsimony of explanation—the theory does not apply, at least without substantial adjustment, outside the context of the ordinary civil case. And why *should* a different standard—and in their view, a different approach to selecting a standard—apply in the criminal case? Pardo and Allen address this question only fleetingly and in a footnote, but what they say is intriguing:

IBE is, at root, based on the notion that explanatory success tracks likelihood of truth—the better the explanation, the more likely true. Because the criminal standard distributes errors unevenly (in favor of the defendant), it should not be surprising that the quality of the explanation needed for a pro-defendant verdict should therefore be lower.<sup>33</sup>

Well, wait a minute. So the ultimate aim appears to be assessing “likelihood of truth,” and in cardinal terms. That sounds pretty much like probability. And why is a relatively low probability of innocence sufficient to justify a verdict for the defendant? Because, they say, our system has decided to distribute errors unevenly in favor of the defendant—which I take as intending to say that our system regards a pro-defendant error as of much lower magnitude harm than a pro-prosecution error.<sup>34</sup> And if that is *not* why, what other plausible reason is there for setting such a pro-defendant standard of persuasion? Pardo and Allen offer none. So we see a large differential in the magnitude of types of error together with a resultant low probability for

---

<sup>30</sup> In speaking of the pro-plaintiff and pro-defendant accounts, I am implicitly assuming aggregation of all pro-plaintiff accounts and of all pro-defendant accounts. *See supra* note 24.

<sup>31</sup> Pardo & Allen, *supra* note 5, at 267.

<sup>32</sup> Pardo & Allen, *supra* note 5, at 239 n.45.

<sup>33</sup> Pardo & Allen, *supra* note 5, at 239 n.45.

<sup>34</sup> I do not think their articulation is quite accurate. The aim of our system should not be to create any particular distribution of errors; that distribution will depend in part on the population of cases. The aim instead is to minimize the total cost of errors, and a way to think of doing that is to move in a pro-defendant direction until the marginal cost of pro-defendant errors caused equals the marginal benefit of pro-prosecution errors avoided. In one private conversation, Allen told me this point was gibberish; in a later communication, he said that it was correct and that he had made it before. Whatever.

## 2018] PERSISTENCE OF THE PROBABILISTIC PERSPECTIVE 1599

a standard of persuasion. It's beginning to look a lot like a probabilistic account, everywhere you go.

The apparent congruence with a probabilistic model becomes even more pronounced when one considers intermediate standards such as “clear and convincing.” Pardo and Allen acknowledge that the jury should not find in favor of a party bearing such a burden merely because the best explanation among available ones favors that party. Instead, the explanation favoring that party must be “sufficiently more plausible” than others—that is, “clearly and convincingly more plausible than those favoring the other side.”<sup>35</sup> And this is not circular, they say, because it simply reflects the nature of the “clear and convincing” standard; the explanation must be “good enough to cause and justify the desired inference.”<sup>36</sup> That appears to be a clear appeal to cardinality—to degrees of “plausibility,” most easily understood as ordinary probability.<sup>37</sup> So the proposed standard seems to be something close to “clearly more probable than others.” And why *should* our system adopt such a standard in a given type of case? Presumably the reason is that an error on one side is considered substantially worse than an error on the other side (but not by as great a margin as in a criminal case); if they have another explanation, Pardo & Allen do not offer it.

Now consider Pardo and Allen's treatment of individual items of evidence. They quote in a footnote, but then ignore, the “more probable or less probable” language of Rule 401.<sup>38</sup> In their view, evidence is relevant “if it is explained by” a given explanation and “in turn justifies that explanation as correct, assuming the explanation concerns a fact that matters to the substantive law.”<sup>39</sup> But what does that mean? I do not think it is consistently useful to say that a given account “explains” a piece of evidence, and a single piece of evidence does not in itself usually “justify” a given account. For example, if a witness says that the assailant was about 6'2", I believe it is rather strained to say that the hypothesis that the accused was the assailant *explains* evidence that the accused is 6'2", and it seems to be an overstatement to say that the evidence *justifies* the hypothesis. But it is perfectly comprehensible and sensible to say that the evidence makes the hypothesis *more probable* than it was before the evidence. And that, of course, ties in nicely with the overall structure of the trial, because the prosecutor is trying to show that—taking into account all the evidence and

---

<sup>35</sup> Pardo & Allen, *supra* note 5, at 239–40.

<sup>36</sup> Pardo & Allen, *supra* note 5, at 240 n.46.

<sup>37</sup> See Richard D. Friedman, “E” Is for Eclectic: Multiple Perspectives on Evidence, 87 VA. L. REV. 2029, 2047 (2001).

<sup>38</sup> Pardo & Allen, *supra* note 5, at 242 n.53. Since Pardo and Allen published their article, this language has been amended for style, by deleting the first “probable.”

<sup>39</sup> Pardo & Allen, *supra* note 5, at 241.

everything we know that is material to the case—the probability of guilt is very high indeed. I do not believe that putting the relationship between the evidence and the hypothesis in the terms that Pardo and Allen favor offers anything useful that is not already achieved by speaking in probabilistic terms.

\* \* \*

There is much more that can be said, but this is supposed to be a brief essay. Silence on other points does not suggest assent. I do not believe that Callen's piece undermined the usefulness of a probabilistic model of fact-finding, and I am not sure that Pardo and Allen's invocation of IBE adds anything useful that is not already captured by such a model. But this is an ongoing conversation, and I suppose it will continue; I am just sorry that Craig Callen's voice will no longer be a part of it.