Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Conjeturas y Refutaciones
N. 3 | 2022 pp. 253-272
Madrid, 2021
DOI: 10.33115/udg_bib/qf.i3.22597
Marcial Pons Ediciones Jurídicas y Sociales

ISSN: 2604-6202
Received on 23/03/2021 | accepted on 06/04/2021 | published online 00/00/2021

NATURALIZED EPISTEMOLOGY AND THE LAW OF EVIDENCE: A REPLY TO PARDO, SPELLMAN, MUFFATO, AND ENOCH

Ronald J. Allen*

ABSTRACT: In «Naturalized Epistemology and the Law of Evidence Revisited», the original target article for the various refutations that I comment on here, I revisited through a slightly different lens the subject of the article that I coauthored with Brian Leiter close to twenty years ago. That article has prompted four responses from Professors Pardo, Spellman, Muffato, and Enoch. Professors Pardo and Spellman basically accept the implications of the original article and offer useful but friendly amendments. Prof. Muffato apparently does not want to dispute over my ground and so changes the subject, and in doing so offers a number of interesting points. Only the fourth, Prof. Enoch, has the same doubts about the utility of my original article as I do of the genre that gave rise to it. I can thus be quite brief in discussing professors Pardo, Spellman, and Muffato, but it will take bit more effort to lay out the limits of Prof. Enoch's analysis.

KEYWORDS: probability; relative plausibility; epistemology; safety; sensitivity.

RECOMMENDED CITATION: RONALD J. ALLEN, 2021: «Naturalized Epistemology and the Law of Evidence: A Reply to Pardo, Spellman, Muffato, and Enoch», publicados en Quaestio Facti 2/2021», in *Quaestio facti*, 3: 253-272. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i3.22597

---

* John Henry Wigmore Professor, Northwestern University Pritzker School of Law; President, International Association of Evidence Science; Fellow, The Forensic Science Institute, China University of Political Science and Law. I am indebted to William Lawrence for excellent research assistance and to the Julius Rosenthal Foundation Fund for supporting this research.

In «Naturalized Epistemology and the Law of Evidence Revisited», the original target article for the various refutations that I comment on here, I revisited through a slightly different lens the subject of the article that I coauthored with Brian Leiter close to twenty years ago.[1] At the time, we focused on the distinction between purely a priori conceptual analysis on the one hand, and empirical inquiry on the other, and the admonition that «should implies can.» I had been in the grip of this methodology long before we wrote that article, although I would not have used the language of epistemology to describe my research program, and remain in its grip today, although with a slight evolution in focus if not in meaning. In the original article, I described my research program today as involving «inquiries—analytical or empirical—into how the human mind engages with its environment and the implications of that form of engagement for western legal systems—in particular the legal systems within the United States.»[2] The original article was intended to do two things. First, show the fecundity of this approach by briefly describing one of its major achievements, which has been to replace probabilism as the best explanation of juridical proof in the United States[3] with explanationism. The second objective was to cast doubts on one common methodology used to explicate juridical proof that does not subscribe to such an approach and remains focused on a priori conceptual analysis.

These latter efforts come in two general flavors. One focuses on the interior demands of the discipline in question and is not concerned with instructing the judicial system or its scientists in what to think or do. As I made clear, my original article is not addressed to them; I would not presume to instruct people working within other disciplines how to go about their tasks. The other flavor involves purported explanation of or prescription for juridical proof—which is the manner, from beginning to end, in which a system structures and resolves legal disputes—for, I presume, real legal systems, including the American system. This flavor does presume to be instructing jurists about the object of their inquiry, and on occasion the instruction is interesting and valuable. But often it is not, and I identified three variables common to the a priori approach that seem to explain this particular slip between cup and lip. First, much of this work is focused on weird hypotheticals that have no relationship to the American legal system. Second, and relatedly, this theorizing commonly neglects critical aspects of the actual state of affairs in the object being theorized about. Third, and I must candidly say quite remarkably for many of these commentators are well versed in the intricacies of epistemology, much of the theorizing entails literally impossible epistemological demands. As I also explained, my suspicion is that these three variables are derivative of a deeper conceptual problem that often inflicts the analytic disciplines of approaching juridical proof as a static rather than dynamic system that will yield simple and direct prescriptions, solutions, analyses, whatever, and this is false. This completely misconceives legal systems (and

[1] Allen and Leiter, 2001.
[2] Allen, 2021: 1.
[3] I return to the focus on U.S. law below in my comment on Enoch's paper.

here I dare to go beyond my own) which are complex and adaptive, not static and rule bound. They are, after all, complex human institutions.

This was all done in what I referred to as an admittedly «solipsistic» fashion, by which I meant, as I think was entirely clear, that I was explaining why I find much of this work of little value in my task of trying to advance knowledge of the actual legal system that I happen to know something about—the American legal system—whatever its value from an internal perspective of the various disciplines producing this work or for other legal systems. As I say, I would not presume to comment on such foreign territory. The editors of *Quaestio Facti* solicited that paper for their «Conjectures» section, anticipating that they would also solicit various refutations, which they have done. I am very grateful to the editors for having created this opportunity, and even more so to the commentators who have spent valuable hours critiquing the original paper.

And now I get to respond. There is, however, an immediate difficulty. Two of the papers, those of professors Pardo and Spellman, basically accept the implications of the original article and offer useful but friendly amendments.[4] The third, Prof. Muffato, apparently does not want to dispute over my ground and so changes the subject, and in doing so offers a number of interesting points. Only the fourth, Prof. Enoch, has the same doubts about the utility of my original article as I do of the genre that gave rise to it. I can thus be quite brief in discussing professors Pardo, Spellman, and Muffato, but it will take bit more effort to lay out the limits of Prof. Enoch's analysis.

1. PARDO[5]

I am grateful to professor Pardo for taking the time to provide his thoughtful refutation. Having introduced into the literature the concept of weird hypotheticals, perhaps it is fitting that I begin a response to a refutation in the somewhat weird way of quoting the refuter's conclusions, because I agree with them entirely:

> The methodological picture I have described is one way in which epistemology may contribute to the law of evidence. It is intended to be an optimistic one and an invitation to epistemologists. For anyone taking up this invitation, I think there are three important points that Allen's critique reveals. First, it is important (and sometimes difficult) to accurately describe the underlying legal details. Second, there is a complex relationship between epistemology and law. It is unlikely that useful connections will involve a simple application of some epistemic concept or issue to law—a particular concept or issue may have different features, or play different roles, in the legal context than it does in other epistemological concepts. Third, it is probably time to stop trying to «solve»

[4] Professor Spellman, apparently, does so somewhat begrudgingly based on my (quite friendly) efforts to give her some feedback on her developing skills as a moderator. It just goes to show no good deed goes unpunished.

[5] A disclosure is in order here. Professor Pardo was a student of mine at Northwestern and I have worked closely with him throughout his academic career.

> issues related to the Blue Bus and Gatecrasher hypotheticals and, for those interested in contributing to our understanding of the law of evidence, to focus elsewhere.

In Pardo's always erudite and at the same time sensible fashion, he puts the discussion into a very helpful framework. His analysis cuts right to the heart of the concerns that I spent considerably more time elaborating. The first problem with the weird hypotheticals is, as he says,

> that the *question* they are being used to answer is not a genuine problem for law in need of analysis or explanation. […] Because the hypotheticals are being put in the service of answering an ill-formed question, it is not surprising that after decades of trying no one has successfully answered the question or «solved» the problem. The second problem is that, even if the question being addressed were genuine, the examples do not provide enough information to answer it. The hypotheticals are framed so as to imply that the evidence appears to be sufficient on its face to satisfy the applicable standard of proof. Implicit in this framing, however, is the assumption that the standard of proof is a probabilistic threshold, […] which the statistic in the example surpasses. But this assumption is itself highly questionable. If, instead, standards of proof are better understood as explanatory thresholds, as Allen and I argue, then the evidence is no longer necessarily sufficient on its face to meet the standard of proof.

Exactly, and I could end this response right here. But… I will make two other points. I think that Pardo attributes to me a skepticism that I do not possess, and more importantly for purposes of this exchange, has no relationship to the original paper. He says «I remain more sanguine about the genre of scholarship that Allen critiques—in my view, epistemology may provide a greater source of understanding for legal evidence and proof than Allen's article may suggest.» My article is being misread if taken as suggesting something general about the relationship between epistemology and inquiries into the nature of juridical proof—although I will happily attribute the error to my exegetical skills. I meant to concentrate, in a laser-like fashion, on one and only one well defined issue—and that is the use of weird hypotheticals that ignore critical parts of the legal system and that make literally impossible epistemological demands—no more and no less. To be sure, one of the reasons for focusing on this limited set is that the employment of such hypotheticals is quite prevalent in the scholarship purportedly analyzing juridical proof, especially from a philosophical perspective, but at the same time that work does not uniformly employ this methodology. My intent, whether accurately conveyed or not, was to carefully limit my analysis and I was not making any broader claims. I am indebted to Pardo for bringing to my attention the need for this clarification.

The second point is that the use Pardo makes of hypotheticals actually shows even greater agreement between us than perhaps Pardo suggests. For example, in his discussion of the chemist hypothetical, he makes it less weird by adding another layer to the original problem: «Suppose that there is a small range in which the process is unreliable; sometimes the result is slightly over the true amount and sometimes it is slightly under, *but (given our best available knowledge) the former tends to happen more often than the latter*» (emphasis mine). The more layers one adds, the less weird the hypothetical becomes, and it moves in the direction of responding to my critique.

The original hypothetical posited certain facts of the matter, but it was unclear who knew those facts (another ubiquitous problem in this genre, actually). If no one knew them except some ethereal entity (like God or the people who write Gettier problems), they are completely irrelevant to the case at hand. If the facts are known in the real world, they can be presented to the fact finder and the parties can argue over their significance. Pardo has now added the last aspect, thus, as I say, moving it in the direction of my critique.

But I am still not sure what is added to our understanding of juridical proof by invoking the language of safety and sensitivity. The problem that processing evidence depends on generalizations that have complexity is not new to the legal system. The problem of false memory, for example, or stereotypical thinking, which epistemologists would call insensitive evidence, involve issues that the legal system has dealt with for hundreds of years. And so has evidence that involves a range that may or may not include the relevant variable (was the car being driven in a reckless manner, what was his blood alcohol level), which the epistemologists would call safe or unsafe depending on the case. Pardo or others may think that invoking the epistemological framework facilitates understanding the implications of such matters, and even though I do not, perhaps they are right. I have no basis nor any desire to deny that possibility. One should present one's analysis in the most effective manner possible.

His second hypothetical is not a weird hypothetical at all. It is based on a real case and presents a real problem. But the solution to the problem does not require reference to or resolution of issues of knowledge (thankfully given the record of epistemology on that score). The Commonwealth is essentially asking the appellate court to try the case on the basis of evidence not presented to the jury and that does not fit within the concept of judicial notice. Doing so would violate the Sixth Amendment right to a jury trial with a finding by a jury of guilt beyond reasonable doubt.[6] That is true whether whatever the fact finder does involves «knowledge» or not. It is not «knowledge» that is at stake, but a decision based on the evidence presented, which under universally accepted conventions would include the kind of data involved in this case.[7] Still, perhaps presenting it as he does helps clarify the issues at stake, and again perhaps he is right.

The upshot of this is that Pardo uses hypotheticals with a keen awareness of their doctrinal setting, of, in other words, their relationship to reality. Using hypotheticals in the way that he does, at least with the elaboration of the case of the chemist, is not at the core, and maybe not even at the periphery of my critique of the use of weird hypotheticals that ignore critical aspects of the legal system and that make impossible epistemological demands.

---

[6] The court did not elaborate the doctrinal basis of its decision, and it could have reached the same conclusion on the basis of state rather than constitutional law. The decision would be the same in either event.

[7] What is "evidence" is more complicated than it appears. Allen, 1994.

2. SPELLMAN

I am grateful for Professor Spellman's very interesting guided tour of the history of experimental cognitive psychology, and her situating it in the context of the present discussion. The difficulty I have in responding is that I have even less disagreement with her than with Pardo. For example, although she is in agreement that relative plausibility is currently a better explanation than probability of the nature of juridical proof, «thinking about how other people do think, and can think, (i.e., metacognition) is even a more essential foundation of the system than is relative plausibility.» I agree.[8] Various strands in the psychological literature were constituents in the complex set of variables that led to the abduction of the best explanation of juridical proof. Thus (and maybe the only place I actually disagree with her), it is not «ironic» at all that «all of these findings» of the work she discusses on causal attribution and the story model among other topics «are consistent with […] relative plausibility theory.» Quite to the contrary, they were partially responsible for the emergence of the theory.

I slightly disagree that what I call weird hypotheticals are examples of or very closely analogous to the thin hypotheticals that have often been used to start research programs in cognitive psychology. They are thin in the sense of having stripped down facts, but they are not weird in the senses of being completely divorced from reality, ignoring critical variables pertinent to the object under investigation, or making impossible epistemological demands.

But even if I am wrong about that, note that Spellman flips the script so to speak. The phenomenon that I analyzed involves invoking intuitions based on these peculiar scenarios, and then engaging in heroic efforts to justify them. Spellman is describing the opposite of this process. In her cottage industry, the responses to thin hypotheticals become the occasion to dig into the reasons for the responses to try to understand the underlying thinking process rather than engaging in an effort to rationalize the initial responses. In my cottage industry, the participants do not try to understand what motivated the intuition, but instead it is taken as the gold standard that must be justified. I have no idea what explains this, and it is all very peculiar. As I noted in the original article, there is essentially no evidence in support of the operant intuition about statistical evidence within the American legal system. No matter how often this intuition is evoked in response to weird hypotheticals and reported in various literatures, the actual American legal system applies normal evidentiary standards to statistical proffers and continues to domesticate new kinds of evidence, including things that begin to look close to naked statistics like DNA evidence. If the process that Spellman discusses were in play, the conclusion of all the products of this cottage industry would be that the intuition has not borne up to scrutiny and is rejected. I do not know of a single article in this genre that reaches that conclusion.

[8] Allen, 2011.

I will give just one example of the different dynamics at play. In Spellman's presentation of the blue bus hypothetical, she goes through various ways in which one might try to understand why people give the answers that they do to the experimental psychologists, or as she says «I don't care much about what people think of as answers to weird hypotheticals, but I do care why they think it.» That is exactly the right question to ask. The weird hypothetical folks I focused on, by contrast, literally never ask that question but instead try to construct rationalizations that would justify the response. In my opinion, this is a perfect example of the possible beneficial effects of internalizing the lessons of naturalized epistemology.

The different tasks undertaken explains one other area where we may disagree slightly. Spellman says «Allen's claim that weird hypotheticals ignore important aspects of actual legal systems—because they do not capture all the information in trials and they tend not to represent all the values that go into a trial decision—is absolutely true. But it is mostly irrelevant.» It is irrelevant to some extent to Spellman's tasks, but as I have laid out, empirical inquiry that must begin somewhere is not what I was criticizing. I was criticizing the opposite.

3. MUFFATO

I am particularly grateful to Professor Muffato for his contribution. As I understand his research program, it is the most distant of all the commentators from the central objectives of mine. He is a philosopher, apparently with a particular interest in linguistics, doing philosophy; I am an empiricist trying to understand a complex, sprawling system. He, with excess of kindness, suggests that I am «a profound connoisseur of statutory and common law of evidence in the U.S.A.» and so far as I can tell, so is he of the complexities and mysteries that go under the general label «epistemology.» I suspect it is because of these different sets of knowledge that he chooses not to dispute over the central focus of the original article and my research program, which is understanding the nature of juridical proof; indeed, he apparently is convinced not only of its value but of its discoveries. No, he shifts ground to give a penetrating analysis of the conceptual framework that he believes I have adopted or promoted, and then critiques how well that framework is actually reflected in or constrains the work that has been done. The reader can immediately see where this is going, of course—in my opinion he should pursue his interests, as I do mine. Regardless of any external judgments about its compatibility with philosophical or other norms, the turn to empirical and analytical inquiries into propositions with truth value that respect the admonition that «should implies can» has been quite beneficial to reinvigorating the field of evidence. That for me is the measure that matters.

I could end this comment with the simple recognition that different disciplines, and different scholars, pursue differing objectives, but there is a bit more to say. Muffato goes beyond elaborating the implications of his own field to give advice to those pursuing one of the interesting strands in the field of evidence today that

plays an important part in the original article—the ongoing shift in conceptualizing juridical proof as probabilistic to explanatory, an example being the theory of relative plausibility. Here he unfortunately is subject to a version of the central argument in the main paper—to wit: that advice about a legal system and research into it might do well to accurately reflect the object of the inquiry.

How is this reflected in Muffato's refutation? My original paper extolled the virtues of a research agenda for the field of evidence that is an example and reflects the lessons of naturalized epistemology. It compared that methodology and one of its main products, the theory of relative plausibility, to the results of a different methodology, purely a priori conceptual work heavily informed by probability theory, that has had little success recently in advancing understanding of juridical proof. Muffato has two objections to this: first, that the evolution of the theory of relative plausibility involves at best a pale form of «naturalized epistemology,» and second, that we are not doing it very well, as we have not conducted any controlled studies of juror reasoning.

I begin with the second and more important point to demonstrate that he misunderstands what he is criticizing and misevaluates its empirical support. Thus, the criticisms that he advances miss their target. I then conclude with a few words about what comprises «naturalized epistemology.» According to Muffato, «the empirical base of [the relative plausibility theory lies] [a]lmost entirely in the […] "story model" resulting from the researches in social and cognitive psychology conducted by Nancy Pennington and Reid Hastie.» He goes on to say that «Allen might reply that explanations in a trial context are not stories, and that holistic reasoning supplies the lacking part in RPT: but in that case he should go into details and make explicit testable hypotheses about the holism he endorses.» And, he complains, I have not provided any new evidence from controlled studies that goes beyond what is presently in the literature, and thus the only empirical support for relative plausibility is woefully inadequate. This misconceives the project.

Muffato is under the misimpression that juror (human) reasoning involving stories and their generalization, explanations, is the overwhelmingly important aspect of the relative plausibility theory, and thus its success may depend on whether well-structured empirical studies support a particular theory of juror reasoning. This is simply false. Relative plausibility arose as an abduction influenced by a cacophony of variables, of which how people reason is only one. That so far as we know humans do and must think in a certain way consistent with the theory obviously provides support for an abduction about how a sensible legal system employs that form of reasoning. If further evidence is forthcoming about that matter, it may influence the inferences drawn about the nature of juridical proof, but the fact remains that the present best explanation of juridical proof, including how people reason, is explanatory rather than probabilistic. To be sure, it would be wonderful if further evidence were forthcoming about how humans reason, but it is odd to critique a theory

that is relying on the extant evidence as one data point for failure to gather more through studies in a completely different discipline.

Perhaps recognizing the unpersuasiveness of his complaint about lack of evidence, Muffato goes on to list seven of the variables that contributed to the abduction that juridical proof is best understood as explanatory, and then comments: «It seems then that Allen and Pardo aim at explaining empirically each one of the seven objectives enumerated and their mutual connections.» And we are criticized for not having done that work through carefully controlled studies. This has it completely backwards. Those variables and others contributed to the abduction; they are not the subjects of the analysis. The objective of relative plausibility theory is not to «explain empirically» the seven (and other) variables and their almost infinite variety of interactions—and it is a good thing, too, as I doubt that could be done in the way Muffato insists it must proceed. Rather, those are data points that led to the abduction in question.

But still the strength of the evidence matters for the persuasiveness of the abduction, and according to Muffato, the only empirical evidence for relative plausibility is the psychological evidence concerning the story model. And that is simply not good enough. This is an astonishing claim both about the legal system and more generally about what it means to be evidence. So far as the legal system is concerned, evidence scholarship has explained how the variables he identifies fit into a coherent picture of juridical proof. If there is a better explanation, it has not been forthcoming. That better explanation in turn would rest on counterexamples to the various points that have been made about the actual operation of the legal system, which again would be forthcoming if they existed. And again, they have not been offered. This is not an example of random sampling that he suggests is the sole basis of empirical knowledge, but he is simply wrong that knowledge of the sort humans actually use does not come from other sources than controlled studies—like observation and induction (no matter what Hume thought).

More pointedly, all of this overlooks the massive amount of evidence that supports relative plausibility as a better explanation than probability of juridical proof (which, do not forget, is the question under examination—which is the better explanation). To identify just some of this set: probability explanations of burdens of proof lead to unresolvable paradoxes; those same explanations are internally inconsistent (a probabilistic burden of persuasion can lead to any error distributions at trial); no human can reason about real life like a Bayesian; virtually never do relative frequencies about litigated matters exist; people do, in fact, reason through a focus on the alternatives available; parties form almost without exception, and judges often make them form if there is any resistance, a theory of the case; the theories of the case are virtually always presented as alternatives to the fact finder (judge or jury); the rules of evidence nudge the trial process in the direction of competing explanations; the parties are the masters of their cases, and courts do not interfere with party presentation for the most part; the Supreme Court of the United States has instructed the lower federal

courts to bear in mind that «a syllogism is not a story»; closing arguments weave the evidence together in as compelling an explanation as possible; appellate review of fact finding is differential, and so on.

Still, the complaint might continue, most of this has not been discovered through controlled studies. That is true to some extent,[9] but in part because it would often be a huge waste of time. Many of the variables that have been identified are ubiquitous (and as noted above subject to counterexamples that have not been forthcoming). It would be asinine for a scholar to ask for funds to do a study of whether parties in their pleadings have to articulate plausible explanations (again in federal courts by mandate of the Supreme Court and the rules of civil procedure); the literature on advocacy and «story telling» is rich and uncontradicted; one does not have to do a random sample of cases to «know» that the standard of review of fact on appeal is deferential (one only has to know how to do legal research), and again and so on. It was not discovery of these points and their interactions that is the contribution of relative plausibility, but the weaving of them together into the best extant explanation of the nature of juridical proof.

More generally, think of the implications of the view that empirical knowledge can only come from well-structured controlled studies. This precludes geology, astronomy, much of epidemiology, paleontology, sociology, macroeconomics among other disciplines from the reach of knowledge, which to say the least is curious. None of this is to disagree with the admonition to conduct controlled studies if that is where one's interests lie, and in fact one of the advantages of the relative plausibility theory is that it produces interesting questions that can be studied and confirmed.[10] By comparison, the probabilistic explanation produces verifiable propositions that have been systematically disconfirmed.[11]

There is another ironic similarity between Muffato's critique and the original article. The original article pointed out that the work being critiqued often made impossible epistemological demands. Muffato is arguing that the work on relative plausibility has not lived up to impossible standards given the object of inquiry. But note what happens when he engages in his own factual exegesis of a sprawling phenomenon analogous in that manner to the nature of juridical proof. To support his argument about the inadequate empirical basis of the relative plausibility theory, he draws a distinction between how «social theorists and historians» on the one hand, and «the officials of a legal order» on the other go about their affairs. With respect to the latter, he asserts:

---

[9] But see the next footnote.

[10] For example, see Diamond et al., 2015, examining the role of alternative explanations in actual jury deliberations. This is recent empirical work of that sort that Muffato calls for providing evidence confirming the relative plausibility theory, but which he rather surprisingly ignores.

[11] Allen, 2014, 2015, 2017.

> The officials of a legal order […] do not care about all these descriptive facets (and do not praise explanations that preserve the «pluralism» of the practice) when they try to set what the law is or «says» on a specific matter. Neither do formal legal doctrine and descriptive jurisprudence/legal theory, even if they work on a different level of systematic and generality. This happens because, while operating with a (not too precise) common sense concept of law/legality, they are concerned with deriving from the law itself—not from the concept of law—the (more precise) technical criteria for its description—rectius: for the recognition of its sources and the determination of its normative contents—in order to find out the legal solution to a practical problem. Officials and formal legal doctrine assume law's criteria of self-delimitation as given: an assumption an empirical social scientist cannot accept, but must substitute with a testable theory about the effectiveness of the preferred rule of recognition.

How does Muffato know what «officials of a legal order» do or do not do, what they care or do not care about, or what they assume? And to make matters worse, which «legal order» is he talking about? Is he asserting universals applicable to all imaginable «legal orders»? How did he make such remarkable discoveries? Through controlled empirical studies with appropriate randomization as he is demanding of relative plausibility? Of course not. It is not even clear how a serious study of the sort necessary to examine such questions could be organized.

When Muffato turns to making assertions about the legal system, he reports the results of his observations over time, and thus he must think that to be a usable version or reflection of «knowledge,» certainly good enough for the human condition in these circumstances. I agree. One of the central themes of the work that goes under the umbrella of «relative plausibility» specifically or «naturalized epistemology» as it pertains to legal systems more generally is that the tool should be tailored to the task. Many questions are amenable to examination by controlled studies; others are not. That some matters resist such forms of inquiry does not mean we can know nothing about them.

Perhaps the most surprising aspect of Muffato's critique is that, while he mentions that the objective of the work on relative plausibility has been to show that explanatory considerations better explain the nature of juridical proof than does probabilism, he then proceeds to ignore the point. His critique is a series of criticisms of the empirical adequacy and conceptual purity of the efforts to date that are entirely unexplanatory of the explanatory merits of explanationism compared to probabilism. He also neglects that science is a work in progress, and in this case it is a work in progress on a complex adaptive system, which complicates matters considerably (and substantially undercuts the value of the methodology he is promoting). Perhaps his various criticisms will prove useful to guide the evolution of the field, although frankly I could provide a more compelling list of the holes in knowledge that it would be helpful to fill than those that he provides. We shall see. Providing useful criticism in this context, however, cannot neglect the comparative nature of the enterprise.

This leaves the first of Muffato's critiques, to wit: that the work leading up to the theory of relative plausibility is at best a weak and pale form of naturalized epistemol-

ogy.[12] He says for example: «without the support of precise laws of supervenience (or of concept- and theory-reduction), sheer reference to empirical or scientific evidence in the context of the explanation of a social practice, even if useful, is not sufficient to qualify a theory as naturalistic or empirical, in the sense of naturalized epistemology (and jurisprudence).» I have no idea what «precise laws of supervenience» means, but this is tantamount to saying that Alvin Goldman—the most important scholar in the field—is not doing naturalized epistemology. Rather obviously various scholars, like Brian Leiter and Alvin Goldman, disagree with Muffato about the extension of the field. Muffato and whomever can argue about it. For reasons I have advanced, approaching the field of evidence from the perspective that Leiter and I identified, and I slightly embellished in the original article, whatever it is called, has had a positive effect on understanding the nature of juridical proof, which for me is the objective. I most certainly have no intent to instruct epistemologists on matters internal to their own field. Moreover, I think that he knows exactly what Leiter and I were saying (and I essentially repeated), to wit: that inquiries into the field of evidence would do well to concentrate on propositions with truth value and methods that may verify them, and that any normative advice should adhere to «should implies can.» I wonder if Muffato really wishes to argue with either proposition, or whether he actually thinks it might matter that technical but debatable criticisms internal to philosophy might be directed at our characterization—or more importantly the work that has followed in its wake. I also wonder how the methodology he employs in his refutation can possibly assist in deepening our knowledge of juridical proof. Of course, that is my agenda, not his.

4. ENOCH

Criticism is central to the growth of knowledge, and that is why I am particularly grateful for the time and effort Professor Enoch has expended on his refutation. Of all the refuters, he has the greatest doubts. As I do concerning his refutation. Before explaining why, I want to emphasize that I am approaching the critique of Enoch's refutation in the same spirit of my original paper, which is to say viewing it from my perspective and not his. Part of that perspective involves focusing on the payoff

[12] Muffato makes various discrete points that doth protest too much. For example, I describe my research program as generally within the purview of naturalized epistemology as it pursues propositions with truth value either analytically or empirically, to which, in what is supposed to be a criticism I believe, Muffato responds: «the engraved disjunction—«analytical or empirical»—is a bit confounding, given that naturalization and rejection of the analytic/synthetic distinction are generally seen as coextensive.» To me, the significant point is the emphasis on propositions with truth value, whatever the internal debates are within his field. On that score, much work on the nature of juridical proof has advanced through straight forward logical analysis, such as the substantial clarifying of burdens of proof. See Allen, 2014. I see no point to worrying about whether that process is hermitically sealed off from empirics in a fashion that might falsify Quine's argument about the analytic/synthetic distinction. Quine, 1951.

of different forms of work for my, not his, interests. Thus, his extended inquiry into whether intuition mining can ever be of value is simply a distraction. My focus was not on intuition mining per se but instead on the role of intuition mining based on weird hypotheticals that ignore critical parts of the legal system and that make literally impossible epistemological demands, and why such work has had a minimal payoff with regard to advancing knowledge of the nature of juridical proof. I thus ignore the many discrete points that are problematic in Enoch's article[13] and just concentrate on the ones that might matter.

From that perspective, Enoch's argument begins on a curious note, and gets as Alice said, curiouser and curiouser. He begins by saying that he fails to recognize his aims in my original article, which focuses on the utility of a certain kind of work for advancing the understanding of the American legal system. No, his aims are much different; they are «to reach knowledge and understanding of important relevant truths.»

Put aside that his characterization means that understanding the American legal system does not involve «important relevant truths» on the order of such questions as whether some theory is «true» or «enlightening» or «promote[s] insight into the nature of statistical evidence, and indeed evidence more generally.» If he were just saying that his motivations differ from mine, I would simply agree and urge him to pursue his interests, but he goes beyond that and refers to my analysis of the American legal system as «the academic analogue of American exceptionalism.» As I said in the introduction to this article, I think it is clear why I focused on the American legal system, and that is because it is the only one that I know in depth. But regardless, as I would think Enoch would know, «American exceptionalism» is the idea that the United States is exceptional in its pursuit of freedom and liberty and is a beacon providing an ideal for other cultures to emulate. There is in my original paper not a word of comparison of any aspect of the American legal system to that of another culture or country. Unlike Enoch, I am trying to understand an embodied system, not an ethereal one.

But again, curiouser and curiouser. In Enoch's pursuit of disembodied truth in the articles that I cite, he and his coauthors have scores of citations to cases (50 or so, with some multiple citations to the same case), and every single one of them save two is to American case law. Those two exceptions are English. Perhaps that is all it takes to transform an article into a cosmopolitan pursuit of truth rather than a slightly scandalous focus on American law. When Enoch et al. articulate what motivates their analysis at least in part, they assert that «the powerful, uncompromising intuitions of pretty much all of scholars» is that there is a critical difference between

[13] Like for example, his suggestion that perhaps I am so misguided as to think that one can "dismiss … the relevance of the probability calculus entirely when commenting on legal evidence." Perhaps this is not just a subtle *ad hominem* but simply reflects a lack of command of the legal literature. See, e.g., Allen, 2017.

statistical and nonstatistical evidence and that the law is consistent with that intuition.[14] Those «intuitions» for decades have come from scholarship focused on American and, to a lesser extent, English case law—just as his citations demonstrate. Indeed, what would it mean to have intuitions about the use of evidence in «the law» unless the phrase «the law» were referring to something real—like American law for example?

This might appear as little better than the «academic analogue» to name calling, but to the contrary there is a serious point lurking here. Maybe it does not matter where one's examples come from if one is investigating timeless truth, which Enoch informs that he is. But the truths that he is reaching for are not disembodied truths concerning statistical or other forms of evidence; the issues that he focuses on—sensitivity, safety, incentive effects, deterrence rationales for procedural and evidentiary matrices—are well known in the literatures of philosophy, economics, among others, and yes, even law. His efforts do not advance understanding of these in disembodied forms but instead prescribe how courts should use evidence in a certain way and attempt to justify the reported but erroneous intuitions about how they would use it.

But, what courts? All the courts in the world? Does he think the intuitions he relates would apply to those structuring or working within Kadhi courts? And how do those in Kenya differ from those in Zanzibar? How about the courts in Germany, Azerbaijan, China or Tanganyika? Sharia courts? Rabbinic courts? Are the scholars whose intuitions he is mining up on the evolution and present conditions of the world's varied systems of justice so that we can count on their intuitions to be informed and universal? Or is this just the paradigmatic example of what I am trying to avoid in my work and advising against in the work of others if they want to advance the understanding of real legal systems—sitting in a well-appointed office in the Western tradition thinking about «evidence» or maybe «evidence law» and dreaming things up?

I suspect so. I have worked on reform of the law of evidence and procedure in multiple countries and on multiple continents entailing widely divergent histories and economic, social and political contexts, and not to put too fine a point on the matter, it verges on the ridiculous to analyze juridical proof in a disembodied fashion. The critical question facing legal systems across the world is not some undiscovered truth about statistical or any other kind of evidence; rather, the critical questions are how to structure dispute resolution and use different kinds of evidence given the historical, economic, social, and political context of very much embodied legal systems.

Enoch thinks to the contrary. He thinks that examining an idealization such as naked statistical evidence may, like the mathematics of frictionless surfaces, yield insights into the real world. Further, he says:

[14] Enoch et al., 2012

> think about the suspiciousness of relying on statistical evidence in the real world; now assume all these «friction»-like factors away, so that you remain with an unrealistic clean case of statistical evidence. Does it feel like you've assumed away the problem? Of course not. These other factors are much more like friction to Newtonian mechanics, and much less like air resistance to aerodynamics.

Maybe you have not assumed away the problem, but you have ended the argument. The argument is about, and only about, naked statistical evidence. Once you assume away naked statistics and have non-naked evidence, no one—not even the individuals whose intuitions he is relying on—have any doubts about the approach of American courts—normal standards of relevancy, materiality and probative value apply; there is nothing unique about statistical evidence. I realize that what American courts actually do may not matter because knowledge of such things does not rise to the level of disembodied, timeless truths for which Enoch and his colleagues are striving. But if real court systems are not behaving consistently with the intuitions at the base of Enoch's argument, it is a complete mystery what the point of the intuitions might be.

There is, however, an ambiguity in Enoch's presentation. At times, Enoch refers to «unrealistic» examples of naked statistics, yet at other times he seems to think that the unrealistic case exists in the real world. For example, he criticizes my and Pardo's realistic examples of statistical evidence to demonstrate the erroneousness of the reputed intuitions in play because he claims that they mix up different types of statistical evidence and evidentiary contexts. He is focusing, he says, on «the phenomenon sometimes called base-rate evidence, sometimes market-share evidence, or sometimes naked statistical evidence…» as though these were synonyms. Yet both market share and base-rate evidence are potentially admissible evidence (and thus not «unrealistic») in every jurisdiction in the United States; thus, precisely what is being distinguished from what and why is somewhat opaque. [15]

Thankfully, Enoch's views on the ontology of naked statistical evidence does not matter, because the unrealistic case litigated on the basis purely of naked statistics does not exist in the real world. It is literally not possible to have and only have such evidence in an American case. At a very minimum, and perhaps unknown to Enoch, there is one and only one universal law of (American) evidence, even though it is nowhere written down, and that is that everything needs to be shown to be what it purports to be. Evidence, naked or otherwise, does not walk into court on its own power, hop up onto the witness stand and offer itself up for analysis. Someone purportedly with pertinent knowledge has to present it, and thus the idealization of a case decided solely on naked statistics never in fact materializes. Consequently, and critically, for Enoch's argument to be analogous to frictionless surfaces and to

[15] However, he may not be keeping evidentiary niceties straight, admissibility, sufficiency, and probative value being different concepts, and the question of the use of any particular statistical proffer could vary over these legal principles. For more on failing to keep distinct different evidentiary matters, see n. 19, infra.

have any meaningful content, it would need to extend to the realistic condition of «friction,» which here means non-naked evidence, and show how the idealization shines light onto the real world. That demonstration is entirely lacking in his and his colleagues' work.[16] Nowhere of which I am aware, at any rate, is any insight offered from an examination of the frictionless naked statistics to the world of friction inflicted non-naked evidence, and thus the entire body of work fails to live up to Enoch's own articulated standards.

Enoch's argument about the incentive effects involved with statistical proffers fares no better than his more abstract attempt to justify rootless intuition mining as a method of obtaining disembodied knowledge. Ironically, given that my primary point was the lack of utility to analyzing weird and unrealistic hypotheticals that make impossible epistemological demands, Enoch adds to the genre in responding to my demonstration that he and his coauthors' employment of incentive effects does not justify the intuitions in question. Their arguments made the critical mistakes of failing adequately to explore crucial distinctions, for example between bus drivers and bus companies, and offering wildly implausible assertions concerning likely human responses to differing conditions.[17] Consider his replay of the gate-crashing scenario:

> But if you know that you may be convicted on the basis of statistical evidence alone, you know that your own action has little effect on your legal fate, or anyway, less effect than it has in a legal regime that is more suspicious of statistical evidence: After all, if so many people crash the gates, then the statistical evidence will be available against you even if you happen to buy a ticket. And if so, whatever incentive we may have hoped the criminal law supplied to buy the ticket rather than crash the gates is seriously eroded by the willingness to rely on naked statistical evidence. Notice how modest our claim here is: We're not saying that accepting statistical evidence always, necessarily, eliminates the relevant incentives. What we are saying is that the evidence law regime—and in particular, the attitude towards statistical evidence—may affect the relevant incentives, and at least often, in a counter-productive way.

---

[16] Remember that the intuitions he is mining include that courts would treat naked statistics as though there is a critical difference between naked statistics and something else. If naked statistics do not exist, what possible meaning can be given to an intuition about how courts would treat the nonexistent thing? It is like saying intuitions about the size and speed of unicorns may render insights into thoroughbreds. This is quite unlike friction that may range from a coefficient of 0.0 to 1.0. Indeed, rather than relying on people's untutored intuitions, he might have consulted the views of people who have actually studied the American court system and who fairly uniformly announce that this «intuition» is inconsistent with reality. For a discussion and citations, see Pardo, 2019: 262 n. 128.

[17] Tongue in cheek, I referred to my analysis of some of these issues as my own form of a priori analysis. I see that the joke was not understood. Making predictions as to how humans would react based on decade upon decade of observations of humans is not a priori. It is a form of clinical knowledge. I thought the point to be obvious. If Enoch and the other scholars he relies on have substantial experience with «courts» that lead to the pertinent intuitions (inferences, really), he should disclose how that can be done without observing actual courts—like those in the United States for example. It seems rather mysterious to me. Moreover, he ignores that the difficulty is not with conceptual analysis per se but with purely conceptual analysis. See Allen and Leiter, 2001.

I will put aside the internal inconsistency in arguing that this is a «modest claim» even though the appropriate incentives of the criminal law will be «seriously eroded» «at least often […] in a counter-productive way» and just stick to the facts. It is difficult to imagine a less perspicacious assertion than the claim about the American legal system that if statistical evidence is admitted, «you know that your own action has little effect on your legal fate.» A person will «know» this and have her incentives eroded even slightly, let alone seriously, only if the person is massively ignorant of the American legal system (massive stupidity by contrast is required for his blue bus arguments to go through, which I will turn to next). Buying a ticket always creates admissible evidence, from the receipt to a record of the form of payment (credit or debit card for example) to a friend who attended with the person to—I would have thought obviously—the admissible firsthand testimony of the ticket buyer. As one is walking toward the rodeo gate, reflecting on the legal system's treatment of statistical evidence, the operative incentive is to create admissible evidence so that you will not be harassed with legal proceedings concerning whether you crashed the gate. Like remember what you did.

No judge in the land would let a case survive peremptory motions that involves only naked statistics of the gatecrasher variety and any plausible contrary evidence including firsthand knowledge of the person in question, for both epistemological and policy reasons, and not because there otherwise would be incentive crushing consequences that need to be suppressed in order to secure compliance with the law.[18] This is just a case of lousy evidence caused by the plaintiff offset by a plausible explanation.[19] Nothing more complicated than that would be needed to explain the phenomenon, were it ever to arise. When the proffered evidence is not ridiculous,[20] the courts treat statistical evidence more or less like any other form—as my list of ex-

[18] Enoch might respond that the gatecrasher hypothetical includes that NO other source of evidence is admissible, including the firsthand knowledge of the defendant. That is why it is a weird hypothetical that yields no insights into the real world. To the extent that there is an interesting incentive argument here, it has to do with parties not creating ridiculous phenomena like rodeos without gates or fences and then trying to take advantage of the public. There is no doubt that the government function of evidence law often involves the effort to create appropriate incentives, but that is not support for the argument Enoch is making about naked statistical evidence. On incentives and the governmental functions of the law of evidence, see Allen, 2015. Note that DNA is not a counterexample to any of this. DNA identifies a person on occasion when there is no other evidence, and the party offering the evidence is not responsible for that state of affairs, which among other things is missing in the gatecrasher hypothetical.

[19] I am not sure that evidentiary technicalities are being kept straight, so I will simply note that relevance, admissibility and sufficiency are different matters. For another example of Enoch and his colleagues jumbling up evidence law, see Pardo, 2018: 62. Although Enoch and I reach similar results in the gatecrasher case, the routes are considerably different. He is saying that there is something in our intuitions about the case that are informative about a class of evidence—naked statistics presumably. I am saying that the hypothetical just involves lousy evidence in a context that argues for incentives to rodeo operators to organize their affairs differently—in other words, normal approaches to probative value, the production of evidence, and policy concerns generate what would be the outcome. And of course, it is absurd to think that there would be no other evidence, as I pointed out in the original article.

[20] Nonstatistical proffered evidence can be ridiculous, too, just to be clear.

amples in the original article demonstrates. That in turn explains why to my knowledge no American court has ever excluded evidence or entered a preemptory motion on any ground remotely like Enoch is suggesting is a fundamental disembodied «truth» about statistical evidence. Enoch's explication of the gatecrasher case simply reinforces the inutility of studying weird hypotheticals. There is literally no evidence supporting his intuitions other than implausible assertions about how humans in the real world (in the United States) would react.

Enoch's argument in his refutation concerning the other aspect of the incentive argument that I critiqued involving blue buses is equally curious, although for different reasons. In essence, I made two points: First, that bus drivers had to be carefully sorted out from bus companies, and second, although the analysis would differ over the two categories in neither case would one get the socially undesirable incentive effects upon which Enoch et al. based their entire argument, unless the bus companies involved were massively stupid. As for bus drivers, it is just silly to think that statistics concerning the accident rate of bus companies will have any effect on their behavior (not based on purely a priori analysis but instead the observations of human beings over a lifetime), given the other interests of the drivers at stake.

Enoch now apparently agrees with this, and thus I presume that he rejects his earlier argument that bus drivers would be noticeably disincentivized from avoiding accidents, but he goes on to say, as though he were criticizing my argument generally:

> This, however, is just not how incentive-explanations work, partly because this is not how incentives work. Sure, one way in which agents—people, organizations, etc.—respond to incentives is by explicitly deliberating about the incentives' significance for them and their interests. But it is an extremely impoverished picture of action and its responsiveness to incentives (and other reasons) to think of this as the only way in which incentives may work. Incentives serve to structure choice situations, and they infiltrate actions in other ways as well.

Reading this passage, one would think that I had neglected the bus companies, but rather than critiquing my original argument, this passage adopts it. As I pointed out, one has to take into account the adaptive nature of the phenomenon, and as again one observes over a lifetime in the economy, businesses adapt to new conditions (they are usually not massively stupid). And here the adaptation would be simple and cheap.

That is not what Enoch et al. originally thought. Inadvisedly relying on what must be one of the worst misapplications of microeconomics in the legal literature, they thought:

> In addition to these precautionary distortions, the admissibility of statistical evidence may also impair market competition: Each of the bus companies will be incentivized to hold less than fifty-one percent of the market share so as not to suffer from the evidentiary disadvantage that a larger market share imposes. Moreover, the company holding the larger market share will absorb higher liability costs, which may lead to a decrease in its market share and to a possible exit from the market.[21]

[21] Enoch and Fisher, 2015: 583, n. 83, relying on an argument made by Richard Posner.

Rather than relying on a misguided application of microeconomics to the law, Enoch et al. would have been better advised to adopt one of the basic principles of microeconomic theory that having parties internalize costs contributes to efficient outcomes. As I pointed out in the original paper, the result of this form of liability would surely be (observational data again) lower social costs due to the reduction of accidents and the correct allocation of their costs.[22] Were the law to instead suppress statistical or other evidence in order to avoid what Enoch et al. feared, the law would indeed be misguided. Which probably explains why it does not do so. At least not in the United States. [23]

In any event, all the angst over «naked statistical evidence» has yielded literally nothing in decades[24] on non-naked evidence, which is the only form of evidence that exists in the United States.[25] It has not yielded anything of juridical import on naked statistical evidence because either the intuition is wrong that the courts would reject it were it actually to exist (DNA cold hits being the best recent close example)[26] or because such evidence does not exist (DNA evidence is produced by securing evidence and testing, either of which can introduce the possibility of error, and has to be authenticated at trial).

Although I am pleased to have this opportunity to continue the dialogue with Enoch and his colleagues, I must say that I am disappointed that he did not respond to the deeper points about the nature of evidence and the differences between statistical and nonstatistical evidence. As I pointed out, the distinguishing feature of statistical and nonstatistical evidence seems to be primarily how much one knows about the evidence in question. Realistic statistical evidence[27] for the most part carries its

---

[22] For example, a cheap adaptation (cheaper than leaving the market surely) that would both reduce accidents and assist in allocating costs correctly would be the installation of cameras on buses.

[23] There are points at which Enoch's argument becomes incomprehensible. After critiquing my explanation of how incentives would actually play out in the real world, he says:

> Does Allen really want to suggest—on the basis of a priori reasoning, no less!—that nothing in evidence law supplies any incentive at all for primary behavior? Indeed, if his objection to our incentive story is to be taken seriously, he must reject anything remotely resembling law-and-economics explanations of pretty much everything, and vast parts of the social sciences.

Having already demonstrated that Enoch et al. misapplied microeconomics and that the incentive structure would differ (not disappear) from what he and his colleagues asserted, I will let this pass without further comment. As I will his statement, because I literally have no idea what it is referring to, «if Allen is sincere about his attempts to implement in legal theory the lessons of naturalized epistemology, he cannot consistently reject all attempts to incorporate philosophical insight into studying the law.»

[24] When these arguments started in the juridical context fifty years or so ago, it was a different matter.

[25] I am putting aside a different literature on Bayesian approaches to evidence that have nothing to do with the present discussion.

[26] For what might be the most recent example, see https://www.cnn.com/2020/11/19/us/norcal-rapist-roy-charles-waller-convicted-trnd/index.html.

[27] I say «realistic» statistical evidence because it is impossible to decipher the kind of assertions typical in the weird hypothetical literature, like «The witness identified the bus as blue and is 70% reliable.»

ambiguity on its surface, whereas nonstatistical evidence for the most part does not. In that sense, statistical evidence often is better than nonstatistical evidence, as its limitations are more obvious. Thus, the arguments about preferring one form of evidence to the other often seem to have it backwards in preferring the more to the less ambiguous evidence. I would love to see individuals with the talents of Enoch and his colleagues tackle that issue, including pointing out my mistakes, rather than being distracted by weird hypotheticals that ignore critical aspects of real legal systems and make impossible epistemological demands.

BIBLIOGRAPHY

Allen, R.J., 1994: «Factual ambiguity and a theory of evidence», in *Northwestern Law Review*, 88(2): 604-640.

Allen, R.J., 2011: «Rationality and the Taming of Complexity», in *Alabama Law Review*, 62(5): 1047-1068 (2011).

Allen, R.J., 2014: «Burdens of Proof», in *Law, Probability and Risk* 13(3-4): 195-219.

Allen, R.J., 2015: «A note to my philosophical friends about expertise and legal systems», *Humana. Mente Journal Of Philosophical Studies,* 8(28):79-97.

Allen, R.J., 2017: «The nature of juridical proof: probability as a tool in plausible reasoning», in *International Journal of Evidence and Proof*, 21: 133-142.

Allen, R.J., 2021: «Naturalized epistemology and the law of evidence revisited», in Quaestio facti. Revista Internacional sobre Razonamiento Probatorio, 2: 1-32.

Allen, R.J. and Leiter, B., 2001: «Naturalized epistemology and the law of evidence», in *Virginia Law Review*, 87(8): 1491-1550.

Diamond, Shari S., Murphy, Beth and Rose, Mary R., 2015: «*The "Kettleful of Law" in Real Jury Deliberations: Successes, Failures, and Next Steps*», in *Northwestern University Law Review*: 106(4) 1537-1608.

Enoch, D., Spectre, L and Fisher, L., 2012: «Statistical evidence, sensitivity, and the legal value of knowledge», in *Philosophy & Public Affairs*, 40(3): 197-224.

Enoch, D. and Fisher, T., 2015: «Sense and sensitivity: Epistemic and instrumental approaches to statistical evidence», in *Stanford Law Review*, 67(3): 557-611.

Pardo, Michael S., 2018: «Safety vs. sensitivity: possible worlds and the law of evidence», in *Legal Theory*, 24 (1): 50–75.

Pardo, Michael S., 2019: «The paradoxes of legal proof: A Critical Guide», in *Boston University Law Review, 99(1): 233-290.*

Quine, Willard V.O., 1951: «Two dogmas of empiricism», in *The Philosophical Review* 60 (1): 20-43.