University of Southern California Law School

Legal Studies Working Paper Series

Thin Empirics, Comment on Allen & Pardo: Relative Plausibility and Its Critics

Dan Simon[*]

[*]USC Gould School of Law, dsimon@law.usc.edu

Thin Empirics, Comment on Allen & Pardo: Relative Plausibility and Its Critics

Dan Simon

Abstract

In the target article of this symposium, Ron Allen and Michael Pardo advance the empirical claim that Relative Plausibility is the best account of juridical proof. While I tend to agree with the relative plausibility approach and endorse its holistic underpinnings, the article suffers from three weaknesses. First, the authors fail to substantiate their empirical claim. Second, the authors cite too casually to the Story Model. For all its brilliance, the story model provides too narrow a basis to serve as a general model of legal fact-finding. Finally, the authors fail to appreciate the adverse effects of holistic cognition on legal fact-finding.

Thin Empirics

Comment on Allen & Pardo: Relative Plausibility and Its Critics

In press: International Journal of Evidence and Proof

Dan Simon[†]

This comment starts on a high note. In the symposium's target article, Ron Allen and Michael Pardo set out to explore the fundamental features of legal fact-finding. Specifically, they compare the erstwhile dominant *probabilistic* account, by which evidentiary strength is determined by serially updating the probabilistic values of the constituent pieces of evidence, against the author's preferred *relative plausibility* account, by which evidentiary strength derives from its explanatory power. Notably, the authors anchor their argument in empirical grounds: "The primary disagreement, in other words, is empirical: what is the best explanation of juridical proof." (p.3). Given that evidence scholarship has traditionally been grounded in normative, mathematical, and doctrinal approaches, the turn to empirics is plainly heartening (at least for readers who are partial to an empirical take on legal analysis). The tune is brightened further (again, for readers with compatible proclivities) by the fact that the empirical claim focuses on psychological processing by fact-finders, and even further by the fact that core insights are grounded in a holistic view of human cognition (more on this partiality below). Thus, my overarching reaction to the target article is to re-endorse the core insights of relative plausibility framework,[1] to welcome the empirical approach and to laud the authors for focusing on holistic processing of information.

The seasoned reader might suspect that reaching a crescendo in the first paragraph of a symposium commentary portends wrangling down the road. Indeed, it does. Notwithstanding my fundamental agreement with the core of Allen and Pardo's conclusions, I find myself having a difficult time accepting two key components of their argument: both the proposed empirical inquiry and their reckoning with its implications feel thin.

We start with the target article's empirical claim. While it would be unfair to expect that a legal analysis of this nature meet the scrutinous standards of a scientific examination, any empirical pursuit must clearly identify the unit of analysis being examined: a nation, a household, an emotion, an electron. But Allen and Pardo refrain from identifying a clear unit of analysis. They casually list a slew of focal interests, including "juridical proof" (pp. 1-2, 3, 4, 5); "the structure of

[1] For an early expression of agreement with the core features of the relative plausibility approach, see Simon, D. (2004). A third view of the black box: Cognitive coherence in legal decision making. *University of Chicago Law Review, 71*, 511-586, pp. 559-569.

proof at trial." (pp. 2, 4); a "theory of proof" (fn. 1); a "system" [of proof] (p. 4); the "entire litigation process" (p.4); and the "evidentiary proof process" (pp. 4, 5). This multiplicity of focal points defies any meaningful empirical endeavor, especially given that each of these constructs is a heterogeneous mix of features, and testing each of the features would require employing a dedicated empirical project. Moreover, Allen and Pardo insist that their project is driven by data: "This task is empirical: what is the best explanation of the data, where 'the data' are observations of how the American legal system structures proof at trial." (p.4). Yet, the authors offer no data. With no definition of the unit of analysis, no testable hypotheses, and no data, the empirical claim is plainly befuddling.

Instead, Allen and Pardo seem to rest the empirical claim on referencing psychological research pertaining to the manner by which fact-finders process evidence. In particular, they point to Nancy Pennington and Reid Hastie's Story Model, according to which people make sense of complicated bodies of evidence by constructing narratives, with the best story leading to the chosen verdict.[2] Allen and Pardo explain that "jurors typically attempt to construct narratives to fit evidence..." (p. 17) and that the construction of narratives is "an instantiation of how virtually everyone reasons about the world at large."[3] Allen and Pardo consider it a virtue that their "account fits with the leading empirical account of evidence processing." (p. 35).

But this broad reliance on the story model has been a cause for concern.[4] To appreciate the issue, we must examine what the model does and what it does not. The story model posits that narratives are structured around episodes of human action that are held together by causal and physical relationships and are driven by mental states: protagonists encounter and react to a chain of events, each of which triggers additional reactions, eventually culminating in an action that constitutes the climax of the story. The story model was constructed from simulated jurors' treatment of a real life murder case that began at a local bar when tensions arose between Alan Caldwell and Frank Johnson about the former's girlfriend. Caldwell reacted angrily and threatened Johnson with a razor. Johnson backed off and went home. Later on, a friend invited Johnson back to the bar. Johnson agreed, and brought his fishing knife along. At the bar, Johnson encountered Caldwell and the two stepped outside to settle the dispute. A fight erupted. At some point,

---

[2] For almost three decades, relative plausibility's reliance on psychological research has amounted mostly to perfunctory references to the story model, most often to the single article Pennington, N. & Hastie, R. (1991). A Cognitive Model of Juror Decision Making: The Story Model, Cardozo Law Review, 13, 519-574. Occasionally, the authors allude also to work by Shari Diamond and colleagues and by Vidmar & Hans (2007), who too tend to refer to the story model. See Diamond, S. S., Rose, M. R., Murphy, B., & Smith, S. (2006). Juror questions during trial: A window into juror thinking. *Vand. L. Rev.*, *59*, 1925-1972, fns. 27, 52; Vidmar, N., & Hans, V. P. (2007). *American juries: The verdict*. Prometheus Books, pp. 135-137.

[3] Allen & Pardo (2008, fn. 34). See also Allen 1991, p. 403.

[4] See Simon (2004), *supra* note 1, p. 563-4. For an early critique of narrative theories of evidence, see William Twining, *Rethinking Evidence: Exploratory Essays* 225 (Basil Blackwell 1990).

Caldwell punched Johnson in the face and Johnson plunged his knife into Caldwell's chest, killing him. The key experimental finding was that jurors who rendered a murder verdict tended to imbue the physical acts with congruent mental states and goals: Johnson's humiliation, vengeance, and intent to kill Caldwell. Conversely, those who chose the self-defense verdict imbued the story with fear, avoidance of trouble, and self-protection.[5]

Notwithstanding its brilliance and invaluable contribution to the literature, the story model's shoulders are simply too narrow to support the full breadth of the relative plausibility framework. The model is most capable of discerning "human purposive action sequences"[6] and "causal and intentional relations among events"[7] which were indeed prominent in that particular case. But, the model seems underpowered to handle the vast array of cases that inhabit our legal world. Notably, the model is of limited use in cases where there is not much of a story to tell, such as when the evidence is fragmented (probably the majority of evidence produced in criminal cases): a neighbor's statements about hearing shouts, a fingerprint analysis, a lineup identification, and a credit card receipt. The same would apply to cases that hinge on technical judgments of patents, trademarks and antitrust, discerning causes of physical events, setting damage awards, and on.[8] Indeed, these limitations have not been lost on critics of the relative plausibility framework. Kevin Clermont, for example, has noted that the theory "comes with baggage."[9]

In response, Allen and Pardo have had little choice but to qualify the grounding of their framework in the story model. They state that relative plausibility theory "deals with explanations, not stories," (fn. 7) a protestation that repeats twice more in the target article,[10] and they too list domains that are not amenable to the construction of stories.[11] Thus, it seems that Allen and Pardo find themselves in an awkward position of anchoring their theory in the story model, while also bending over backwards to distance themselves from it. Indeed, the authors seem to have turned to rely more heavily on holism, a basic psychological mechanism that is deemed to underlie the story model and narrative construction. The authors claim that it is clear that fact-finders "do in fact proceed holistically—this is how information is processed" (p. 34), and insist that the sequential updating presumed by probabilistic models is inconsistent with holism. (p. 11). But

---

[5] P&H, 1993.

[6] Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision making. *Journal of personality and social psychology*, *51*(2), 242-258.

[7] Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 521

[8] See Simon (2004), p. 564.

[9] Clermont, K. M. (2015). Trial by Traditional Probability, Relative Plausibility, or Belief Function. *Case W. Res. L. Rev.*, *66*, 353-392, pp. 359-560.

[10] See fns. 86, 186-188 and accompanying text.

[11] The authors list patent litigation (fn. 86), and "anti-trust litigation. Or no-fault divorce. Or much contract litigation." (p. 35).

Allen and Pardo do little to explicate their reliance on holism. They state merely: "We refer to this process throughout this article as 'holistic,' but all we mean by the use of that term is that explanatory considerations structure the process of proof." (fn. 7). This laconic and opaque treatment seems odd given that holism serves as the bedrock of the entire framework. In sum, even as one welcomes the authors' reliance on holism,[12] their empirical aspiration remains utterly unfulfilled.

My second major critique of the target article is that by refraining to pry deeper into holism and to reckon with its implications, Allen and Pardo deprive the reader of the opportunity to appreciate both the promise and perils of the cognitive process they espouse. I attempt to make up for this omission in the limited confines of this comment.

We begin with the basic function of holistic processing. In all but easy cases, legal questions present fact-finders with sprawling evidential arrays that are rife with uncertainty, incommensurability, and conflict. To illustrate this complexity, take Ron Allen's account of the seemingly simple task of judging deception from a witness' demeanor. This task can involve observing the witness' sweating, twitching, slouching, gaze aversion, and voice inflection.[13] The task becomes much more expansive and more complicated by the fact that the witness' demeanor constitutes just one aspect of her credibility, and that the credibility of a single witness will usually constitute just one piece of the larger evidential array. According to one analysis of the famous trial of Sacco and Vanzetti, the evidential array contained no fewer than 338 facts and propositions.[14] Inevitably, some evidence items point in one direction while others point in different directions (that, after all, is what makes cases difficult). Thus, at bottom, the key challenge facing factfinders is to form a discrete judgment from vast and cacophonous arrays of evidence. Yet, thankfully, fact-finders do manage to perform the task and even meet the lofty evaluative criteria of a evidentiary conclusions—consistency, coverage, simplicity, coherence, and the like.[15] How do they (and we) do that?

Grounded in the principles of Gestalt psychology, holism is best captured by the notion that the properties of a whole are different from, and typically more pronounced than, the sum of the

---

[12] A great deal of research supports the pervasiveness of holistic processing. See e.g., Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41,* 258-290; Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology, 21,* 107-112; Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13,* 295-355.

[13] Allen, R. J. (1993). Factual ambiguity and a theory of evidence. *Nw. UL Rev., 88,* 604-640, pp. 625-626.

[14] Kadane, J. B., & Schum, D. A. (1996). *A probabilistic analysis of the Sacco and Vanzetti case,* pp. 80, 286–337. New York: John Wiley & Sons. According to this count, the prosecution's case contained 139 evidence items and the defense's case 199.

[15] On the criteria for explanation evaluation, see Allen and Pardo, fn. 7, fn. 74 and accompanying text.

properties of its constituent parts.[16] Based on theories of connectionistic cognition, holistic tasks are deemed to be represented in the brain in the form of networks, wherein the activation of a node denotes the strength of believability of the corresponding proposition, and each node is connected to other nodes based on the associations between the propositions. Notably, propositions will be positively linked to all other propositions that support the same verdict, and negatively linked to propositions that support the vying conclusion. Much like electric networks, nodes interact with and cross-activate one another, with the overall pattern of activation drifting towards states of equilibrium. A network will be considered at equilibrium—or *coherent*—when one subset of propositions, say, the evidence associated with one verdict is strongly believed, while the evidence supporting the opposite verdict is deemed week. Over the course of the evaluative process, fact-finders (unwittingly) alter their perception of the propositions, enhancing the ones that fit with the emerging choice and diminishing those that support the to-be rejected choice. In other words, the cognitive process boils down to transforming states of conflict-laden complexity into states of coherence, a process that can be captured by the framework of *coherence based reasoning*.[17] The lopsided representations in states of coherence are what provide the network its stability and, crucially, they enable fact-finders to reach discrete judgments with sufficient resolve and confidence. Indeed, high levels of confidence in the chosen decision, despite the difficulty of the task and the equibalance of the options, is one of the central and persistent findings in this line of research.[18]

---

[16] Wertheimer, M. (1924/1967). Gestalt theory. In Ellis, W. D., *A source book of Gestalt psychology* (Vol. 2). Psychology Press.

[17] In broad terms, this transformation is measured by obtaining participants' evaluation of the same items of evidence on different formats: first, as isolated vignettes, and second as items of circumstantial evidence in the context of deciding a large case. Studies find consistently that the judgments shift from being roughly neutral in the first measurement, to strongly polarized subsets in the second. These transformations are called *coherence shifts*. See, e.g., Glöckner, A., Betsch, T., & Schindler, N. (2010). Coherence shifts in probabilistic inference tasks. *Journal of Behavioral Decision Making, 23,* 439–462; Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General, 128,* 3–31; Simon, D., Snow, C. J., & Read, S. J. (2004). The redux of cognitive consistency theories: Evidence judgments by constraint satisfaction. *Journal of Personality and Social Psychology, 86,* 814–837; Simon, D., Stenstrom, D. M., & Read, S. J. (2015). The coherence effect: Blending cold and hot cognitions. *Journal of Personality and Social Psychology, 109,* 369-394. For a review, see Simon (2004).

Another framework that could explain these findings is based on Bayes Nets. See Lagnado, D. A., & Gerstenberg, T. (2017). Causation in legal and moral reasoning. In Waldmann, M. (Ed.). *The Oxford Handbook of Causal Reasoning*, 565-601.

[18] Studies show that the reported confidence in the decisions tends to be close to the high end of the scale, regardless of which verdict is chosen. See Holyoak & Simon (1999), Simon et al., (2004), Simon et al. (2015).

A key lesson from the research is that holism is at once highly effective and adaptive, but also rife with features that have inherent and unintended implications.[19] This realization bellies Allen and Pardo's coupling of holism as superior on both descriptive and prescriptive grounds. One should be able to endorse holism as a correct account of legal fact-finding while acknowledging that some of its implications are quite problematic from a normative point of view. I focus here in brief on only three such implications.

First, Allen and Pardo criticize the probabilistic model for its subjectivity and inaccuracy: they explain that numeric values, such as relative frequencies and known statistical distributions, are simply unavailable for most items of evidence (p. 9). They explain that "Subjective probability is truly subjective. It ... has no necessary relationship to advancing accurate outcomes." (p. 10). They protest that prior probabilities and likelihood ratios are "literally just made up by the decision maker" (p. 38), and reject Bayes nets because they "generally reflect the standpoint of the individual constructing the graph, rather than ... objective truth." (fn. 208). Conversely, the authors seem to imply that relative plausibility is immune to human subjectivity and maps relatively closely onto accuracy, primarily thanks to the criteria for explanation evaluation (consistency, coverage, simplicity, coherence and the like).[20] But missing from this optimistic viewpoint is the unavoidable truism that fact-finders have no access to the ground truth, and thus have no Archimedean point from which to leverage a true perception of the facts or to assess the actual validity of their own explanation evaluation. Holism, not unlike any other theory of cognition, is entirely reliant on subjective assessments. A mental model that is perfectly coherent and that meets all the criteria for explanation evaluation can easily be wrong. By the end of an adjudicatory process, every discerning juror is bound to hold a coherent view of the facts. Though, whenever two jurors disagree, one of those coherent mental models—if not both—will be mistaken. Indeed, one can assume that all mistaken verdicts ever made were borne by holistic judgments that were deemed by the fact-finders to pass the muster of the explanation evaluation criteria. In sum, while holism promotes confidence in the decision, it guarantees neither objectivity nor accuracy.[21]

---

[19] At various points, Allen has noted some of the features to be discussed here. He refers to connectionist processing and the interconnectedness of evidence (1991, p. 401), bi-directional processing (1991, p. 403, fn. 103), the findings by Schum & Martin mentioned below (Allen, 1991, p. 401), and the important work of Paul Thagard (see Allen, R. J. (1997). Rationality, algorithms and juridical proof: A preliminary inquiry. *The International Journal of Evidence & Proof, 1* (special), 254-275, 275). To the best of my knowledge, though, the implications of these themes were not pursued.

[20] Allen and Pardo suggest that the criteria will enable an assessment of the "sufficiency of evidence or the reasonableness of particular findings." (p. 17). The authors concede that a good story may be false, but go on to state, without explanation, that that "has nothing to do with the relative plausibility theory, which deals with explanations, not stories." (fn. 7).

[21] See Simon, D. (2012). *In doubt: The psychology of the criminal justice process*, 169-170. Cambridge: Harvard University Press.

Second, the holistic process itself has the potential to distort verdicts. Holism does not fall from the sky. Rather, it is constructed via a cognitive process that entails a transformation of the fact-finder's mental representation of the evidence and drives the evidence towards a more extreme view of the case. Morphing the evidence—actually, distorting it—is the lifeblood of the holistic process.[22] This means that fact-finders will typically reach conclusions that are stronger than warranted by the underlying evidence: evidence that supports the chosen verdict will seem stronger than it should and evidence that supports the rejected option will not receive the credence it deserves. The cognitive relegation of the non-fitting evidence might actually subdue true evidence that could have alerted the fact-finder to a looming mistake. This form of evidence distortion might not have serious implications in civil cases, where the verdict choice does not normally hinge on whether the evidence surpasses or falls short of the standard of proof threshold by a smidgeon or by a mile. The same would apply in criminal cases where the fact-finder leans towards acquittal. Though, when a juror is leaning slightly towards a guilty verdict, the relegation evidence can actually quash a would-be meritorious conclusion of reasonable doubt. In other words, the diminishment of the evidence that supports the rejected verdict might lead the fact-finder to conclude that there is little merit to the defense's position, or in the terms of relative plausibility theory, that "there is no plausible defense explanation" (p. 15). Criminal convictions could similarly be influenced by the high levels of confidence observed throughout this body of research. It important to acknowledge that the typically strong confidence is, to a large degree, an epiphenomenon of the cognitive process rather than a property of the evidence itself. In fact, the research shows that the reported confidence correlates with the magnitude of the distortion of the evidence.[23] This epiphenomenal inflated confidence could potentially make all the difference by boosting (however honestly) a weak propensity to vote for conviction up to a level of confidence that amounts to feeling firmly convinced in the defendant's guilt.[24]

A third implication stems from the intricate interconnectivity within holistic networks. An unavoidable feature of interconnectivity is that the believability of a proposition does not depend entirely on its own truth value, but is influenced also by its cross-activation with other propositions in the network: a proposition will be deemed more reliable when it is associated with other believable propositions, and less reliable when associated with losing ones. This *non-independence* can help explain why Schum and Martin found that participants interpreted contradictory testimony as either probatively valueless or even corroborative, and "double

---

[22] Thus, evidence distortion is not exclusive to probabilistic frameworks. *Cf.* Allen, R. J., & Stein, A. (2013). Evidence, Probability, and Burden of Proof. *Ariz. L. Rev.*, *55*, 557-602. pp. 572, 577.

[23] The correlation of confidence with distortion is measured within-subjects. See Simon et al., (2004); Simon, D., & Spiller, S. A. (2016). The Elasticity of Preferences. *Psychological science*, *27*(12), 1588-1599.

[24] See Simon (2012), p. 175.

counted" corroboratively redundant testimony.[25] This finding is actually quite common. For example, studies have observed the (otherwise inexplicable) phenomena whereby weakening the defendant's alibi results in stronger belief in the prosecution evidence,[26] increasing the strength of a prosecution witness' testimony results in lower judgments of police coercion during an interrogation,[27] and learning that the suspect confessed to the crime increases (the erroneous) belief that his handwriting matches the writing on the note used in the robbery.[28] A recent study also shows that altering fact-finders' emotion towards the defendant changes their perception of the entire array of evidence in the case.[29] Given the conditional non-independence of the evidence items in these cases, the observed interactions are non-normative. Moreover, these findings also indicate that extra-evidential evidence—particularly loaded information, whether factual or emotional—can alter the fact-finder's perception of the correctly-admitted evidence in the case.[30] Indeed, this helps explain why lawyers are so intent on exposing fact-finders to those types of information.

Space constraints preclude a full discussion of the implications of holistic processing to legal fact-finding. Though, it should not be difficult to see how reckoning with these issues could offer new insights to our conception of the accuracy of legal factual findings, the limited effect of heightened standards of proof, the contours of rules of admissibility, the effectiveness (of lack thereof) of admonishing jurors to disregard information, judgments of harmless error, and more. While Allen and Pardo ought to be commended for the intellectual pursuit of a deeper understanding of our system of judicial proof, it is in my view even more important that we roll up our sleeves and reckon with the real-life implications of what we find.

---

[25] Schum, D. A., & Martin, A. W. (1982). Formal and empirical research on cascaded inference in jurisprudence. *Law and Society Review*, 105-151.

[26] Smith, B. C., Penrod, S. D., Otto, A. L., & Park, R. C. (1996). Jurors' use of probabilistic evidence. *Law and Human Behavior*, *20*(1), 49-82

[27] Greenspan, R., & Scurich, N. (2016). The interdependence of perceived confession voluntariness and case evidence. *Law and human behavior*, *40*(6), 650-659.

[28] Kukucka, J., & Kassin, S. M. (2014). Do confessions taint perceptions of handwriting evidence? An empirical test of the forensic confirmation bias. *Law and Human Behavior*, *38*(3), 256-270.

[29] Simon et al., (2015).

[30] See Simon (2012), p. 171-172.