CHALLENGES FOR COMPARATIVE FACT-FINDING

Sean P. Sullivan[*]

September 15, 2018

I often wish … that I could rid the world of the tyranny of facts. What are facts but compromises? A fact merely marks the point where we have agreed to let investigation cease. – Anonymous[1]

I agree with Allen and Pardo that a paradigm shift is underway in scholarship on legal fact-finding.[2] So much recent work points the same direction—that persuasion is the product of purely comparative assessments of factual propositions—that those unable to perceive this shift could only be those who refuse to see.[3] I worry, though, that some second-order debates risk obscuring this growing consensus.

---

[*] University of Iowa College of Law. Contact the author at sean-sullivan@uiowa.edu.

[1] *On Having Known a Poet*, 97 THE ATLANTIC MONTHLY 711, 712 (1906).

[2] *See* Ronald J. Allen & Michael S. Pardo, Relative Plausibility and Its Critics 1, 1, available at https://ssrn.com/abstract=3179601.

[3] Proposals to approach fact-finding in comparative terms include many qualitative models of fact-finding. *See, e.g.*, Michael S. Pardo & Ronald J. Allen, *Juridicial Proof and the Best Explanation*, 27 L. & PHIL. 223 (2008) (comparing factual theories in terms of their explanatory power); Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511 (2004) (comparing factual theories in terms of the concept of coherence); Ronald J. Allen, *Factual Ambiguity and a Theory of Evidence*, 88 NW. U. L. REV. 604 (1994) (comparing factual theories in terms of relative plausibility); Nancy Pennington & Reid Hastie, *The Story Model for Juror Decision Making*, *in* INSIDE THE JUROR: THE PSYCHOLOGY OF JUROR DECISION MAKING 192 (Reid Hastie ed., 1993) (comparing factual stories in terms of their psychological salience). Other proposals include formal models of comparative fact-finding. *See, e.g.*, Sean P. Sullivan, *A Likelihood Story: The Theory of Legal Fact-Finding*, 90 COLO. L. REV. (forthcoming) (comparing factual theories in terms of likelihoods); Kevin M. Clermont, *Trial by Traditional Probability, Relative Plausibility, or Belief Function?*, 66 CASE. W. RES. L. REV. 353 (2015) (comparing factual theories in terms of belief functions); Edward K. Cheng, *Reconceptualizing the Standard of Proof*, 122 YALE L.J. 1254 (2013) (comparing factual theories in terms of probabilities).

I do not, for example, perceive any antagonism between formal and qualitative models of the fact-finding process. As I have written elsewhere, I agree with Allen and Pardo that Bayesian probability concepts are often inappropriate as a model of legal fact-finding.[4] But that is because probability concepts lend themselves to absolutist or propositional reasoning, not because of any inherent incompatibility between formalized concepts of uncertainty and the objective of trying to better understand the legal system.[5]

To the extent that proponents of formal models are proposing fully quantitative *substitutes* for the type of reasoning that Allen, Pardo, and others have described,[6] I agree that the formal approach lacks the depth and nuance of actual trial fact-finding.[7] But to the extent that formal models are merely a tool of abstraction—used to think clearly and logically about a problem—I fail to see the urgency for distinguishing them from less formal models. Thus, while I accept that subjective Bayesian probabilities are not drop-in substitutes for the epistemic credentials at focus in the relative plausibility model,[8] I cannot imagine why it would be bad if formal credence concepts *could* quantify relative plausibilities. If formal theory and something closer to empirical description were to align in description of closely similar models of fact-finding, that would strike me as all to the good. A true paradigm shift, I would think, *should* be evident at *every* strata of analysis.[9]

---

[4] *See* Sullivan, *supra* note 3, at § IV.

[5] *Cf.* People v. Collins, 68 Cal. 2d 319, 438 P.2d 33 (1968) (famously commenting that "we discern no inherent incompatibility between the disciplines of law and mathematics and intend no general disapproval or disparagement of the latter as an auxiliary in the fact-finding processes of the former").

[6] *See generally* Allen & Pardo, *supra* note 2; Allen, *supra* note 3; Pennington & Hastie, *supra* note 3.

[7] I take this to be the point of explaining that relative plausibility does not inherently involve numeric reasoning. Allen & Pardo, *supra* note 2, at 13, 17.

[8] *Cf. id.* at 10 (rightly noting the vagaries of prior probabilities, but also appearing to suggest that likelihood ratios needn't have any correspondence to the probative value of evidence).

[9] *See supra* note 3 (observing just such a trend).

And it is. At every level of research, from the flightily formal to the grittily empirical, an unyielding shift in understanding is taking place: moving progressively away from absolutist or propositional concepts of what it means to find a fact, and progressively toward comparative definitions of facts as the most plausible (least rejected) alternative among the possibilities in consideration.[10]

This comparative approach is a departure from what long appeared to be settled legal thinking. For decades, conventional probability models posited fixed thresholds which a given proposition had to cross before it had the credentials of a legal *fact*.[11] General notions of absolute, propositional fact-finding are even older. Thus, common articulations of the preponderance standard instruct jurors that this burden has been met when they are "persuaded that [the asserted fact] is more probably true than not true."[12] Even greater propositional certainty is demanded in the requirement that prosecutors prove the essential facts of a crime to be true "beyond a reasonable doubt."[13] There is undeniable appeal to the idea that *facts* should at least be *probably true* before they suffice to establish a claim or defense,[14] and devotion to this view of fact-finding is consequently deep-rooted.

It is also deeply flawed. Allen and Pardo critique the absolutist view as inconsistent with "how proof proceeds at trial,"[15] as inconsistent with "the goals underlying the standards of proof,"[16] and as inconsistent with

---

[10] Again, see *supra* note 3.

[11] *See* Allen & Pardo, *supra* note 2, at 8–9 (summarizing important works on the probability approach).

[12] 7th Cir. Fed. Civ. Jury Instructions § 1.27 (2017), http://www.ca7.uscourts.gov/pattern-jury-instructions/7th_cir_civil_instructions.pdf.

[13] *See* In re Winship, 397 U.S. 358, 364 (1970) (holding "that the Due Process Clause protects the accused against conviction except upon proof beyond a reasonable doubt of every fact necessary to constitute the crime with which he is charged.").

[14] *See, e.g.*, J. P. McBaine, *Burden of Proof: Degrees of Belief*, 32 CAL. L. REV. 242, 248–49 (1944) (noting the possibility of defining the preponderance standard in comparative terms, but arguing it would be unwise not to at least require facts favoring the plaintiff be shown more probable than not).

[15] Allen & Pardo, *supra* note 2, at 12.

[16] *Id.*

how jurors actually evaluate factual propositions.[17] I agree on all counts. But I would push the critique even further.

In law, as in life, it is a mistake to set impossible goals. Nearly a century ago, Karl Popper made the logical argument that theories cannot be proven true; they can only be falsified.[18] What's worse, in most cases even falsification is impossible, and all that can be done is to show one hypothesis more refuted than another.[19] This closes the door to all but comparative inference, and the hope that observable evidence will show one proposition more *truthlike* than its alternatives.[20]

This is bitter medicine for a system still clinging to the idea that evidence could prove a factual proposition *true* to any degree of certainty. But it is hardly fatal to the larger project. Science embraces the comparative paradigm—scientists never prove their hypotheses true, but only disprove competing hypotheses—yet no one disputes that great strides of scientific understanding have been made.[21] The same goes for statistical inference. Statisticians cannot prove any hypothesis true in isolation; all that can be done is to show that the data falsify one hypothesis (or group of hypotheses) more than another. Yet because the least refuted hypothesis is also the most confirmed,[22] comparative statistical inference has produced concrete results as well.

Trial fact-finding is a different animal in many respects, but nothing about it enables the impossible. Evidence cannot prove isolated factual

---

[17] Michael S. Pardo & Ronald J. Allen, *Juridicial Proof and the Best Explanation*, 27 LAW & PHIL. 223, 232 (2008) ("Explanations rarely explain why A; they explain why A rather than B.").

[18] *See generally* KARL POPPER, THE LOGIC OF SCIENTIFIC DISCOVERY (1956).

[19] *See* Mark L. Taper & Subhash R. Lele, *Evidence, Evidence Functions, and Error Probabilities, in* 7 HANDBOOK OF THE PHILOSOPHY OF SCIENCE 514, 515 (Prasanta S. Bandyopadhyay & Malcolm R. Forster eds., 2011) ("Another difficulty with the falsificationist approach is the fact that not only can you not prove hypotheses, you can't disprove them.").

[20] *See id.* at 515 ("Popper … was the first to realize that although all theories are false, some might be more truthlike than others").

[21] John R. Platt, *Strong Inference*, 146 SCI. 347, 347 (1964) ("Any conclusion that is not an exclusion is insecure and must be rechecked.").

[22] Taper & Lele, *supra* note 19, at 517 ("[T]he falsificationist paradigm is really a paradigm of relative confirmation—the hypothesis least refuted is most confirmed.").

propositions absolutely true to any degree of certainty. It can only show some propositions comparatively more truthlike than others.

That may be enough. Research on comparative fact-finding is rapidly developing a sensible and coherent framework for understanding how persuasion works in a trial setting. These purely comparative models of fact-finding brush past venerable paradoxes and align theoretical description with intuition and practice experience in ways that probability-based models of fact-finding never did, and never could.[23]

But there is still more work to be done. In the following pages, I outline two ongoing challenges for the comparative paradigm, using Allen and Pardo's theory to illustrate. The first is the challenge of articulating the highest burden of persuasion in purely comparative terms. The second is the challenge of understanding what it means for fact-finders to weigh, and possibly accept, disjunctive or unspecific factual claims.

I.   BEYOND A (RELATIVELY) REASONABLE DOUBT

An initial challenge for any understanding of legal fact-finding is to interpret and express common burdens of persuasion in terms intrinsic to the model.[24] For the comparative paradigm, this means articulating purely comparative versions of all the usual fact-finding standards. The task has proven manageable for the preponderance standard, and also for the clear-and-convincing-evidence standard. But proof beyond a reasonable doubt resists easy translation.

To start with the good news, the preponderance standard is readily interpreted in comparative terms. At the theoretical level, I have argued that the preponderance standard can be understood as keyed to the most likely factual hypothesis on the available evidence.[25] Generalized

---

[23] *See, e.g.*, Sullivan, *supra* note 2, at § IV (discussing problems and paradoxes with the conventional probability-based models of fact-finding); Brian Leiter & Ronald J. Allen, *Naturalized Epistemology and the Law of Evidence*, 87 VA. L. REV. 1491, 1503–10 (2001) (similarly discussing problems with probability-based theories).

[24] *See* McBaine, *supra* note 14, at 242, 244–45 (commenting on the importance of clear and consistent articulations of the burdens of persuasion).

[25] Sullivan, *supra* note 3, at §§ III.B.1, III.C.

to other concepts of evidential support,[26] the idea is that the party with the burden of persuasion should only prevail on a claim or defense if the most supported factual hypothesis favoring that party has greater evidential support than the most supported factual hypothesis favoring the unburdened party.

The above theoretical description is dense, but Allen and Pardo provide a persuasive account of how it unfolds in practice, with evidential support defined by the explanatory power of holistic factual narratives for either side of a case.[27] I generally accept that Allen and Pardo's account is how the preponderance standard does, and should, operate in practice, and my sense is that most others agree as well.[28]

The intermediate clear-and-convincing evidence standard also fits the comparative model of fact-finding. At the theoretical level, I have argued that the difference from the preponderance test is that merely having the most likely account of the facts does not entitle the burdened party to a claim or defense under the intermediate standard. Instead, meeting this heightened burden requires the production of evidence that *strongly* supports the burdened party's factual hypothesis over the opposing party's alternative.[29]

Again, Allen and Pardo's account aligns with this idea, and describes how the intermediate standard does, and should, work in practice.[30] There are still important details left unspecified at both the formal and informal levels: for example, how great must the plausibility differential be before the burdened party has met the standard? But the intermediate burden has never been a model of clarity,[31] and imprecision is inevitable for the time being.

---

[26] *See generally* Taper & Lele, *supra* note 19 (describing generalized evidence functions); Clermont, *supra* note 3, at § I (describing generalized belief functions).

[27] *See* Allen & Pardo, *supra* note 2, at 14–15 (describing the preponderance standard under the relative plausibility account).

[28] *See id.* at 29 n.153 (collecting citations).

[29] Sullivan, *supra* note 3, at §§ III.B.2, III.C.

[30] Allen & Pardo, *supra* note 2, at 16 ("[T]he plaintiff's explanation must be not only better than the defendant's but … clearly more plausible than the defendant's.").

[31] *See id.*, at 16 n.77 (noting ambiguity in the intermediate standard).

Now, what about the highest standard—beyond a reasonable doubt? Again, comparative models of fact-finding require purely comparative explanations of how this standard is met. I have proposed such a test as a theoretical matter: a further refinement of the strength of evidence requirement of the intermediate standard. The test is keyed to the idea that if a strong enough comparative showing is demanded to convict, then the risk of erroneous conviction will be mitigated to some degree.[32] I stand behind this idea but concede that is suffers two disadvantages. First, the connection between strength of evidence and risk of error is complicated and indirect.[33] Second, there is room to debate whether the positive legal standard—proof *beyond a reasonable doubt*—articulates this type of purely comparative test today.

I wish that Allen and Pardo's descriptive account supplied a better translation between the comparative approach and the current legal standard, but the relative plausibility model actually appears to depart from the comparative framework on this count. At least as I understand it, Allen and Pardo's version of the reasonable doubt standard works as follows: the prosecution only wins if it asserts a factual proposition that is plausible on the evidence *and* if no plausible alternative proposition favors the defendant.[34] This collides with the comparative paradigm in two respects: (1) it evaluates propositions according to what appears to be an absolute credence status, *plausible*, and (2) it accords dispositive weight to any *plausible* defendant-favoring proposition.

The first of these problems is easy to state. If *plausible* status attaches to a factual proposition without reference to any alternative proposition, then it constitutes a species of propositional proof at odds with the comparative paradigm. This definition of what it means to be *plausible* is analogous to the probability-threshold idea of fact-finding. It would also seem to face the same impossibility razor: just as an isolated factual

---

[32] Sullivan, *supra* note 3, at §§ III.B.3, III.C.

[33] *See id.* at § III.B.3.

[34] *See* Allen & Pardo, *supra* note 2, at 30 ("The [beyond a reasonable doubt] standard is met when there is a plausible explanation consistent with guilt and no plausible explanation consistent with innocence."); *id.* at 15 (similar); Pardo & Allen, *supra* note 17, at 238–39 (similar).

theory cannot be proven true, I struggle to see how an isolated factual theory could be proven *plausible* in more than the trivial sense of not being definitively falsified by the evidence. And because definitive falsification is often practically impossible as well,[35] even this trivial sense of plausibility is little help to the theory.

A fair response to that critique may be that *plausible*, as the term is used here, is itself a relative concept: something like "not too strongly out-supported by an alternative factual proposition." This is consistent with Allen and Pardo's broader commitment to comparative analysis, and with their description of the explanatory process.[36] But aspects of their articulation of the reasonable doubt standard make it difficult to see a way clear to this comparative interpretation:

> … even when there is no defense explanation proffered (*or constructed by fact-finders*), the standard is not met if the fact-finder concludes the prosecution's explanation is not plausible (*regardless of whether it is better than the defendant's*, which may be none at all.)[37]

This leads to the second and related point. If *plausible* status attaches to a factual proposition regardless of what the alternatives may be, then the dispositive role that a plausible defendant-favoring proposition has in this framework means that the beyond-a-reasonable-doubt standard is comparative in only a limited and technical sense. One might say that the standard is comparative in the sense that it involves "considering and comparing the possible alternative explanations on each side."[38] But no comparison is necessary if the defendant is able to put forth any *plausible* defensive theory. In this sense, the test as articulated does not satisfy the comparative paradigm.

---

[35] *See supra* note 19 and accompanying text (asserting this impossibility).

[36] *See* Allen & Pardo, *supra* note 2, at 12–13 (arguing for the need for comparative reasoning); Pardo & Allen, *supra* note 17, at 27 (commenting that explanations are usually comparative by nature).

[37] Allen & Pardo, *supra* note 2, at 30 (emphasis added).

[38] Allen & Pardo, *supra* note 2, at 31-32.

The point, here, is not to single out an issue with relative plausibility theory. Articulating a truly comparative account of the highest burden of persuasion is a challenge for any comparative model of fact-finding.[39] The problem, which perhaps falls heavier on Allen and Pardo's descriptive project than it does on more theoretical endeavors, is that decades of scholarship and judicial writing have sought to describe this burden in absolutist, propositional terms.[40] Whether existing language can be molded into a truly comparative articulation of the test is an open and important question for the comparative paradigm.

II. COMPARING ALTERNATIVES WITHIN ALTERNATIVES

Another vexing wrinkle is current lack of clarity on how fact-finders assess the evidential weight of propositions that themselves encompass alternative propositions. Examples of this type of proposition include negative claims ("not what the plaintiff said") and disjunctive or unspecific factual claims ("my injury was caused by something the defendant did"). For the purely comparative paradigm, the challenge is to explain how comparison operates when the alternatives in question encompass alternatives themselves.

An example helps to illustrate the point. Suppose a negligence action arises from D's car exiting the road and striking P on a suburban sidewalk. The issue at trial is negligence. P has no first-hand knowledge of what D was doing but argues that the jury can infer negligence from the circumstances:[41] for example, the jury is invited to consider whether D might have been daydreaming at the time of the collision, or whether D might have been reading a text message. In response, D claims prudence behind the wheel, and attributes the accident to wet road conditions and D's reasonable effort to avoid a large, unmarked pothole. In evaluating these competing narratives, how is the fact-finder supposed to assess the plausibility of P's unspecific factual theory?

---

[39] *See, e.g.*, Clermont, *supra* note 3, at 376–77 (similarly incorporating absolutist concepts into an articulation of this standard).
[40] *See, e.g.*, Allen & Pardo, *supra* note 2, at § I.B (summarizing probability models); Sullivan, *supra* note 3, at § III.B.3 (discussing qualitative models).
[41] Specifically, P invites the jury to make a *res ipsa loquitur* inference.

Taking the relative plausibility model as an example, what would it mean to conclude that P's *unspecified* factual proposition provides a better *explanation* of the evidence than D's specific factual proposition? What is the explanatory content of an amorphous set of disjoint possibilities? Allen and Pardo clearly contemplate that unspecific assertions can have explanatory weight.[42] But what defines this weight?

Logically, it seems like the explanatory content of an unspecified factual proposition must depend on the explanatory power of the individual factual propositions it contains.[43] That may be true, but it doesn't advance the answer that much. To see why, suppose the three theories described above are the only possibilities that the fact-finder can imagine, and suppose further that the explanatory power of each of these three theories can be represented by ordinal values: 3 for daydreaming, 5 for texting, 7 for the diligent-driving accident. Larger numbers indicate greater explanatory power.[44] Does P win or lose?

It isn't clear. If the fact-finder operates in purely comparative terms, conducting two pairwise comparisons (diligent-driving vs daydreaming and diligent-driving vs texting), then D wins. The diligent-driving proposition is a better explanation than either of the individual propositions favoring P. But is this the right way to handle the disjoint set of propositions advanced by P? One might reason that P should win because the sum of the two negligence theories (3+5=8) is a larger number than the diligent-driving theory (7),[45] but nothing so far has indicated

---

[42] *See, e.g.*, Allen & Pardo, *supra* note 2, at 15 n.75 ("[J]urors are free to reject all of proffered explanations and simply conclude 'something else must have happened,' without needing to formulate a specific alternative."); Pardo & Allen, *supra* note 17, at 236–37 ("[T]he doctrine of *res ipsa loquitur* allows plaintiffs to recover even by offering explanations such as 'My injuries were caused by something done by the defendant' when such a theory provides the best explanation of the evidence.").

[43] *See* Allen & Pardo, *supra* note 2 (seeming to suggest this by noting that human fact-finders will evaluate unspecified theories, in part, by reference to "other possible explanations that [their] knowledge suggests").

[44] These numbers don't mean anything outside of indicating which theory would win in a pairwise comparison. The daydreaming theory is less plausible than the texting theory, for example, as indicated by the value 3 being less than the value 5.

[45] *Cf.* Allen & Pardo, *supra* note 2, at 27 (suggesting this type of addition of disjunctive theories in the different context of theories with assumed probability values).

that explanatory values can be added in this way. What would it mean to aggregate explanations? If addition makes sense, is there any reason why each proposition should be given equal weight in the addition? If not, what weight should be assigned to each of the propositions? Taken to the extreme, this path leads back to Bayesian probabilities and thus to non-comparative reasoning about factual propositions.[46]

This is not to suggest that fact-finders cannot evaluate unspecified or disjunctive propositions.[47] They certainly can. It is also not to suggest that allowing for the evaluation of disjunctive propositions necessarily causes relative plausibility theory to converge with Bayesian analysis.[48] It doesn't. The suggestion, here, is simply that we do not currently have a clear idea *how* fact-finders do, or should, evaluate these types of propositions in a comparative framework. A fallback to Bayesian probability analysis is one possibility, but I hope and suspect that further research on comparative fact-finding can do better than that.

Again, these challenges are not unique to relative plausibility theory. Complications with the evaluation of unspecified and disjoint factual propositions arise not from the explanatory account, but from the comparative paradigm itself. The further that comparative theories venture away from propositional probability concepts, the more difficult it becomes to understand how unspecific and disjoint claims fit within the framework. This is a serious challenge, but also a valuable opportunity. The ease with which probability concepts aggregate alternatives has long suppressed serious thinking about this issue. And that means that in addressing this challenge, future research will not only be improving comparative models, but actually charting entirely new territory in our understanding of legal fact-finding.

---

[46] *See id.* at 10 (discussing problems with the use of Bayesian probability concepts in fact-finding); Sullivan, *supra* note 3, at §§ IV.C–D (same).
[47] *See* Allen & Pardo, *supra* note 2, at 25 and n.132 (discussing such a critique).
[48] *Id.*