Unraveling the conjunction paradox

Mark Spottswood[†]

Associate Professor, Florida State University College of Law

The conjunction paradox arises when a claim requires proof of multiple elements and the likelihood of some elements are at least partially independent of the likelihood of others. In that situation, probability theory may dictate that the conjunction of the elements is less likely than their disjunction, implying that a defendant should not be found liable, even though each element is probably true when considered in isolation. Nonetheless, American jury instructions reject this implication, and many scholars of proof have sought to construct normative theories to justify that rejection.

This article collects and critiques two families of arguments about the conjunction paradox. First, I explain why an explanatory conception of proof cannot eliminate the paradox. Second, I show why various mathematical alternatives to standard probability theory are normatively deficient when applied to legal fact-finding. Instead, I suggest that the best way to resolve the paradox is through instructions that encourage juries to make appropriate adjustments for conjunctive and disjunctive likelihoods without having to frame their analyses in mathematical terms.

Keywords: conjunction paradox; fact-finding; probability; conjunction; disjunction; jury instructions; Bayesian; inference to the best explanation; explanationism; fuzzy set theory; legitimacy; procedural justice.

1. Introduction

Ever since L. Jonathan Cohen first described it,[1] the conjunction paradox has troubled scholars of proof.[2] Stated in simple terms, the paradox arises as follows:

First, probability theory requires that, when the likelihoods of multiple events occurring are mutually independent, the likelihood that the elements occurred simultaneously is the product of their independent likelihoods.[3] For example, say that we want to know whether the outcome of three flips of a fair coin will yield three 'heads'. To determine this *conjunctive* likelihood, we would multiply the three separate probabilities of 0.5 to yield a combined probability of 0.125. More generally, probability

---

[†] Corresponding author. Email: mspottsw@law.fsu.edu

[1] L. Jonathan Cohen, The Probable and the Provable 58–67 (1977).

[2] The literature on the paradox is voluminous. *See, e.g.*, Dale A. Nance, The Burdens of Proof: Discriminatory Power, Weight of Evidence, and Tenacity of Belief 74–78 (2016); Ronald J. Allen & Alex Stein, *Evidence, Probability, and the Burden of Proof*, 55 Ariz. L.J. 557, 594–99 (2013); Edward Cheng, *Reconceptualizing the Burden of Proof*, 122 Yale L. J. 1254, 1263–66 (2013); Kevin Clermont, *Death of Paradox: The Killer Logic Beneath the Standards of Proof*, 88 Notre Dame L. Rev. 1061, 1094–105 (2013); Michael S. Pardo & Ronald J. Allen, *Juridical Proof and the Best Explanation*, 27 L. & Phil. 223, 247–56 (2007); Dale A. Nance, *Naturalized Epistemology and the Critique of Evidence Theory*, 87 Va. L. Rev. 1551, 1566–95 (2001); Saul Levmore, *Conjunction and Aggregation*, 99 Mich. L. Rev. 723 (2001); Alex Stein, *Of Two Wrongs That Make a Right: Two Paradoxes of the Evidence Law and Their Combined Economic Justification*, 79 Tex. L. Rev. 1199 (2001); Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 Harv. L. Rev. 1357, 1387–90 (1985).

[3] *See* Joseph Y. Halpern, Reasoning About Uncertainty 122–23 (2005).

theory teaches us that combinations of events will typically be less likely than any of the individual events would be considered alone, and that the more events we combine, the more dramatic the reduction will be.

Second, in most legal cases, a claimant will need to prove multiple facts or elements in order to prevail, each of which is disputed and uncertain. In some of those cases, the probabilistic product rule we just discussed will dictate that, although each of the separate elements is more likely than not to be true, the conjunction of the elements is probably false. For example, in a case with two perfectly independent elements, a fact-finder might have a confidence level of 0.7 regarding each one of those elements, but the product rule implies that she should have a confidence level of only 0.49 that both are true at once.

Finally, prevailing law appears to ignore probability theory's recommendations regarding conjunctive likelihoods, instead allowing—or often encouraging—fact-finders to find liability whenever each element is more likely than not, regardless of the likelihood of the elements' occurring together. As a result, the legal system appears to endorse findings of liability in some cases where it is more likely than not that the defendant has not acted wrongfully. This is the conjunction paradox.

The theoretic responses to the problem of conjunction can be grouped into three categories. First, some scholars have attempted to defuse the apparent paradox by arguing that juries will actually apply probabilistically correct reasoning in most cases.[4] Second, other scholars have responded to the paradox by advancing accounts of the fact-finding process that eschew any quantification of evidentiary strength, in order to avoid the apparent contradiction.[5] Finally, a third group of scholars have used the paradox to argue in favor of mathematical accounts of the proof process that do quantify inferential strength, but which employ different rules for conjoining those quantities than the standard multiplicative rule of ordinary probability theory, so as to provide a mathematical framework that can justify the apparently paradoxical jury instruction.[6]

I argue that all three responses are misguided. The paradox likely arises reasonably often in ordinary litigation. Although practices vary across jurisdictions and even between individual courtrooms, a common form of jury instruction given in many American courtrooms encourages juries to violate the conjunctive probability rules. In a subset of cases involving many elements, complicated factual theories of liability, or closely contested issues, taking these instructions literally will lead juries to find defendants liable even though they probably did no wrong. Moreover, the tendency of some theorists to optimistically assume that poorly guided jurors will nonetheless reason correctly flies in the face of ample evidence that everyday people are likely to err when faced with probabilistic reasoning tasks.

Re-describing the task of fact-finding in non-quantitative ways will not make the conjunction paradox disappear. The need to account for conjunctive and disjunctive likelihoods is inherent in the structure of reality, not a mere artifact of our descriptions of it. For this reason, any normatively attractive account of juridical inference must offer jurors a way to notice such issues and respond to them appropriately. To illustrate this, I will walk readers through an in-depth consideration of how one such approach, the explanationist account of proof offered by Ronald Allen and Michael Pardo,[7]

---

[4] *See, e.g.,* Richard D. Friedman*, "E" is for Eclectic: Multiple Perspectives on Evidence*, 87 Va. L. Rev. 2029, 2041 (2001); Dale A. Nance, *A Comment on the Supposed Paradoxes of a Mathematical Interpretation of the Logic of Trials*, 66 B.U.L. Rev. 947, 947–52 (1986).

[5] *See, e.g.*, Pardo & Allen, *supra* note 2, at 253–56; *see also* Cohen, *supra* note 1, at 217.

[6] *See, e.g.*, Clermont, *supra* note 2, at 1094–105; Cheng, *supra* note 2, at 1263–66; Nesson, *supra* note 2, at 1387–90.

[7] *See generally* Pardo & Allen, *supra* note 2.

implies that stories incorporating necessary factual conjunctions will usually have less explanatory power than similarly supported but non-conjunctive accounts, especially when the necessary conjunctions involve unusual or surprising combinations of events. Their theory should therefore agree with the standard probabilistic account and suggest that conjunctive hypotheses have systematically lowered explanatory power, particularly when the conjunctions would be described probabilistically as 'independent'. Thus, it seems that their theory is just as 'paradoxical' as the standard account, in that it likewise recommends a pattern of results that varies sharply from what American juries are encouraged to do.

Furthermore, those who advocate using different mathematical rules that more closely align with standard jury instructions risk confusing the discussion by subtly conflating descriptive and prescriptive accounts. These alternative models, in other words, leap too quickly from describing judicial practices to validating those practices as desirable. In fact, there are real reasons to worry that fact-finders would *naturally* make mistakes when considering cases involving multiple, relatively independent elements, so that they are too willing to find defendants liable in close cases. Since the mathematical fixes would operate to suggest a lowered burden of persuasion in these cases, they would seem to aggravate the problem rather than correcting it. I illustrate this with reference to three such theories that have been proposed by influential scholars: Charles Nesson's 'acceptable verdicts' theory,[8] Edward Cheng's probability-ratio theory[9] and Kevin Clermont's use of the theory of fuzzy sets to explicate burdens of proof.[10] As I will explain, each theory would be more convincing if adjusted to recommend that jurors appropriately discount the likelihood of partially independent conjunctive events.

In the end, the best way to address the conjunction paradox is not through theory, but through improved jury instructions. By instructing judges and jurors that they should only find a defendant liable when the plaintiff has met the burden of persuasion for all of the elements *taken together as a whole*, we can unravel the conjunction paradox and strike a better balance regarding the risks of errors favoring plaintiffs and defendants in our legal system. At the same time, we must be careful to instruct juries in a way that will be transparent and useful, which rules out presenting them with technical definitions of dependent and independent probabilities and expecting them to reach justifiable results through quantitative reasoning. Happily, the fact that a careful use of an explanationist approach reaches probabilistically appropriate results without quantifying likelihoods may make it possible to draft non-mathematical instructions that do away with the paradox of conjunction. To that end, I offer tentative initial suggestions regarding the form that such instructions should take, and encourage other scholars to invest their energies in identifying what type of instructions might best unravel the conjunction paradox.

2. Clarifying the problem of conjunctive proof

Before considering the varying scholarly responses to the conjunction paradox, it will be useful to lay some initial groundwork. Some readers may wonder whether the conjunction paradox is a mere theoretic curiosity or whether it gives rise to problems in real-world litigation. It will be difficult to answer these questions directly, as real-world judges and jurors do not report probability or confidence estimates regarding either individual elements or overall claims. Instead, we shall have to make do with a close examination of the tasks that judges and jurors engage in, in which we attempt to identify conditions under which they are either more or less likely to reach erroneous decisions. This

---

[8] *See generally* Nesson, supra note 2.
[9] *See generally* Cheng, *supra* note 2.
[10] *See generally* Clermont, *supra* note 2.

exploration will involve two closely related questions: First, how often do real-world cases involve a sufficient independence among theories or elements to make the conjunction paradox a potential source of decisional error? And second, how worried should we be that typical judges and jurors are in fact assigning liability without regard to conjunctive likelihoods? In this section, I shall attempt to give the best answers I can to these questions given the scanty empirical and experimental evidence.

2.1 *Understanding conjunctive and disjunctive likelihoods*

First, we must explore in more depth the notions of conjunctions and disjunctions, as well as the notions of statistical dependence and independence. To start with, consider that for any two events A and B of whose existence we are uncertain, we can ask whether B is more or less likely to have occurred if we knew for sure that A had occurred. Thus, we can ask what is the probability of B given that A is true (which we can write as P(B|A), and express the joint probability P(A&B) as follows:

$$P(A \& B) = P(A) \times P(B|A)^{11}.$$

From this, several other propositions follow quite readily. We can determine the conjunctive likelihood of two events if one only occurs in combination with the other, or where one event always excludes the other:

$$\text{If } P(B|A) = 1, \text{ then } P(A \& B) = P(A)^{12}.$$

$$\text{If } P(B|A) = 0, \text{ then } P(A \& B) = 0^{13}.$$

Between these two extremes of perfect statistical dependence, there lies the situation where two events have no relation with each other, so that knowing that one has occurred makes it neither more or less likely that the other has occurred, a condition known as statistical independence:

$$\text{If } P(B|A) = P(B|not - A) = P(B), \text{ then } P(A \& B) = P(A) \times P(B).$$

And of course, between the extremes of perfect dependence (either positive or negative) and perfect independence lie ranges of *partial independence*, in which correlations may be positive or negative but

---

[11] Using the conventional notation that represents the probability of B given A as P(B|A), we can derive the above identity as follows:

$$P(B|A) = P(A \& B)/P(A).$$

$$\text{Therefore } P(A \& B) = P(A)P(B|A).$$

See generally Halpern, *supra* note 3, at 72–73.

[12] This identity can be readily derived via substitution:

$$\text{If A implies B, then } P(B|A) = 1.$$

$$P(A \& B) = P(A)(P(B|A)).$$

$$\text{Therefore by substitution, } P(A \& B) = P(A) \times 1 = P(A).$$

[13] Once again, the derivation is straightforward:

$$\text{If A implies not-B, then } P(B|A) = 0.$$

$$P(A \& B) = P(A)(P(B|A)).$$

$$\text{Therefore by substitution, } P(A \& B) = P(A) \times 0 = 0.$$

less than perfectly so. Thus, if we were trying to estimate the joint likelihood that a randomly chosen person is female and more than 5'10" tall, we would expect that probability to be greater than zero, but also less than the product of the independent likelihoods of a randomly selected person having either quality on its own, because tallness and being female are negatively correlated. Conversely, the probability that a randomly chosen person is male and taller than 5'10" would be less than the probability of either standing alone, but quite a bit higher than the product of the two probabilities, because being male is positively correlated with height.

Finally, we must also briefly consider the concept of *disjunctive likelihoods*. Imagine that we have two events, and we are interested in knowing whether either one of them, or both together, are true.[14] As a general rule, we can determine the likelihood of a disjunction as follows:

$$P(A \, or \, B) = P(A) + P(B) - P(A \, \& \, B)^{15}.$$

For example, if we wish to identify the probability that a randomly selected person is either blue eyed, female, or both, we can simply add up the percentage of blue eyed people and the percentage of women in the overall population, and then subtract any people who have both traits (to avoid double counting). From this equation we can easily derive the two limiting conditions for perfect positive and negative correlation:

$$If \, P(B|A) = 1, \, then \, P(A \, or \, B) = P(B)^{16}.$$

$$If \, P(B|A) = 0, \, then \, P(A \, or \, B) = P(A) + P(B)^{17}.$$

Once again, there is a range of cases in the middle (since perfect positive or negative implications are rare situations), with the intermediate case of statistical independence described by the following equation:

$$If \, P(B|A) = P(B|not - A) = P(B),$$

$$Then \, P(A \, or \, B) = P(A) + P(B) - P(A \, \& \, B) = P(A) + P(B) - P(A)P(B)^{18}.$$

Last but not least, there is one final observation that will be important in the discussion below, concerning the relationship between conjunctive and disjunctive likelihoods.

---

[14] *See generally* Ray Jennings & Andrew Hartline, *Disjunction*, STAN. ENCYC. PHIL. (2008), *available at* http://plato.stanford.edu/entries/disjunction/.

[15] *See* MORRIS H. DEGROOT & MARK J. SCHERVISH, PROBABILITY AND STATISTICS 15 (3d ed. 2002).

[16] This identity can be derived as follows:
$$If \, A \, implies \, B, \, then \, P(B|A) = 1 \, and \, P(A \, \& \, B) = P(A).$$
$$P(A \, or \, B) = P(A) + P(B) - P(A \, \& \, B).$$
$$Ergo, \, P(A \, or \, B) = P(A) + P(B) - P(A) = P(B).$$

[17] This identity is also easily derived by simple substitution:
$$If \, A \, implies \, not\text{-}B, \, then \, P(B|A) = 0 \, and \, P(A \, \& \, B) = P(A) * 0 = 0.$$
$$P(A \, or \, B) = P(A) + P(B) - P(A \, \& \, B).$$
$$Ergo, \, P(A \, or \, B) = P(A) + P(B) - 0 = P(A) + P(B).$$

[18] Once again, the derivation is straightforward:
$$If \, P(B|A) = P(B|not - A) = P(B), \, then \, P(A \, \& \, B) = P(A)P(B).$$
$$P(A \, or \, B) = P(A) + P(B) - P(A \, \& \, B)$$
$$Ergo, \, P(A \, or \, B) = P(A) + P(B) - P(A)P(B).$$

$$P(A \,\&\, B) \leq min\,[P(A),\ P(B)] \leq max\,[P(A),\ P(B)] \leq P(A\ or\ B)^{[19]}.$$

To restate this equation in plain language, the *conjunction* of two events can never be more likely than the *least* likely of those events, but the *disjunction* must always be at least as likely as the *most* likely event. Ergo, adding a necessary conjunction to the description of an existing event must either decrease its likelihood or leave it the same, while adding a disjunctive alternative has the opposite effect.

2.2   *Conjunctive likelihoods in the context of litigation*

As we apply this mathematical framework to the litigation environment, two things are worth noting. First, it will be rare for many events of interest in litigation to be either completely dependent (positively or negatively) or completely independent from other events. Consider, for example, two events of interest in a typical automobile negligence suit: (1) Did the defendant drive carelessly, and (2) did the defendant's driving cause the plaintiff to be injured? The probability of (2) given (1) is certainly higher than it would otherwise be, because careless driving is more likely to lead to collisions than careful driving. But it is also true that people often drive carelessly without actually getting into crashes, so the probability of (2) given (1) is not equal to 1 (which would imply that careless driving *always* causes a car crash). Instead, it lies somewhere on the spectrum between P(2) and 1, meaning that the likelihood of a collision has risen due to the carelessness, but it has not become certain. By implication, P(1 & 2) will typically be lower than either P(1) or P(2) alone, but higher than the product of P(1) and P(2).

Second, the conjunctive likelihoods of interest in litigation will rarely involve only two events. Most civil claims or crimes have more than two elements, and there will often be situations in which multiple facts must be proven separately in order to find in a plaintiff's favor on a single element. For instance, if a plaintiff wishes to prove that a defendant was driving negligently in a car collision case by means of the testimony of a bystander eyewitness, she must prove a conjunction of several facts that are at least partially independent. First, did the bystander have a clear view of the accident? Second, does the bystander remember clearly what he saw that day? And third, is the bystander testifying truthfully to what he remembers?[20] Unless each of these facts are true, the bystander's testimony is at best unhelpful and at worst actively misleading. Ergo, the fact-finder would ideally consider, not just the individual likelihoods that the witness perceived clearly, remembered accurately, and testified truthfully, but also the likelihood that all three of those things were true at once. Conjunction, in other words, is not merely a consideration with respect to combining the likelihood of different elements of claims; it arises whenever multiple facts must be simultaneously proven before a conclusion can be reached in a case.[21]

Nevertheless, even though nearly every case will involve conjunctive proof and partial independence, the likelihood that a mistake in conjunctive reasoning will lead to an erroneous *outcome* will vary depending on some contextual factors particular to each case. First, and most obviously, we should worry more about conjunctive proof in cases with more elements or more independently necessary

---

[19] The left-hand inequality represents the upper-bound of the conjunctive probability, derived above. *See* n.12 and accompanying text, *supra*. The central inequality is trivial. The right-hand inequality represents the lower-bound of the disjunctive probability. *See* note 16 and accompanying text, *supra*.

[20] *See* Edmund M. Morgan, *Hearsay Dangers and the Application of the Hearsay Concept*, 62 Harv. L. Rev. 177, 177–78 (1948).

[21] *See also* Stein, *supra* note 2, at 1205.

facts. There is less worry about the conjunction paradox, for instance, when a plaintiff seeks to establish legal liability based upon an intentional trespass to land, because under common law principles such liability exists merely by proving two things: that the defendant entered land intentionally, and that the plaintiff possessed that land.[22] In a case with such a simple structure, a plaintiff need only prove each element to a likelihood of 0.71 to make their conjunction more likely than not, even if the elements are completely independent.[23] By contrast, consider a private actions for securities fraud under SEC Rule 10b-5, which require plaintiffs to prove six elements: (1) a misrepresentation or omission, (2) scienter, (3) a connection between the statement and the purchase or sale of a security, (4) reliance, (5) economic loss, and (6) loss causation.[24] If we were to make the same assumption regarding independence, then each element would need to be proven to a likelihood of approximately 0.9 before their conjunction becomes more probable than not![25] This example shows the speed with which conjunctive proof can become an obstacle to a claim as we add additional independent elements to be proven. And similar obstacles will arise in cases where plaintiffs seek to prove claims that have a simple legal structure by means of elaborately conjunctive factual theories.

We must remember, however, that this conjunctive escalation hinges on the extent to which elements or factual theories are independent. In many cases, theoretically separate elements will in fact tend to reinforce each other, so that conjunctive probabilities are closer to the lowest of the individual probabilities than to the product of each. Consider, for instance, a battery case in which the plaintiff claims that the defendant punched him in the face during an argument. To prevail in such a case, the plaintiff typically needs to prove that the defendant both (1) intended to cause a harmful or offensive contact with his body and (2) succeeded in causing such a contact.[26] Although I doubt that precise or reliable statistics have been collected on the subject, it seems safe to speculate that (1) people who intend to punch other people in the face are more likely to actually do so than people who have no such intent, and conversely that (2) people whose fists make contact with other people's faces are more likely to have intended that result.[27]

Now consider a rather different example: A plaintiff sues a defendant for battery, charging that the defendant, while intending to shoot an entirely different person, shot the plaintiff instead. Under the common law doctrine of 'transferred intent', such a claim could still be viable.[28] But with this new claim, we are no longer setting forth elements that go together so naturally as in the more common sort of battery claim described above. To be sure, aiming and firing a gun with the intent to strike one person does raise the likelihood of hitting a bystander, but not nearly so much as if the goal was to hit that other person. This claim comes much closer to having the truth of its elements be mutually independent, and therefore, should require a higher level of proof for each element on its own.

Finally, consider an oft-referenced 'Florida law'—actually an urban legend—which would make it a misdemeanor, punishable by a fine, for an unmarried woman to use a parachute on Sunday

---

[22] See RESTATEMENT (SECOND) OF TORTS § 158.

[23] Alternatively, the plaintiff could prove one element to a higher degree, and the other to a lesser degree. For example, if the former element were proved to a likelihood of 0.9, the conjunction would still be more likely than not if the latter element were proved to a likelihood of 0.6.

[24] See, e.g., Erica P. John Fund, Inc. v. Halliburton Co., 131 S. Ct. 2179, 2181 (2011).

[25] We must make the same caveat that some likelihoods could be lower if others were higher.

[26] See, e.g., RESTATEMENT (SECOND) OF TORTS § 18.

[27] The notion that the intent to cause a result and the actual occurrence of that result tend to be correlated is reflected in the common law maxim that a person is "presumed to intend all the natural and probable consequences flowing from his deliberate acts." See, e.g., Pearson v. Component Tech. Corp., 247 F.3d 471, 505 (3d Cir. 2001).

[28] See Vincent R. Johnson, Transferred Intent in American Tort Law, 87 MARQ. L. REV. 903, 911–12 (2004).

afternoons.[29] If a legislature were indeed to tie liability for fines to such a peculiar set of circumstances, problems with conjunction would be unusually severe, because the four elements of this offense—being female, being unmarried, parachuting, and doing so on Sunday afternoon—have no particular relation to each other. Proving each element, in other words, does almost nothing to prove the remaining ones; in fact, some may even be *negatively* correlated, such as being female and going parachuting.[30] Because of the legislature's decision to punish such a rare co-occurrence, the conjoint probability would be even lower than in the last example, and probably even lower than the pure product of each probability on its own.

We can take from these examples a more general idea: when the facts that must be proven to sustain a claim are the kind that ordinarily go together—when they follow the pattern of an everyday story, in other words—there will be less likelihood that a failure to account for conjunctive likelihood will result in an error, than would be otherwise be the case. The reverse is also true. As either the elements of a claim or the facts needed to prove it involve events that rarely go together, a failure to account for this would be especially likely to lead to an outcome error. On reflection, this should come as no surprise, because all this talk of correlations, dependence and independence is merely a statistical language we use to quantify the reality that some events tend to go together while others do not. For this reason, when legislatures or litigators choose to tie liability to unusual combinations of circumstances, it becomes more important to assess, separately from the independent likelihood of each fact by itself, the likelihood that they would arise in such an unusual combination.

To sum up, probability theory dictates that jurors should be less willing to find a defendant liable or guilty, all other things being equal, in cases where the proof against the defendant requires more rather than fewer conjunctions, where the likelihood of some of the events in question is less likely (even if more likely than not), and where some or all of those conjunctions are unusual or surprising. This leads us naturally to ask a related question: Do ordinary judges and jurors really fail to adjust for these factors when it would be appropriate to do so? Some authors have suggested that, in practice, conjunctive reasoning is less of a problem than it would seem.

2.3 *Jury instructions regarding conjunctive likelihoods*

In the legal literature, there has been some debate concerning the instructions that judges give to jurors regarding conjunctive proof, as well as the ways that jurors are likely to interpret those instructions. Dale Nance conducted what he described as an 'informal survey' of jury instructions, and came to the conclusion that most of the guidance given to jurors regarding conjunctive likelihoods was at least somewhat ambiguous. He focused on a federal pattern jury instruction, which told jurors that a plaintiff must prove 'every essential element of his claim by a preponderance of the evidence', and argued that a juror might understand this to mean either 'each' element by itself or 'all' of the elements, taken together.[31] The latter, but not the former, interpretation would

---

[29] *See, e.g.*, KHSB.com, *No Unmarried Woman Can Parachute on Sundays Is Just One of Florida's Most Bizarre Laws* (Nov. 22, 2011), *available at* http://www.kshb.com/news/ local-news/water-cooler/no-unmarried-woman-can-parachute-on-sundays-is-just-one-of-floridas-most-bizarre-laws. No such law can be found among Florida's codified statutes.

[30] *See* Jennifer Olson, *Support for Women in Skydiving* (3 December 2014), *available at* http://www.womensadventuremagazine.com/extreme-outdoors/support-women-skydiving/ (noting that only 13.8% of licensed Australian skydivers are female).

[31] Nance, *supra* n.4, at 950–51 (citing Devitt & Blackmar, Federal Practice and Jury Instructions § 71.14 (3d ed. 1977)). Nance gives another example in which the court would tell the jury that the plaintiff is responsible for proving 'the following facts: *First*, that the defendant was negligent . . .; and *second*, that the defendant's negligence was a proximate cause of some

correctly account for conjunctive likelihoods. Others have agreed with Nance that many such instructions are ambiguous.[32]

Ronald Allen and Sarah Jehl took issue both with the implementation of Nance's survey and his interpretation of its results. They point out that other federal pattern instructions from the same source that Nance relied upon tell the jurors that if the plaintiff has proved 'each of these elements by the preponderance of the evidence', then the jury 'should return a verdict for the plaintiff'.[33] If this instruction were given in addition to the instruction quoted above, it would seem to reduce the ambiguity and reinforce the paradox. To be sure, it is at least possible that some statistically sophisticated jurors might realize that, as more elements accumulate, the probability that each was true *simultaneously* was lower than each of them being true *individually*, and might decide to adjust their decision-making accordingly. But to do so would be to go against the grain of the instruction, which suggests that one may stop once 'each' element has been sufficiently proven in isolation. Allen and Jehl conducted a broader survey of instructions, and found that the emphasis on proving 'each' element is common in both pattern jury instructions and actual practice.[34] By contrast, they did not locate any instructions that either required or encouraged the jurors to think separately about conjunctive likelihoods in cases involving multiple contested elements.[35]

Finally, there is one additional area in which the paradox plays an unusually explicit role: the use of special verdict forms under Federal Rule of Civil Procedure 49(a) or similar rules at the state level. That device permits a judge to ask a jury to answer factual questions regarding separable elements of the case, and then to enter a judgment in favor of the plaintiff if the jury has entered pro-plaintiff findings on each separate question. Some judges approve of this device precisely because it avoids the possibility that the jury might find in a plaintiff's favor on each element but then award the overall verdict to the defense,[36] even though such a decision would be probabilistically appropriate in many close cases involving partially independent proof of multiple elements. So on a closer look, it would seem that most American jury instructions are at best ambiguous regarding conjunction, and that the general tendency is to encourage jurors to think only about the separate likelihood of each element, while ignoring the possibility that the likelihood of the conjunction might be lower.

2.4 *Jurors' intuitions about conjunctive likelihoods*

That still leaves a second possibility, also raised by Nance, which is that jurors naturally and intuitively adjust for conjunction even when judges given them no encouragement or instruction to do so. As

---

injury . . . sustained by the plaintiff.' *Id.* (citing Devitt & Blackmar, *supra* this note, at § 80.17). Like the example discussed above, this could potentially be understood either to require assessment of individual or conjunctive likelihood.

[32] *See* Levmore, *supra* n.2, at 724–25 note 1; *see also* Friedman, *supra* note 4, at 2041.

[33] *See* Ronald J. Allen & Sarah A. Jehl, *Burdens of Persuasion in Civil Cases: Algorithms v. Explanations*, 2003 L. Rᴇᴠ. M.S.U.-D.C.L. 893, 900 (2003) (quoting Edward J. Devitt et al., 3A Federal Jury Practice and Instructions §83.02 (4th ed. 1987)).

[34] *Id.* at 900–02.

[35] *Id.* at 902. Nance briefly returned to this debate as an aside to his broader project of using the concept of Keynesian weight to explicate burdens of proof. In this more recent treatment, he reiterates that some instructions are ambiguous, but does not point to any instructions that unambiguously require jurors to analyze conjunctive likelihoods separately from individual element likelihoods. Instead, he now emphasizes that judges and rule-makers are simply unaware of the problem, and thus any suggestion given by the rules is accidental rather than intentional. *See generally* Nance, *supra* note 1, at 74–78.

[36] See e.g.., Hon. John R. Brown, *Federal Special Verdicts: The Doubt Eliminator*, 44 F.R.D. 245, 339–40 (1967) (noting that to ask for a general verdict as well as answers to specific questions offers 'nothing but trouble' due to the 'high likelihood of a conflict which extinguishes both').

Nance puts it, '[i]t is not at all uncommon to have reasonably correct intuitions but to be unable to explain them'.[37] Indeed, modern psychological theory does suggest that, for some types of problems, people can often reason better by relying on common sense and intuition than they would if they went through a more deliberative process of decision-making.[38] But we should not be too hasty to assume that all is well, because the same body of research also shows that our intuitions are often mistaken, and that these mistakes can take a predictable form.[39]

I have written elsewhere about the modern framework by which psychologists understood the interaction between intuition and reason, and the ways that either mode of thinking can give rise to errors in the legal process,[40] so I will not belabor those subjects here. Nevertheless, there are two important principles we can draw on to help assess the likelihood that jurors' intuition adjusts for the deficiencies in formal jury instructions regarding conjunctive likelihoods.

As a general matter, people's intuitive judgments are often fairly unreliable when it comes to reasoning about conjunctive likelihoods. One of the original errors of judgment shown in Tversky and Kahenman's classic investigation of heuristics and biases was the 'Linda problem', in which participants first read a general background description of a hypothetical person named Linda and then rated the likelihoods of a number of additional statements about that person. Two statements that subjects were asked to consider were 'Linda is a bank teller' and 'Linda is a bank teller and is active in the feminist movement'. Since the latter statement is conjunctive in form and contains the first state-ment as half of the conjunction, its likelihood must be less than or equal to the likelihood of the former statement. Nevertheless, *eighty-five percent* of their participants rated the latter statement as more likely than the former—even when they tested doctoral students who had taken advanced courses in probability and statistics![41]

Nor is this particular experimental result an unusual outlier. In fact, a wide array of experiments have illustrated that people often fail to conform their judgments to basic principles of normative rationality. Thus, for instance, people also tend to neglect whether an occupation is widespread or rare when estimating the likelihood that a described person works in that field, even when they have been warned that the particularized evidence they had received about the individual's preferences was unreliable.[42] Moreover, when Gettys and his co-authors tested participants on their ability to estimate one uncertain parameter and then use it to estimate a second in a more abstract reasoning task, they observed a marked tendency for their participants to act as if their 'best guess' at an earlier stage was

---

[37] Nance, *supra* note 4, at 952.

[38] *See* Gerd Gigerenzer, Gut Feelings: The Intelligence of the Unconscious 7–8, 27–30 (2007) (describing studies in which intuition outperformed analysis); *see also* Daniel Kahneman, Thinking, Fast and Slow 240–42 (2012) (describing how experts can acquire very accurate intuitive methods for solving highly complex problems).

[39] *See* Kahneman, *supra* note 38, at 25 (noting that our intuitive reasoning system is prone to make 'systematic errors . . . in specified circumstances' and that it 'has little understanding of logic and statistics').

[40] *See, e.g.,* Mark Spottswood, *The Hidden Structure of Fact-Finding*, 64 Case W. Res. L. Rev. 131, 171–93 (2013); Mark Spottswood, *Bridging the Gap Between Bayesian and Story-Comparison Models of Juridical Inference*, 13 Law, Probability & Risk 47, 47–49 (2014).

[41] Kahneman, *supra* note 38, at 156–57. Gigerenzer criticized the original experiment on the ground that some subjects may have understood the isolated statement 'Linda is a bank teller' to include an implied negation of her participation in the feminist movement. *See* Gerg Gigerenzer, *On Narrow Norms and Vague Heuristics: A Reply to Kahneman and Tversky*, 103 Psychol. Rev. 592, 593 (1996). Notably, however, subsequent studies found similar errors when asking questions that avoided this and other potential problems with the original design. *See* Kahenman, *supra* note 44, at 161–63 (reviewing alternative formats that also generated high rates of conjunction errors).

[42] *See* Daniel Kahneman, *Maps of Bounded Rationality*, 93 Am. Econ. Rev. 1449 (2003) (describing the 'Tom W.' experiment).

certain when moving on to subsequent stages of the task, indicating that serious conjunction errors can occur even when the probabilistic nature of the task is made explicit.[43] In short, our gut intuition can be quite good for solving a variety of tasks, but it seems particularly unreliable when it comes to complex decisions that incorporate multiple kinds of uncertainty.

At the same time, there is some evidence that these defects are not inherent in human reasoning, but instead reflect our limited ability to use intuition to solve unfamiliar types of problems. When abstract reasoning tasks are represented in ways that map onto problems we might encounter in daily life, researchers find that people do a markedly better job at avoiding logical errors.[44] Likewise, some common errors in statistical reasoning can be significantly moderated by presenting information in more intuitive formats, such as by using frequencies rather than percentages.[45] This may provide some reassurance, but we should remind ourselves that standard jury instructions do not encourage jurors to consider their uncertainty about potential outcomes in terms of the frequencies with which they will make mistakes across large numbers of similar cases.

Given the unusual nature of the task we ask jurors to perform when deciding cases, it seems farfetched to assume that they will deftly adjust for a mixture of partially dependent probabilities by intuition alone. After all, the multiple forms of uncertainty at play in a typical car accident may fall within the realm of common sense, but it is far less likely that jurors will find a complex business dispute or an antitrust case to be similarly transparent. What is worse, even when jurors encounter situations, like a car crash case, that do fall within the realm of everyday experience, they probably lack the sort of experience that would make their probabilistic judgments well-calibrated. After all, whether or not a situation is familiar, our estimates of uncertain likelihoods usually arise without the provision of prompt feedback regarding what actually happened, and in the absence of such feedback it will be hard for our intuitive judgments to become well-tuned.[46]

To sum up, the law often requires that jurors combine multiple uncertain judgments before they can issue a verdict. Because such judgments are usually at least partially independent, normative rationality would suggest that jurors should discount the likelihood of their conjunction to account for this partial independence. But standard jury instructions rarely discuss how jurors should go about doing this, and often seem to suggest that accounting for the independent likelihoods of elements on their own is all that is necessary to reach a decision. Nor should we be confident that jurors instinctively do this without help or encouragement from judges, because psychological experiments have shown that ordinary people often fail to reason appropriately when incorporating multiple levels of uncertainty into their decisions, especially when the type of decision they must make is an unfamiliar one and they lack explicit guidance or knowledge regarding the proper decision-making procedure.

---

[43] *See* Charles F. Gettys et al., *The Best Guess Hypothesis in Multi-Stage Inference*, 10 ORG. BEHAV. & HUMAN PERF. 364 (1973).

[44] *See* Leda Cosmides & John Tooby, *Cognitive Adaptations for Social Exchange*, *in* THE ADAPTED MIND: EVOLUTIONARY PSYCHOLOGY AND THE GENERATION OF CULTURE 163, 182 (Barkow, J. et al. eds., 1992). More recent research suggests that the impacts of frequency formats may be situational, depending on their ability to make class-inclusion relations more transparent in a particular context. *See* Valerie F. Reyna & Charles J. Brainerd, *Numeracy, Ratio Bias, and Denominator Neglect in Judgments of Risk and Probability*, 18 LEARNING & INDIVIDUAL DIFFERENCES 89, 93-94 (2008).

[45] Leda Cosmides & John Tooby, *Are Humans Good Intuitive Statisticians After All? Rethinking Some Conclusions from the Literature on Uncertainty*, 58 COGNITION 1, 22–37 (1996); KAHNEMAN, *supra* note 38, at 163.

[46] KAHNEMAN, *supra* note 38, at 241; PHILIP E. TETLOCK & DAN GARDNER, SUPERFORECASTING: THE ART AND SCIENCE OF PREDICTION 180–84 (2015).

3. Addressing conjunctive proof within an explanationist framework

The prior section discussed conjunctive proof within a probabilistic framework, but one need not invoke probabilities to discuss how jurors can or should conduct the task of fact-finding. A competing mode of describing the proof process eschews any attempt to quantify the likelihood of individual events, and instead focuses on the relations between the overall field of evidence items and the comparative strength of the parties' competing *explanations* of those evidence items. The leading advocates of this approach contend that it is superior to a probabilistic account on a number of grounds, one of which is that it eliminates the conjunction paradox.[47]

Contrary to this claim, I will seek to show in this section that a careful application of explanationist theory leads to the conclusion that conjunctive explanations should typically be penalized in the inferential process compared with simpler or disjunctive explanations, in a manner that closely parallels the recommendations of classical probability theory. Moreover, in cases where the plaintiff must prove a conjunction of multiple elements to prevail, their explanations will need to contain facts that cover each of those elements, implying that plaintiffs will typically offer conjunctive explanations of the evidence while defendants may offer simpler, or even disjunctive, explanations. Unfortunately, jurors will often fail to adjust for conjunctivity when assessing comparative explanatory strength. What is more, an explanationist theory cannot justify that refusal, and it gives us no warrant to treat such refusals as an epistemic ideal. Therefore, an explanatory conception of proof may give us a different way to describe the mismatch between what jurors tend to do and what they ought to do when reasoning about conjunctive likelihoods, but it cannot make the mismatch go away.

3.1   *The explanationist model of juridical proof*

Although others have explored similar terrain,[48] Ronald Allen and Michael Pardo have been the leading advocates of an explanation-based account of trial inference as both descriptively accurate and normatively desirable. Their framework borrows liberally from the idea of 'inference to the best explanation' in the philosophy of science, which describes the process of choosing among competing theories as a search for the theory that best explains the existing data.[49] According to Allen and Pardo, jurors do not compute the likelihoods of each of the disputed facts under consideration in a typical trial, because such a task is neither tractable nor necessary. Instead, jurors compare the broader theories of the case offered by each party to the entire field of the evidence, and weigh the comparative explanatory power of the competing accounts. In a civil case, the jury would then issue a verdict in favor whichever party offers the best available explanation of the evidence (or, if the best explanation is not one provided by either party, in favor of whichever party the best explanation favors).[50] The process is similar in a criminal case, except that the defendant can prevail with a weaker explanation, so long as it is a plausible one, so as to accommodate the 'beyond a reasonable doubt' standard of proof.[51]

---

[47] *See, e.g.*, Michael S. Pardo, *Group Agency and Legal Proof: Or, Why the Jury is an "It"*, 56 William & Mary L. Rev. 1793, 1831 (2015).

[48] *See generally* Amalia Amaya, *Justification, Coherence, and Epistemic Responsibility in Legal Fact-Finding*, 2008 Episteme 306; David A. Schum, *Species of Abductive Reasoning in Fact Investigation in Law*, 22 Cardozo L. Rev. 1645 (2001); John R. Josephson, *On the Proof Dynamics of Inference to the Best Explanation*, 22 Cardozo L. Rev. 1621 (2001).

[49] *See* Peter Lipton, Inference to the Best Explanation 55–57 (2d ed. 2004).

[50] *See* Allen & Pardo, *supra* note 2, at 233–35.

[51] *Id.* at 238.

The choice among competing explanations, as Allen and Pardo explain it, would generally be guided by similar factors to those identified in the philosophy of science for choosing among rival scientific theories.[52] Among the relevant considerations, a good explanation should be *consistent*, both internally and with the field of available evidence.[53] Second, all other things being equal, *simple* explanations are preferable to more complex explanations.[54] And thirdly, good explanations will generally be *coherent* with our background beliefs about the way that the world usually works.[55] As the authors repeatedly remind readers, there is no precise formula for combining these criteria; rather, the selection of the best among competing explanations is by necessity a holistic task that balances among each kind of consideration.[56]

3.2. *Pardo and Allen's response to the conjunction paradox*

Pardo and Allen have argued that one of the virtues of the explanatory approach to fact-finding is that it 'avoids' the conjunction paradox.[57] As they see it, a juror need merely consider which of the competing explanations is more plausible based on all the evidence, and then assign liability based on whether that explanation satisfies all the elements of a legal claim. Therefore, there is no stage at which a juror would need to separately consider the independent likelihood that each element would be true in isolation, or combine those likelihoods into a broader judgment regarding liability.[58] Moreover, the authors do not consider the avoidance of these issues a mere detail of their theory; rather, the fact that these issues must be considered in a probabilistic framework, but not in their own, is taken as a reason to reject the former theories and adopt the latter.[59]

To Pardo and Allen, the need for a theory that avoided the conjunction paradox is driven by the fact that they found both of the main existing solutions to the paradox unsatisfactory. One option, which is to follow the grain of standard jury instructions and permit jurors to assess liability based on sufficient proof of each element in isolation, leads to normatively undesirable consequences. In particular, they worry that this approach 'does not distribute errors evenly among parties and therefore is unlikely to lead to accurate outcomes'.[60] To illustrate the strange consequences of such a rule, they compare two hypothetical cases: the first involving two elements proven to 0.6, in which a plaintiff would prevail despite a conjoint likelihood of 0.36, and the second involving probabilities of 0.9 and 0.5, in which the plaintiff would lose despite a strictly *higher* conjoint likelihood of 0.45. The standard legal answer regarding conjunction, therefore, is normatively unsatisfactory.

The authors find the alternative answer to the conjunction paradox, which would be to require juries to assess a conjunct likelihood above 0.5 before finding a defendant liable, equally unsatisfactory. They note several problems with this approach. For one thing, it is inconsistent with the standard jury

---

[52] *Id.* at 235.

[53] *Id.* at 230.

[54] *Id.*

[55] *Id.* The authors also mention the idea that a good explanation should not seem too '*ad hoc*', suggesting that an explanation might also be penalized if it seemed to be artfully designed to avoid disconfirmation.

[56] *Id.*

[57] *Id.*; *see also* Pardo, *supra* note 47, at 1831; David A. Lombardero, *Do Special Verdicts Improve the Structure of Jury Decision-Making?*, 36 JURIMETRICS 275, 285–86 (1996) (agreeing that a relative plausibility approach 'eliminates the conjunction problem').

[58] Pardo and Allen, *supra* note 2, at 255–56.

[59] *Id.* at 261–62.

[60] *Id.* at 254.

instructions discussed above.[61] And for another, it would mean that plaintiffs would need to prove each element to a higher probability as more elements were added to a claim, which they find counter-intuitive.[62] Therefore, since either answer to the problem of conjunction leads them to either descriptive or normative difficulties, they prefer to advocate for an approach that 'neutralizes' the paradox by requiring no choice between the alternatives.[63]

3.3.   *The benefits of separating normative and descriptive theory*

Some of the difficulties Allen and Pardo face in this context may arise, not from anything inherent in the mechanics of conjunctive reasoning, but by their attempt to craft one theory to cover multiple kinds of ground at the same time. First, they wish to describe the ways that judges and juries typically reason.[64] Second, they seek to develop a theory that is (mostly) consistent with existing legal rules and institutions.[65] Third, they wish to propose a framework that can also serve as an epistemic guide towards normatively justified decision-making, so that deviations from the theory in the existing judicial process can be identified and excised.[66]

By contrast, I believe that we can gain far more clarity regarding the problem of conjunction if we are first willing to admit the possibility that descriptive reality and normative aspirations might not neatly align with one another. To be sure, the comforting assumption that the way that jurors actually reason lines up closely with the best way that they could reason is alluring. Nevertheless, we have good reasons to worry that sometimes judges and jurors make decisions that are hard to defend. For one thing, there is a large body of psychological research showing that the average person is likely to make certain kinds of predictable analytic mistakes,[67] particularly when they lack domain-specific expertise.[68] Such research would suggest that jurors will make certain kinds of predictable reasoning errors in real-world cases. What is more, the rise of DNA-based forensic identification has given us a window into certain causes of mistaken verdicts that were previously harder to detect, such as jurors' excessive trust in the testimony of eyewitnesses[69] or their unwillingness to disbelieve confession evidence even in the face of strong counterproof.[70] Since we are faced with apparent gaps between actual and ideal courtroom decision-making, the prudent thing to do would be to separately develop normative and descriptive accounts of juridical inference, so that we can use the former to improve the latter.

With that separation in mind, we can sharpen our inquiry and ask two different kinds of questions. The first question is a normative one: If we are to trust the philosophical theories of inference to the best explanation as a normative guide to high-quality decision-making, what do those theories have to say about the comparative explanatory power of conjunctive and disjunctive explanations, respectively? Alternatively, we might ask a descriptive question: When people assess the relative plausibility of competing explanations in relation to a field of evidence, do they appropriately adjust their

---

[61] *Id.* at 253–54.

[62] *Id.* at 255.

[63] *Id.* at 256.

[64] *Id.* at 235.

[65] *Id.* at 268.

[66] *Id.* at 260–61, 263–67.

[67] *See* Kahneman, *supra* note 38, at 12–13.

[68] *See* Keith E. Stanovich, Rationality and the Reflective Mind 101–05 (2011).

[69] *See* Jon B. Gould et al., *Predicting Erroneous Convictions: A Social Science Approach to Miscarriages of Justice* 76 (working paper); Brandon L. Garrett, Convicting the Innocent: Where Criminal Prosecutions Go Wrong 48 (2011).

[70] *See* Dan Simon, In Doubt: The Psychology of the Criminal Justice Process 121 (2012).

judgments of explanatory power based on the presence of conjunctions or disjunctions among the elements of either explanation? Only if we are sure that both questions have similar answers should we be confident that an explanation-based account of inference will avoid the conjunction paradox.

3.4. *Why explanationists should penalize conjunctive explanations*

Let us start with the normative question. Among the relevant criteria available for weighing the plausibility of a party's explanation of the evidence in a given case, it would seem that several are indeed affected by the presence or absence of partially or completely independent elements in that explanation. For instance, when an explanation incorporates conjunctions among facts that have no particular reason to go together (in other words, facts whose likelihoods would be thought of as probabilistically independent), those coincidental features of the explanation will mean that it is less internally consistent than an otherwise similar explanation that is non-conjunctive. Moreover, to the extent that the rules of probability theory are meant as mathematical generalizations of ordinary experience, the presence of relatively independent conjunctions within a story might also affect its coherence with our background beliefs. Finally, and perhaps most controversially, the presence of necessary conjunctions within an explanation should also undermine its claim to be simpler than competing explanations that either avoid conjunctions or are disjunctive in form.

3.4.1 *The criterion of internal consistency.*

Regarding internal consistency, one possibility would be to say that conjunctive explanations are always more internally consistent than disjunctive explanations, because conjunctions stand or fall as a unit while disjunctions need not do so. But on closer examination, this is not a very plausible position. To be sure, disjunctive explanations may contain components that are mutually exclusive, but this need not be the case, and even when disjunctions are exclusive they should not lead us to downgrade our trust in an explanation's power.

Consider, for example, the example of an automobile crash case in which the plaintiff claims the defendant drove through a red light to strike him while he was in an intersection, but in which the defendant counters with a disjunctive theory: perhaps the light was green (according to the defendant herself) or perhaps the light was yellow (according to a bystander witness), but either way the defendant had the right of way in that particular intersection.[71] To be sure, the testimony of each defense witness tends to disprove the other if believed, and they are to that extent inconsistent. Nevertheless, it is also clear that either witness's testimony is equally antagonistic to the plaintiff's own account. There would seem to be little license, in this sort of situation, to view the inconsistency within the defendant's case as a reason to treat either witness as less believable than the plaintiff; at the least, it would seem that the rational thing to do is to place your confidence in whichever witness presents the most credible account, without regard to which party has offered their testimony, and then use that explanation as a basis for choosing the proper verdict. In this scenario, then, it would seem that a reasonable approach is to consider the plausibility of the defendant's overall case to be at least as great as the *stronger* of the two disjunctive possibilities.

So once we grant (as Allen and Pardo do)[72] that there is no rational reason why parties should not be able to offer disjunctive explanations, there seems little reason to view those explanations as

---

[71] *See* Nance, *supra* note 2, at 1578–79 (discussing a similar hypothetical); *see also* Pardo & Allen, *supra* note 2, at 250–52 (responding to Nance's discussion).

[72] *See* Pardo & Allen, *supra* note 2, at 251–52 & n. 8.

'inconsistent' in a way that puts them on a systematically weaker footing than other kinds of cases. Rather, the kind of 'inconsistency' that should lead to an inferential penalty would be when the defendant's explanation incorporates elements that are less consistent with *each other* than with any corresponding part of the plaintiff's own case. For example, imagine that our car-crash defendant maintained simultaneously that he was not involved in any car crash, but that the car crash he was involved in was the plaintiff's fault. In such a bizarre case, a jury could very rationally determine not just that the theories were inconsistent with each other (as in the green light/yellow light hypothetical), but that any story offered by the defendant was worthy of little credence, given that an inability to offer a consistent position as to whether one had been in an accident illustrates a critical defect in either memory or sincerity. For that reason, the defendant's disjunctive explanation would tend to disprove its alternative options more readily than it could disprove the plaintiff's account. Cases like this alternative hypothetical, however, would be rare indeed, because defendants would usually know whether they had been in a crash or not, and because any competent lawyer would warn the defendant that they had little to gain, and much to lose, by offering such a bizarre account of the facts. Thus, although it is possible for disjunctive defenses to combine contradictory stories in a way that lacks internal consistency, in the main defendants would be likely to offer alternative accounts only when doing so does not destroy their credibility.

So if disjunctive explanations should not generally decrease our judgments of internal consistency, what then of their converse, explanations that require us to believe that multiple facts are simultaneously true? We must tread carefully here, because one way of thinking about this is initially attractive but wrong. Imagine for a moment that we were to take the prior example and add a second plaintiff's witness, a bystander who will testify that the light was red. Perhaps, we might think, two consistent witnesses reinforce each other, while two inconsistent witnesses undermine each other, and thus the consistent combination of two witnesses should receive a plausibility boost from the apparent 'conjunction'. But we must keep in mind that when we ask if an explanation is internally consistent, we are not asking if the evidence that supports it is consistent, but rather if the hypothetical story that it tells about the world is consistent. Here, both witnesses are telling the same story, so the explanation itself is not conjunctive; whether there is one witness or two, the plaintiff's story is still 'the light was red rather than green or yellow'. Adding another witness does strengthen the explanation, but it does so by way of improving its base of support in the evidence, not by making the explanation itself conjunctive.

So for an example of a truly conjunctive explanation we need, not multiple witnesses telling consistent stories, but rather a theory whose efficacy hinges on assuming that multiple, separable facts are simultaneously true. This will be a very common situation; after all, in order to cover all of the elements of a tort, a plaintiff will typically need to provide an explanation that speaks to each element. Thus, to prove that a defendant committed the tort of negligence by hitting her with his car, it would be necessary to offer an explanation that includes the fact that he struck her with his car (to cover the element of causation), facts suggesting that he was driving carelessly (to cover the element of breach of duty to exercise appropriate care) and the fact that she suffered injury as a result (to cover the element of damages).[73] If any of these facts were omitted from the explanation, the overall explanation would no longer suffice to allow the plaintiff to prevail.[74] Therefore, we would say that this explanation is

---

[73] *See, e.g. Hargrove* v. *Roman*, No. CV126032427, 2015 WL 5626182, at *3 (Conn. Super. Ct. 20 August 2015).

[74] Note that the key here is not conjunction among the elements, but among the facts included in an explanation that attempts to cover them. In some cases, one explanatory fact might be used to prove several elements. For instance, in a defamation case, the fact that the defendant addressed an unflattering statement about the plaintiff's sexual character to a large room of her associates would likely contribute to proving that the statement was defamatory, that it was published, and that the publication

conjunctive in nature. What is more, the commonplace nature of the example should help to show that even if a winning explanation need not be conjunctive in every possible case,[75] conjunction will be the rule rather than the exception.

What, then, can be said of the effect of adding conjunctions to an explanation on our judgments of that explanation's internal consistency? Broadly speaking, it would seem that these judgments might actually track the proscriptions of probability theory quite closely. Take, for instance, a set of events whose likelihoods might be judged to be quite dependent on one another, in the sense that the occurrence of any makes the others more likely: (1) The defendant formed a passionate hatred of the victim, to the extent that he wanted to kill him, (2) the defendant entered the victim's home with a firearm, (3) the defendant intended to shoot the victim, (4) the defendant did so, and (5) the victim died. Even though this story is conjunctive, we would likely say that the likelihoods are far from independent, because intending to murder and causing death tend to go hand-in-hand. Similarly, I think most of us would also judge the story to be internally consistent, as each part tends to support the others rather than undercut it.

If we tweak the story to make its parts more probabilistically independent, it would seem to also lose internal consistency, as well as its coherence with everyday experience. For example, consider a second murder charge, this time based on the doctrine of felony murder. Here, the explanation might run as follows: (1) The defendant and his friend wished to steal money from a would-be drug buyer. (2) In accordance with this plan, the defendant pointed a shotgun at the buyer and demanded that he give them money without receiving drugs, but (3) the buyer was actually an undercover police officer, who had a concealed weapon. (4) When threatened, the officer drew his weapon and fired, but missed the defendant. (5) The officer instead struck the defendant's partner in crime, and (6) the co-conspirator died of his wounds.[76] In a case like this, the conjunctions no longer seem to flow together in the same way as in the former example. Rather, there are significant internal tensions among some of the explanatory facts. Most of the people who set out to rob potential drug buyers probably seek to avoid targeting undercover officers, for instance, just as most officers seek to hit their target when firing their weapons, rather than a bystander. These tensions do not make the story impossible, but they do make it somewhat odd. Certainly we no longer see each part of the story as mutually reinforcing the other parts. The separate parts of this explanation, in other words, lack consistency with each other, and therefore we should judge this story as less explanatorily powerful than the story in the previous example, even assuming that all else was equal (such as the evidential support for each individual story item). The key point is that although we have changed the language with which we talk about the problem, the presence of facts with a large degree of probabilistic

---

was intentional or reckless. *Cf.* RODNEY A. SMOLLA, 1 LAW OF DEFAMATION § 1:34 (2d ed.) (1999). Conversely, in other cases a conjunction of facts will be needed to prove a single element. Thus, to prove in a securities fraud case that a statement caused the price of a stock to fall, facts must be provided showing the price before the statement was made as well as the price afterwards, with still further facts being necessary to show that the change was not due to unrelated fluctuations in the market. *See Dura Pharm., Inc.* v. *Broudo*, 544 U.S. 336, 347 (2005).

[75] Non-conjunctive explanations are possible, depending of course on some allowance for levels of granularity in defining what constitute separable facts. An alibi defense to a charge of murder in New York City, for instance, might include an explanation as simple as 'The defendant was in the City of Chicago on the date and time of the murder.' To be sure, the facts required to support this explanation might be complex, but the explanation itself boils down to a single item of information. It will be far easier, however, for defendants to offer non-conjunctive explanations than for plaintiffs to do so, given the plaintiff's need to explain multiple elements of a claim.

[76] These facts are based loosely around a case in which the Illinois Supreme Court upheld a conviction for felony murder. *See People* v. *Dekens*, 695 N.E.2d 474, 475 (Ill. 1998).

independence within an explanation should usually lower our assessment of that story's internal consistency.

3.4.2 *The criterion of consistency with background beliefs.*

Similar points can be made regarding another criteria for preferring explanations over others, which is an explanation's consistency with a fact-finder's background beliefs about how the world works. To some extent, this criterion must overlap strongly with judgments of internal consistency as well as with a story's consistency with the evidence. After all, we cannot arrive at a judgment that some explanation items and some evidence items naturally go together by means any known algorithm; we must draw, instead, on our life experiences and our common sense. To that extent, stories that incorporate what would be called 'independent conjunctions' in the language of probability will be penalized when assessed for consistency with background beliefs, for the reasons already explained. And furthermore, reference to background beliefs might help reinforce the notion that we must be cautious before assuming that coincidences are real rather than fictitious, as such stories may remind us of 'tall tales' rather than everyday life experience.

To be sure, if we were discussing *descriptive* uses of an explanatory theory at this point, we would have to worry about the very real possibility that fact-finders might have some strongly rooted background beliefs that are irrational, and thus fail to notice that some combinations are surprisingly coincidental. A devotee of conspiracy theories, for instance, might easily fail to notice that a party's case required them to believe that a number of independent individuals were acting in concert to accomplish an unlawful end, because that person might already believe in many other similar accounts. But at least so long as we are focusing on the *normative* use of the theory, we may assume that an ideal epistemic agent would not be so cavalier. Such a person would, no doubt, be careful about confusing what is reported in the news for a representative sample of reality, and would be doubly cautious before assuming that fictional stories that are structured in order to garner attention and interest are always a good model of the real world around them. The normative version of the explanationist theory, therefore, would posit that a fact-finder's background beliefs should be a good reflection of real world frequencies and occurrences. To such a person, common fact-patterns should be relatively easier to believe, while uncommon ones should require a bit more evidence to be equally convincing.

3.4.3 *The criterion of simplicity.*

What, then, can we say of the relationship between conjunction of probabilistically independent events within an explanation and the simplicity of that explanation? Here, we must tread more carefully, because (somewhat ironically) the philosophical concept of explanatory simplicity is surprisingly complex and its content is frequently disputed.[77] Some conceptions are *syntactic*, in that they tie the simplicity of an explanation to the brevity with which it can be stated.[78] Some are *semantic*, tying our judgments of simplicity instead to the number and types of ontological entities that an explanation posits, rather than to the length of the explanation itself.[79] And

---

[77] *See* ELIOTT SOBER, SIMPLICITY vi (1975).

[78] *See, e.g.*, Willard Van Orman Quine, *On Simple Theories of a Complex World*, 15 SYNTHESE 103 (1963). For attempts to develop more formalized conceptions of syntactic simplicity, *see, e.g.,* Paul Thagard, *The Best Explanation: Criteria for Theory Choice*, 75 J. PHIL. 76, 86–89 (1978); Nelson Goodman, *The Test of Simplicity*, 128 SCIENCE 1064 (1958).

[79] *See, e.g.*, Alan Baker, *Quantitative Parsimony and Explanatory Power*, 54 BRITISH J. PHIL. SCI. 245, 245–46 (2003); SOBER, *supra* note 77, at 40–44.

still others see simplicity as a pragmatic principle that helps us avoid mistaking noise for signal and developing too-detailed models in the face of messy information.[80] Given this theoretical morass, I will try to avoid any claims in this section that require committing to a particular philosophical theory of simplicity; instead, I shall endeavor to show that some common themes in the simplicity literature might encourage us to penalize conjunctive explanations and give extra credence to disjunctive ones.

One way of deciding whether a theory is simple is to ask whether it can be briefly stated. Under this approach, which some have equated to analyzing a theory's *elegance*, we can simply compare the number and length of the hypotheses posed by two competing explanations, and say that the one that explains the same set of facts more briefly is the simpler of the two.[81] Of course, to do this kind of inquiry fairly we must be able to agree in advance on a common syntactic language in which to describe two competing theories; otherwise, it would be all too easy to redefine the terms of one to be brief and the terms of the other to be more verbose, thus unfairly tilting the playing field.[82] The difficulty of doing this fairly limits the usefulness of this kind of analysis, but nevertheless it seems to have a few simple implications. First, an explanation containing more conjunctions will generally be longer than one with fewer, and therefore we might deem it less simple on that account. Thus, for instance, we might judge that an explanation by a plaintiff that requires us to believe facts A, B, C and D is more complex than a defendant's denial that A is true, simple because the plaintiff's story requires more time to state. Moreover, this approach would seem to imply that a plaintiff's conjunctive assertion that A, B, C and D were true would be seen as equally as complicated as a defendants disjunctive denial of each in turn, because it requires an equivalent number of symbols to express either 'A, B, C & D' or 'A, B, C or D'. So in short, focusing on the syntactic length of an explanation seems to imply that we would prefer explanations with fewer elements over explanations with more, but be indifferent between explanations incorporating conjunctions and explanations incorporating disjunctions.

Another common approach to assessing an explanation's simplicity focuses not on its syntax but instead on its ontology, asking not how many words it takes to express, but rather how many assumptions it makes regarding states of affairs in the world. This formulation is often known as 'Ockham's Razor' in reference to the 14th-century philosopher William of Ockham, who cautioned theorists that 'plurality is not to be posited without necessity'[83] because each additional hypothesis may increase the likelihood that the overall theory incorporates errors.[84] More recent defenders of the principle have cast it in terms of 'ontological economy',[85] with the common thread being that it seems defensible to prefer one theory over another, all other factors being equal, on the ground that the disfavored theory 'assumes the existence of entities that' the favored theory does not.[86] For example, one of Einstein's

---

[80] *Cf.* Malcom Forster & Eliott Sober, *How to Tell When Simpler, More Unified, or Less Ad Hoc Theories Will Provide More Accurate Predictions*, 45 BRITISH J. PHIL. SCI. 1, 8 (1994).

[81] *See, e.g.*, Elliott Sober, *What is the Problem of Simplicity?*, in SIMPLICITY, INFERENCE, AND MODELING (A. Zellner, H. Keuzenkamp, and M. McAleer eds. 2002); Alan Baker, *Occam's Razor in Science: A Case Study from Biogeography*, 22 BIO. & PHIL. 193, 194–95 (2007).

[82] *See* Thagard, *supra* note 78, at 87; Sober, *supra* note 81, at 16.

[83] SOBER, *supra* note 77, at 41 n.2.

[84] Paul Vincent Spade, *Ockham's Nominalist Metaphysics: Some Main Themes*, in THE OXFORD COMPANION TO OCKHAM 100, 101 (1999).

[85] *See, e.g.*, Thagard, *supra* note 78, at 88–89; *see also* Baker, *supra* note 79, at 195.

[86] Thagard, *supra* note 78, at 88–89. The popularity of this new formulation is perhaps best illustrated by the fact that Ockham is now widely misquoted to have warned that '*entities* must not be multiplied without necessity'. SOBER, *supra* note 77, at 41 n.2 (emphasis added).

favored arguments in favor of his new theory of special relativity was that it eliminated the necessity of positing an unobservable 'luminiferous ether' to allow light to travel through a vacuum.[87] Likewise, modern biologists reject the notion of intelligent design in favour of evolutionary theory in part because the former posits a complex and unobservable designing entity that the latter can dispense with.[88] To be sure, there can be some difficulty in deciding how to divide the world into 'entities' when doing such an analysis, but it is undoubtedly an important aspect of how scientists choose among rival theories.

Since court cases rarely involve disputes over the existence of subatomic particles, we might adapt this line of reasoning to say that, when fact-finders must choose among rival explanations of evidence, they should exercise some preference towards accounts that assume the existence of fewer separable *occurrences*, all other things being equal. Thus, we might be hesitant to accept a defendant's claim that a large number of police officers conspired to frame him for a crime on the grounds that such a conspiracy would need to involve a large number of co-occurring events to succeed.[89] Such a defense, we might think, should require a significant quantity of evidential support before rising to the level of reasonable doubt, to counter-balance its inherent complexity.[90]

We might likewise consider the addition of additional descriptive *features* of a single event to be another kind of multiplication of entities, at least to the extent that the additional features do not tend to go together as a natural pattern.[91] For instance, compare the explanations necessary to sustain a plaintiff's verdict in an identical automobile injury case in two hypothetical jurisdictions, with one jurisdiction imposing strict liability and the other requiring proof that the defendant drove negligently. The strict liability explanation could consist simply of the actions taken by the defendant and their causal connections to the plaintiff's harm. The negligence explanation, by contrast, would need to incorporate additional descriptive features of the defendant's actions, such as whether he had notice that the plaintiff was in jeopardy and whether he was aware of a risk that he might harm someone at the time when he acted. Fault might be proven inferentially based on proof of other occurrences, or by having witnesses describe the *manner* in which the defendant was driving or other features of the accident itself. We might consider these additional features to be additional 'entities' on the ground that they add ontological constraints to the situation being described; there are many more ways that a car collision could occur consistent with strict liability than ways it could occur consistent with the defendant being at fault. As a result, the explanatory virtue of simplicity would suggest that, beyond the necessity of proving some additional facts to support the additional features of the more complicated fault story, we should also demand a higher overall quantity of support for the story, so that an identical defense case might prevail in the fault jurisdiction but fail to be more plausible than the simpler explanation offered in the strict liability jurisdiction. Both these examples work to show that adding additional conjunctions to an explanatory story undermines its claim to ontological simplicity.

---

[87] *Id.* at 42.

[88] *See* Thagard, *supra* note 78, at 88.

[89] The O.J. Simpson case provides one notable example of a jury falling sway to just such an argument. 'The defense concentrated its effort on convincing the jury that Furhman and countless other cops conspired together to frame O.J. Simpson. The jury's rapid verdict indicated that it accepted the defense's conspiracy theory.' Stephen D. Easton, *Lessons Learned the Hard Way from O.J. and "The Dream Team,"* 32 Tulsa L.J. 707, 724 (1997) (reviewing Christopher A. Darden & Jess Walter, In Contempt (1996)).

[90] On the other hand, Baker has observed that simpler theories, by their nature, tend to have more trouble explaining the whole universe of observed facts, and thus a preference for simpler theories may involve an inherent trade-off with the inferential virtue of evidential support. Baker, *supra* note 79, at 212, 213.

[91] *See* Sober, *supra* note 77, at 20–22.

How, then, might we apply the ontological conception of simplicity to *disjunctive* explanations? Here there is somewhat less guidance from the philosophy of science literature, as scientists generally seek to prove that particular theories are correct, rather than to offer multiple competing theories on an equal footing.[92] But the basic guiding principle, which is that we should demand more support for theories that are more ontologically constraining, would seem to suggest that disjunctive theories are, all things considered, simpler than theories that have a similar number of conjunctive elements. Consider a defendant who disjunctively denies a three-element claim, saying that either all or some of the plaintiff's allegations are false. For convenience, we will call the elements A, B and C. While the conjunction of all three elements is consistent with only one state of the world, their disjunctive denial is consistent with seven[93] possible states of the world:

Conjunction	Disjunctive Negation
[A, B, C]	[not-A, not-B, not-C]
	[not-A, not-B, C]
	[not-A, B, not-C]
	[A, not-B, not-C]
	[not-A, B, C]
	[A, not-B, C]
	[A, B, not-C]

Initially, we might say that this profusion of *possible* ontological states seems very complicated. But if we are conceiving of ontological complexity as a question of the number of constraints posed by an explanation on historical reality, than this disjunctive explanation is actually much simpler than the conjunctive explanation because it is consistent with a larger number of states that the world could be in. So on the ontological conception at least, we might say that disjunctive explanations are simpler than conjunctive explanations, because disjunctive stories permit the world to take a wider variety of shapes than do conjunctive ones.

Finally, let us consider how the two conceptions of simplicity might combine when we decide how much credence to accord to explanations with different logical structures. First, we might find it hard to decide whether a unitary or a disjunctive explanation is simpler, given that the former can claim more syntactic simplicity while the latter might lay claim to more semantic simplicity. Second, and more importantly, we would find either a unitary or a disjunctive explanation more persuasive than a conjunctive one, which would be penalized on both ontological and syntactic accounts. And this only reinforces the judgments we might reach based on considering the explanation's coherence and consistency with our background beliefs. In short, the more conjunctive an explanation is, the more evidential support we should demand before finding it convincing, especially when the conjunctions involved are unusual or surprising. Conversely, explanations incorporating disjunctions may rub our intuitions the wrong way, but they should typically receive an inferential boost rather than a penalty by virtue of their flexibility. Explanationism, in short, is best interpreted as offering

---

[92] *Cf.* THOMAS S. KUHN, THE STRUCTURE OF SCIENTIFIC REVOLUTIONS 12–22, 102–03 (3d ed. 1996).

[93] More generally, for a set of n fully disjunctive elements, the number of states of affairs consistent with the disjunctive explanation is 2^n-1.

recommendations regarding the consideration of conjunctive and disjunctive likelihoods that closely parallel what probability theory has to say on the subject.

3.5 *Descriptive explanationism and conjunctive reasoning*

Before we conclude that the paradox persists within an explanationist framework, we must briefly consider one further possibility. If, when employed descriptively, Pardo and Allen's framework implies that juries do make the necessary adjustments when considering stories that are conjunctive in form, then we should conclude that there is no paradox after all, because it suggests that juries actually do what they ought to do. We must tread carefully here, because the authors have not themselves provided many suggestions regarding the ways that jurors actually reason when considering conjunctive or disjunctive explanatory accounts. Nonetheless, I will urge here that there are good reasons to worry that, when deciding which party has offered the more convincing story, jurors might fail to penalize—or even reward—parties who offer conjunctive explanations of the evidence.

There are two different ways that jurors might correctly account for conjunctive likelihoods. They might do this intuitively and automatically, so that plaintiffs and prosecutors automatically find it a bit harder than defendants to have their stories be found credible, especially when they involve unusual or surprising combinations of events. Conversely, they might have intuitions regarding explanatory strength that fail to make such adjustments, but then take conscious steps to downgrade their acceptance in stories that involve relatively independent conjunctive elements. As I will explain below, I believe the former occurs to some extent, but most likely not as much as it should, while the latter rarely occurs in ordinary American trials.

Consider first the possibility of an automatic, intuitive adjustment. There are several documented biases in intuitive reasoning that cast doubt on the existence of an innate 'knack' for making appropriate conjunctive adjustments, even for everyday sorts of problem solving. In general, intuition can become an accurate mode of solving difficult problems when decision-makers have experience making similar decisions along with prompt, accurate feedback regarding their performance. With such experience and feedback, chess masters become adept at recognizing winning and losing positions with a moment's glance at a board,[94] and some fire captains develop a 'sixth sense' for how a fire will develop over time.[95]

Unfortunately, just because we do a task routinely does not guarantee that we receive prompt and accurate feedback regarding our performance. Numerous studies, for instance, show that most people barely outperform chance when trying to detect whether another person is lying to them based on surface cues like facial expression and tone of voice,[96] and this persists despite the fact that trying to detect lies is hardly an unusual pass-time. Even more strikingly, those whose jobs it is to detect lies, such as police officers, barely do better, and their practice does not improve with more experience on the job.[97] Although this may initially seem surprising, on reflection the causes are clear: Whether a suspect is lying or telling the truth, an officer may never discover the answer in a clear way. An officer may assume that an acquittal was due to an error by the jury or insufficient evidence rather than factual innocence, reducing the clarity of negative feedback. And on the flip side, even an innocent suspect

---

[94] *See* KAHNEMAN, *supra* note 38, at 11.

[95] *See* GARY KLEIN, SOURCES OF POWER 15–17 (1998).

[96] *See* Mark Spottswood, *Live Hearings and Paper Trials*, 38 FLA. ST. U. L. REV. 827, 837–40 (2011) (reviewing the literature on deception detection).

[97] *See* SIMON, *supra* note 70, at 126.

might plead guilty to avoid worse consequences following a trial,[98] or be wrongfully convicted by a jury.[99] So feedback, when it comes, is ambiguous. What is worse, the time delay between an interrogation and the final verdict allows for selective memory effects to arise, so that the discomfort the officer might feel for having been wrong can be moderated by a tendency to remember herself as having been less sure about the supposed lie than she actually was at the time.[100]

The same sort of problem likely infects our ability to develop skill at adjusting our assessments of explanatory power based on the quantity of conjunctions that might be included in the explanation. Whether in the courtroom or out of it, when we rank possible explanations for an observed phenomenon involving the actions of other people, we rarely receive unambiguous confirmation of whether we were right or wrong, and when we do receive such feedback it is often too late to be of much help. Consider, as one example, one of the more common kinds of conjunctive inferences we must make, which is deciding how long it will take to complete a complex project. Such planning often hinges on the conjunction of multiple predictions each of which are relatively independent from one another: How much time and effort we can make available? How much effort will be required to complete different stages of the project if all goes well? How many unforeseen setbacks are likely to arise? And what quantity and quality of support will we receive from others? Moreover, this might seem like a case where we ought to learn from experience, as we will know how long it took to complete similar projects in the past, and all it should take to evaluate the success of the original prediction is comparing the expected date of completion with the actual date. Nonetheless, ample research suggests that in such situations both individuals and organizations systematically engage in a *planning fallacy* by systematically and dramatically underestimating the amount of time that projects will actually take.[101] (This is particularly telling because one plausible contributing factor that might give rise to the planning fallacy would be a failure to consider that, when delays are possible from multiple independent sources, the likelihood that the overall project will be delayed should be considered to be *higher* than the probability that the single most likely source of delay will be realized.)

The problem may be made worse by the structure of intuitive cognition, which many researchers have identified as using associative and coherence-driven logic to promote focal hypotheses, leading to a risk that disjunctive alternatives may be neglected or suppressed. As I have described elsewhere, our unconscious minds frequently operate by means of pattern recognition, sifting through our sensory inputs and fitting them to the nearest appropriate schema.[102] Once an acceptable schema has been recognized, intuitive reasoning can then shift into a confirmatory mode, promoting consistent observations to our attention while making alternative possibilities seem less likely.[103] In such a reasoning mode, the inclusion of additional conjunctive elements that complete a classic 'story pattern' may make an explanation seem more convincing, while the inclusion of mutually inconsistent disjunctions

---

[98] *Cf.* SAMUEL R. GROSS ET AL., EXONERATIONS IN THE UNITED STATES: 1989 THROUGH 2003, at 12 (2004).

[99] *Cf.* John Roman et al., *Post-Conviction DNA Testing and Wrongful Conviction* at 6 (Urban Institute working paper, 2012); Samuel R. Gross et al., *Rate of False Conviction of Criminal Defendants Who Are Sentenced to Death*, 111 PROC. NAT'L ACAD. SCI. 7230 (2014).

[100] *See also* SIMON, *supra* note 70, at 126 (noting that despite their professional interest in catching liars, and their long experience attempting to do so, police interrogators 'tend to report relying on the same [misleading] deceit cues as lay people'.).

[101] *See* KAHNEMAN, *supra* note 38, at 249–52.

[102] Spottswood, *Hidden Structure*, *supra* note 40, at 22–39.

[103] See STANOVICH, *supra* note 68, at 113–15; Dan Simon, *Cognitive Coherence in Legal Decision Making: A Third View of the Black Box*, 71 U. CHI. L. REV. 511, 512–13 (2004); Raymond S. Nickerson, *Confirmation Bias: A Ubiquitous Phenomenon in Many Guises*, 2 REV. GEN. PSYCHOL. 175, 175–220 (1998).

may make it seem dissonant and incoherent.[104] Unfortunately, this may lead our intuitive minds to neglect the inherent advantages that disjunctive explanations possess, in terms of their consilience with a larger number of possible states of the world.

So in short, there are strong reasons to doubt that judges and jurors automatically and intuitively penalize explanations that incorporate conjunctions by comparison with simpler or disjunctive explanations. What, then, of the alternative possibility, that jurors' notice that their intuitions on these subjects are faulty and apply conscious adjustments to correct them? On this point I think it is even clearer that the descriptive reality fails to line up with the analysis I offered above. On the one hand, it seems hard to doubt that, in extreme cases, a juror's common sense would balk when a party offers a theory of the case that involves highly independent conjunctions. As in the felony murder hypothetical discussed above,[105] when strange coincidences pile up too far the jury may start to doubt that the claimed events are possible or likely. On the other hand, it seems less intuitively obvious that even for cases involving common fact patterns, the multiple elements that a plaintiff must show will rarely exhibit *perfect* dependence, so that the combined story should always be *less* persuasive than its *least* persuasive part. Perhaps the best evidence that this is *not* the case is the fact that judges have continued to give instructions that create a conjunction paradox for many decades, without anyone other than law professors noticing that something might be amiss! Moreover, the psychological studies reviewed above tend to suggest that although ordinary people can do a reasonable job handling problems that demand conjunctive reasoning when the task design specifically facilitates doing so, they will often go astray when the need to do so is not made as clear as possible.

To summarize, there is little reason to think that jurors will, as a general rule, give less credence to stories on the basis of the stories' conjunctive form, either by use of intuition or deliberative reasoning. By contrast, the best interpretation of explanationism as a *normative* theory seems to demand that jurors make such adjustments, as conjunctive stories will generally be less simple than unitary or disjunctive ones. Moreover, to the extent that they incorporate story elements that are relatively surprising or coincidental in combination with one another, they will also suffer defects in their internal consistency and their consistency with a fact-finder's background beliefs. This gap between the theory's normative recommendations and its descriptive account of how jurors' actually reason in courtrooms seems no better or worse than that posed by classical probability theory. In short, whether one is a Bayesian or an explanationist, the conjunction paradox is alive and well.

4. Why novel mathematical accounts fail to offer superior normative guidance to fact-finders

By contrast with explanationism, a different kind of theoretic response to the conjunction paradox has been to reform our normative theories of fact-finding so that they correspond with descriptive reality. Multiple authors have attempted to craft solutions along these lines, which retain the notion of determining the likelihood of particular elements of a case while rejecting the product rule.[106] Each of these models may have real value as *descriptive* tools, in that they attempt to provide a structured account of the ways that juries are currently instructed to decide cases.[107] Nonetheless, the advocates

---

[104] *Cf.* Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 CARDOZO L. REV. 519, 542 (1991).

[105] *See* Discussion, *supra* at sub-section 3.4.2.

[106] *See generally* Clermont, *supra* note 2; Cheng, *supra* note 2; Nesson, *supra* note 2.

[107] *See* Discussion, *supra* at sub-section 2.3.

of such models typically claim that these accounts also provide a normative *justification* for current practices. In this section I will briefly review their theories and then illustrate why such normative use of these models is undesirable.

4.1   *Alternative mathematical approaches to handling conjunctive proof*

Charles Nesson was one of the first legal scholars to suggest that the rule embodied in standard jury instructions was normatively preferable to the guidance offered by standard probability theory. In an oft-cited article, he argued at length that the primary value that legal rules should advance in the context of jury trials was ensuring that the public accepts verdicts as authoritative, whether or not such acceptability closely tracks the historical events that are most probable.[108] He suggests that the conjunction paradox illustrates the desirability of such an emphasis. First, he notes that the formal paradox only arises with respect to elements, and suggests that jurors will generally make appropriate adjustments when the proof of a single element is conjunctive in form.[109] He then notes that the standard instruction encourages juries to choose 'the most believable account of each element' in a case, and suggests that findings consistent with the most believable account of each element will necessarily represent 'the most acceptable' of the possible accounts to the public.[110] Although he seems to accept that the conjunction of two believable elements will often be less probable than not, he also believes that the public will routinely accept this conjunction as historically true, because they will not be interested in 'incoherent' disjunctive possibilities, only a comparison of plausible historical narratives.[111] Whether or not a verdict based on individually probable elements is 'more probable than not', it should nonetheless be more probable than 'any other story about the same elements', in other words.[112] Thus, in essence, Nesson proposes the following alternative mathematical rule for assessing what we might call the 'legal acceptability'[113] of a claim:

If $E = \{a, \ldots, n\}$ are the legal elements of a claim,

Then LA (legal acceptability) $= min\{p(a), \ldots p(n)\}$.

Since Nesson suggests that liability should be assessed whenever this legal acceptability exceeds 50%, it delivers equivalent results to what standard probability theory would recommend if the probabilistic dependence among the elements were maximized, so that proving one element necessarily implied the truth of the remaining ones:

Assume that $E = \{a, \ldots, n\}$ are the elements of a claim,

And that, for some element e in E, $P(E|e) = 1$.

---

[108] Nesson, *supra* note 2, at 1358–59.

[109] *Id.* at 1388. Nesson does not cite any empirical evidence to support this assertion, and in fact the psychological literature casts doubt on the assumption that most people routinely and automatically make appropriate adjustments for the likelihood of conjunctions. *See* Discussion, *supra* at sub-section 2.2.

[110] *Id.*

[111] *Id.* at 1389.

[112] *Id.*

[113] The term is mine, not Nesson's.

$$Then\ P(a\ \&\ldots\&\ n) = P(e) = min\{P(a),\ \ldots,\ P(n)\}^{114}.$$

So in short, Nesson's suggestion is equivalent to treating the probability of a conjunction as equal to the likelihood of its least likely component, which we can shorthand as the 'Least Likely Element Rule' or LLER.

More recently, two other influential scholars have proposed mathematical models of the proof process that tend to encourage similar results. Edward Cheng proposed that the inquiry of a jury deciding a civil case should be understood as a *probability ratio test*, in which rather than assessing the likelihood of the plaintiff's theory of the case in isolation, the jury instead compares it to the likelihood of the defendant's story. As Cheng acknowledges, this formulation is equivalent with the standard framework of probability theory if we assume that the defendant is permitted to offer a fully disjunctive defense and argue simply that the plaintiff is wrong. But he argues that his alternative framework nonetheless resolves the conjunction paradox because 'the defendant, especially in a civil case, may not simply be a contrarian' but must instead 'present an explanation of what happened'. And even if a defendant choses to ignore this advice (because no rule forbids disjunctive defenses), Cheng argues that a jury must evaluate each disjunctive story separately, so as to render a verdict based on whichever conjunctive story is most likely.[115] Once adjusted to ignore disjunctive story possibilities, Cheng's theory functions similarly to the LLER in most (but not all) cases.[116]

Finally Kevin Clermont has proposed an account of the proof process that makes use of the theory of *fuzzy sets*, originally developed by the mathematician Lotfi Zadeh.[117] Zadeh's theory amends classical set theory by permitting set membership to be a matter of degree rather than a binary outcome, so that an element can be partially included and partially excluded from a set. Clermont's theory is addressed to two distinct forms of uncertainty: questions of whether an event *occurred* is what classical probability theory addresses itself to quantifying, while a fuzzy set theory of evidence can also handle the issue of *event imprecision*.[118] Clermont takes things a step further by explaining why the membership

---

[114] The proof of the above proceeds as follows:

$$P(E|e) = P(E\ \&\ e)/P(e)$$

$$Because\ e\ is\ in\ E,\ P(E|e) = P(E)/P(e)$$

$$By\ assumption,\ P(E|e) = 1 = P(E)/P(e)$$

$$Therefore\ P(e) = P(E)\pounds min\{P(a),\ \ldots,\ P(n)\}.$$

[115] Cheng, *supra* note 2, at 1262–64.

[116] In the case of perfect independence, if we assume that $\{P(a), \ldots, P(n)\}$ are each greater than 0.5, then it follows readily that $P(H_{plaintiff})$ will be greater than the probability of any story involving the denial of an element. Under that assumption, $P(H_{plaintiff})$ is equal to the product of each element, and any $P(H_{defendant})$ will involve the denial of at least one element. However, if we substitute a probability *lower than* 0.5 into the overall product, instead of one that is higher than 0.5, this will necessarily lower it. Likewise, cases involving positive dependence will exhibit a similar pattern of results, because any amount of positive dependence will tend to inflate $P(H_{plaintiff})$ by comparison with the case of independence, while lowering $P(H_{defendant})$. It is only in cases involving probabilities close to 0.5 and large amounts of negative dependence among the plaintiff's elements that the two approaches will diverge, such that the probability ratio test endorses similar results to standard probability theory. For instance, imagine that $P(A) = 0.6$, $P(B) = 0.6$, but $P(A\ \&\ B) = 0.2$, a situation of maximal negative dependence. In that case, $P(A\ \&\ B)$ would be much lower than $P(A\ \&\ not\text{-}B)$, which would be 0.4. Thus, although the LLER would recommend a plaintiff's verdict in such a case, Cheng's theory would not.

[117] L.A. Zadeh, *Fuzzy Sets*, 8 INFO. & CONTROL 338 (1965).

[118] Clermont, *supra* note 2, at 1080; *see also* Bart Kosko, *Fuzziness vs. Probability*, 17 INT. J. GENERAL SYSTEMS 211, 213 (1990).

of an element in the conjoint of two fuzzy sets should generally be considered to be the minimum of that elements membership level in either of the two fuzzy sets in isolation, represented by the 'MIN' operator. To borrow his illustrative example, if a person named Tom is '0.3 out of 1' tall and '0.4 out of 1' smart, that is a mathematical way of saying he is 'neither very smart nor very tall, and a bit less tall than he is smart'. We would therefore say that he 'is not a very tall or smart man', which is something like being a 0.3 member in the set of tall, smart men, rather than 'he is extremely dumb and short', which the product of 0.12 would seem to indicate.[119] Although the fuzzy set literature fails to recommend a single solution for combining occurrence uncertainty and imprecision uncertainty,[120] Clermont proposes that in general, a fact-finder ought to treat all their uncertainty, of either kind, as if it were event imprecision uncertainty, and thus use the MIN operator (which readers should recognize as equivalent to LLER) to decide whether to impose liability on the basis of multiple uncertain elements. He defends this choice on two primary grounds: First, he suggests that applying the product rule may undermine decisional accuracy, so that 'plaintiffs would lose strong cases they really should win' and 'defendants at fault would not receive a corrective message'.[121] Second, he argues that litigation fact-finding involves the revelation of 'partial truths' whose truth or falsity can never be perfectly known, rather than 'betting question[s]' with certain outcomes we will later learn.[122] Thus, rather than picturing '60 cases of A and 40 cases of not-A' when we think an element has been proven to a likelihood of 0.6, we should instead approach the case as being '0.6 true' with respect to A, and each additional element, and conjoin those 'partial truths' according to the LLER.[123]

4.2. *The deficiencies of the least likely element rule*

We thus have three different authors, employing three very different analytic frameworks, all converging on one recommendation, which is that the law's existing jury instructions on probability are normatively preferable to what classical probability theory would counsel. Unfortunately, the normative arguments they have provided for this alternative standard are unpersuasive. First, appropriate use of probability theory (rather than a misguided choice to use the pure product rule in all circumstances) would likely reduce the rate of error at trial, not increase it, because the mathematical formalisms it employs are designed to correspond to the nature of real world variations. Moreover, even if we worry that reforming the rule would cause problems by shifting the balance of persuasion towards the defense, we should still seek other ways to advantage plaintiffs to counteract such an effect rather than maintain the existing system, which tends to advantage *less* deserving plaintiffs while giving no help to those who were most likely wronged. Second, the fact that in most cases we cannot be nearly certain of the historical truth that is being litigated does not mean that it is incorrect to use probability theory to handle our uncertainty. Third and finally, we should be very reluctant to retain a rule that we

---

[119] Clermont, *supra* note 2, at 1097–98.

[120] In cases where there is event outcome uncertainty but no event imprecision uncertainty, even strong proponents of the utility of fuzzy set theory admit that the ordinary rules of probability should apply. *See, e.g.*, Kosko, *supra* note 118, at 234. Moreover, some fuzzy set theorists have articulated additional formalisms that permit one to analyze in event imprecision and outcome uncertainty for a single event, thus permitting one to separately formalize both degree-of-membership and outcome uncertainty, rather than blend them together. *See, e.g.*, Didier Dubois & Henri Prade, *Possibility Theory, Probability Theory, and Multiple-Valued Logics: A Clarification*, 32 Annals Math. & Artificial Intelligence 35, 62 (2001); *see also* Giangiacomo Gerla, Fuzzy Logic: Mathematical Tools for Approximate Reasoning 178 (2001).

[121] Clermont, *supra* note 2, at 1100.

[122] *See* Kevin M. Clermont, *Conjunction of Evidence and Multivalent Logic* 3-4, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2472383 (working paper, posted on 5 August 2015).

[123] *Id.*

know is increasing the risk of error at trial merely because it is good public relations for the litigation system—and in any event, the reasons to suppose that correctly applying the rules of standard probability would undermine public confidence in the judicial system are far weaker than Nesson suggests.

4.2.1 *Decisional fairness and accuracy.*   Let's start with the potential worry that enforcing the ordinary rules of probability theory would undermine decisional accuracy. Clermont voices this concern, and provides a hypothetical case in which the plaintiff proves four elements to a probability of 0.7 each, and then states that the plaintiff 'would lose with a miserable 24% showing under the product rule'.[124] He goes on to say that this represents 'a strong case becom[ing] a sure loser',[125] which seems to beg the key question. Stating that you have proved an element to a probability of 0.7 implies that, if you litigated 10 similar cases and relied on the 0.7 to find against the defendant, you would be wrong 3 times out of 10—hardly a slam dunk! Moreover, if we go along with Clermont and assume that the joint probability is truly 0.24 (implying perfect independence among the elements, which will be rare), then that implies that the case involves the surprising combination of four elements that have no particular reason to go together; logically, this leads us to say that if we find the defendant liable based on the specified proof, we will be *wrongfully finding the defendant liable in three cases out of every four*, because we will have guessed *wrong* about at least one element in 76% of factually similar cases.[126]

To make this more concrete, let us try to imagine a case fitting Clermont's constraints. Peter sues his former co-worker Donald for battery, alleging that he deliberately hit him with his car and then drove off before officers could arrive at the scene. To prevail, he must prove that Donald (1) intended to hit him with his car, (2) succeeded in doing so, (3) that the contact was harmful or offensive, and (4) that Peter did not consent to the contact.[127] To keep the probabilities of each element separate and independent, we must make sure to construct the hypothetical so that two elements do not stand or fall together based on the testimony of a single witness, so let us imagine that Donald suffers from amnesia arising out of the crash and that Peter has a recent mental illness that will render him incompetent to testify. Accordingly, the only testimony on each element will be as follows:

- Intent: Walter, a witness with a moderate bias towards Peter, who is his friend at work, will testify under oath that he heard Donald threaten to kill Peter because Peter had slept with Donald's wife.
- Contact: Wanda, a bystander, saw the accident from across a busy street, under good, but not excellent, viewing conditions. She was tentative when she first picked Donald's face out of a police line-up, but now is quite confident that Donald was the hit-and-run driver.

We would normally assume that a person hit by a car was harmed and did not consent, but two more witnesses testify, lowering our ability to assume these facts are true based on the fact of a car collision. Neither is highly trustworthy because each is a long-time friend of Donald's, but they offer sworn testimony to rebut the third and fourth elements, respectively:

- Harm: William testifies that he heard Peter say that he fabricated his injuries to inflate his damages and that he actually suffered no harm in the crash.

---

[124] Clermont, *supra* note 2, at 1100.

[125] *Id.*

[126] *See also* Allen & Stein, *supra* note 2, at 596.

[127] *See* RESTATEMENT (THIRD) OF TORTS: INTEN. TORTS TO PERSONS § 101 (Discussion Draft, 2014).

- Consent: Winona says that she heard Peter beg Donald to run him over with his car so that he could take some time off at home with his family and collect disability checks.

I have tried to construct this example so that we think that each element is only 70% likely. In the first two cases the prima facie proof is strong but not overwhelming, while in the second two cases what would otherwise be an easy inference from the first two elements is complicated by weak counterproof. Moreover, the reasons we might have for being unsure about any individual element do not carry over to the others; being wrong about one has little to do with being wrong about another. This case, I feel sure, sounds absurd to most readers, feeling full of strange, *ad hoc* coincidences, and I doubt that many would actually be inclined to label the plaintiff's case as 'strong'. This is exactly what I wished to use it to illustrate: a case involving four *totally* or *nearly independent* elements, each proven to a 0.7 level of probability, would not feel, in real life, like a strong case, but rather a mess of inconsistencies and sources of doubt. Most typical tort cases would involve fewer seriously contested elements and more dependence among those elements, so that although a downward adjustment from the LLER would still be required by standard probability theory, it would be a much smaller one than Clermont sets forth, and would likely only play a deciding role in cases involving probabilities close to 0.5, unusual levels of complexity, or strange and coincidental stories.

Now let us consider a more plausible case in which the choice between the LLER and the ordinary probability rule might make a difference. Plaintiff Auto Lender (PAL) is suing Doug for a debt owed, claiming that he borrowed $25,000 to finance the purchase of a car and then failed to pay some of it back on schedule. PAL's lending records were lost in a data storage center fire, so the case must turn on the testimony of a lending officer, a call-center employee who processed Doug's loan payments over the phone, and Doug, all of whom are biased. There are two contested issues in the case:

- Issue 1: Whether Doug received a mandated set of lending disclosures before signing, without which the contract is void under state law.
- Issue 2: Whether Doug has actually paid back the money already, as he claims.

Let us assume, for the moment, that the two PAL employees told a slightly more credible tale, but that Doug was fairly convincing as well, so that we feel 60% certain that the company should prevail on each issue. Those who advocate the LLER would say that this is a clear victory for the plaintiff, while probability theory would suggest that we should slow down and consider things more carefully. In particular, we should notice that although the two probabilities are partially dependent—because each depends, in part, on our estimates of Doug's credibility—they are also partially independent, given that they also involve two separate acts by different employees, either of whom might lie independently of the other. The joint probability of the two elements, therefore, is somewhere between 0.6 and 0.36. To get a roughly correct answer we might split the difference, and note that the halfway point in our range of possible probabilities is 0.48, which suggests a victory for the defense. Such a decision, moreover, seems quite justifiable when we reflect on the fact that for each element, if we go with our 0.6 probability in the plaintiff's favor we will be wrong two times for every three times we are right. Moreover, based on fact that our plaintiff must prove their case conjunctively while our defendant can win disjunctively, an error on *either* element would mean that the defendant was really in the right, and did not legally owe the plaintiff what it demands. Epistemic humility counsels us to remember that even though we can make a narrow call as to who is right on each element, we may be wrong, and such risks are amplified as cases accumulate conjunctive arguments.

Some readers may feel that the selection of this example is unfair—I picked a corporate plaintiff and a 'little guy' defendant—but in fact I chose the parties' identities to make an important point, which is to remind readers that most lawsuits are not brought by 'small-time' plaintiffs against 'big-time' corporate defendants. In fact, a recent survey of state court litigation found that contract claims wildly outnumbered tort claims by a factor of about 9:1, and that the largest subcategories of contract claims were debt collection cases (37% of all contract claims), landlord-tenant cases (29%), and foreclosure cases (17%), none of which fit the 'little sues big' paradigm.[128] Moreover, the conjunction paradox may also rear its head in criminal cases, where defendants may be especially vulnerable, although its operation is harder to pin down because of the persistent refusal of courts and rule-makers to define the concept of 'reasonable doubt' in quantitative terms.[129]

Finally, I must note one more concern with the LLER's effects on error distribution, which is that to the effect that it benefits plaintiffs, it benefits the *least deserving class* of plaintiffs, those whose cases involve inherently odd combinations of events, meager proof on key elements, or unusually complex and precarious theories of liability. Even if we think that structural barriers stand in the way of plaintiffs with deserving claims and prevent them from obtaining fair settlements or verdicts, the LLER, as currently embodied in jury instructions, is doing little to help most of them prevail. Those plaintiffs who do benefit from it, when compared with those who do not, are disproportionately likely to be bringing claims that are ultimately unfounded. A wide variety of other reforms, such as fee-shifting rules[130] or easier access to effective civil discovery,[131] might help plaintiffs more even-handedly, and might even skew their benefits towards the most deserving plaintiffs, rather than those with structurally weaker claims.

To sum up, the complaints that the ordinary rules of probability theory are unfair to plaintiffs or would increase the error rate of the system seem ill-founded. The strict product rule would only operate in a rare set of cases, and when it did have its effect a fair-minded observer would be hard-pressed to call the result unfair, at least if they took into account the real likelihood of human error as uncertain decisions accumulate together. To the extent that it makes victories less likely for some plaintiffs, such cases may involve 'little guy' defendants more often than most commentators seem to assume, and in any event the plaintiffs whose claims would be disfavored are precisely those plaintiffs who probably would not be entitled to recover if we had access to an omniscient viewpoint from which all relevant facts could be reliably known.

4.2.2 *Probabilistic uncertainty versus 'partial truths'.*   Now let us turn our attention to Clermont's argument that is that it is inappropriate to use a probabilistic framework to decide cases because fact-finding requires the manipulation of fuzzy, subjective likelihoods, rather than probabilities of bivalent events whose outcome we can eventually know with certainty.[132] Clermont is surely right that there are cases in which the most likely pattern of events will have only a 'fuzzy' fit with what the law requires. Nonetheless, many cases involve primarily factual disputes, rather than fine questions of legal categorization. If a murder defendant admits that the victim was killed deliberately but maintains that

---

[128] *See* National Center for State Courts, Civil Justice Initiative, *The Landscape of Civil Litigation in State Courts* 17-18 (Working Paper, 2015), *available at* http://www.ncsc.org/Newsroom/News-Releases/2015/Civil-Justice-Initiative.aspx.

[129] Cf. Erik Lillquist, *Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability*, 36 U.C. Davis L. Rev. 85, 87, 111–12 (2002).

[130] *See, e.g.*, Issachar Rosen-Zvi, *Just Fee Shifting*, 37 Fla. St. U. L. Rev. 717 (2010).

[131] *See, e.g.*, Scott Dodson, *New Pleading, New Discovery*, 109 Mich. L. Rev. 53 (2010).

[132] Clermont, *supra* note 2, at 1105.

someone else did it, the necessary categorizations will be 'crisp', rather than fuzzy, once the fact-finding process is complete. If the jury thinks the defendant was the killer, the elements of murder will be easily satisfied; it they think it was someone else, a not-guilty verdict will follow just as readily. The questions in such cases really are 'either/or' question of historical fact: Who was it who killed the victim? At least from the perspective of an omniscient observer, such cases have clear right and wrong answers.

Most cases will likely lie between these extremes, combining both historical and categorization uncertainty. That does not necessitate, however, that we merge the questions of how we determine the relevant facts with how we categorize them. We could quite easily employ the standards of fuzzy set theories to describe how to fit an agreed set of facts into complicated and murky legal categories, while following the dictates of probability theory to choose among multiple competing accounts of what the facts of the cases actually are.[133] So there is no formal obstacle that would prevent us from applying classical probability rules to determine whether the plaintiff's conjunctive factual account was more likely than the defendant's disjunctive one, and then using the LLER to decide whether the determined facts have an appropriate level of fit with the fuzzily defined legal elements that determine liability or non-liability.

Clermont resists this possibility by arguing that thinking of litigation uncertainty in probabilistic terms is inappropriate because probabilistic reasoning 'assumes that facts are either completely true or completely false', while in real cases we will never know the underlying facts to a certainty.[134] Although he is right that in most real-world cases there will never be a single source of evidence that can resolve uncertainty regarding historical facts beyond any practical doubt, it does not follow that using the LLER in place of ordinary probabilistic reasoning would better help us achieve our goals. In other words, we can easily grant Clermont's premise while rejecting his conclusion.

To see why, consider an analogous situation involving Doctor Evil, a nefarious dictator. Imagine that Doctor Evil has imprisoned someone you love, and will treat her either well, or poorly, depending on the outcome of an ongoing game he will play with you. This game will proceed as follows: First, the Doctor will thoroughly shuffle a standard deck of cards. Second, he will ask you to guess whether the first two cards on top of the deck are both number cards. Third, he will check to see if your guess was correct. If it was, he will give your loved one food. If it was wrong, he will instead torture her. Finally, Doctor Evil will take the rest of the day off to pursue other dictatorial business, and return to play the game again with you the following day, thus iterating the game for the indefinite future.

Several things are true about this scenario. First, from the perspective of an omniscient observer, the 'probability' of the top cards both being number cards is either 0 or 1 each time you play a round, because the deck has already been shuffled and the actual cards on the top of the deck are either number cards or face cards with certainty. There is, in other words, a 'fact of the matter' each time that you are asked to play a round.

Second, from your own perspective as a less-than-omniscient observer, who lacks any information regarding the particular ordering of the deck, the situation nonetheless appears uncertain. Given that you have no way of knowing which cards are on top of the deck, it feels appropriate to approach the problem by considering all the different ways that decks of cards *in general* can be organized, and

---

[133] Zadeh himself argued that mathematics could better track natural language conventions if we used separate formalisms to denote the *truth-value* of an event (which might be fuzzy), the *probability* of an event's occurrence, and the level of *possibility* that an event might occur, depending on which kind of uncertainty we are concerned about. *See generally* Lotfi A. Zadeh, *Fuzzy Sets as a Basis for a Theory of Possibility*, 1 FUZZY SETS & SYSTEMS 3, 10 (1978).

[134] Clermont, *supra* note 122, at 4.

thinking about the *frequency* with which an arbitrarily large number of decks would have two face cards on top.

Finally, if you want to maximize your chances of giving a correct answer, you will follow the standard rules of probability theory in giving an answer. Across a large number of shuffled decks, the probability that the top card is a face card is about 0.69, and the probability that the second card is an ace is also about 0.69. Therefore, the conjunctive probability is about 0.47. This implies that, although you will not know when you make mistakes, over the next 100 trials you can expect that, if you always pick the 'both number cards' option your loved one will be tortured six more times than if choose 'at least one face card.'' I submit that under such conditions, the only sensible thing to do is to pick 'at least one face card' every time and spare your loved one some suffering. As this example illustrates, the mere fact that outcomes are not *revealed* to a decision-maker does not mean that the decision-maker should not reason probabilistically. Rather, it makes sense to account for conjunctive likelihoods in a probabilistic way, whether or not the outcomes can be confirmed through observation, whenever we are trying to maximize a particular outcome of interest.

Now that we have finished exploring the example, note the following similarities that real-world litigation has with it. First, as in the above example, from the standpoint of an omniscient observer there is no factual uncertainty attendant to litigated cases, because whatever the parties may argue and whatever the witnesses say on the stand, there was a real set of historical facts that gave rise to the suit, which our hypothetical omniscient observer could perceive directly. To be sure, this observer might have trouble *characterizing* that reality within existing legal categories, but that sort of fuzziness is an ontological problem, not an epistemological one, and so it persists even though the observing entity has no factual doubts at all.

Second, real-world fact-finders will generally lack the kind of proof that can resolve any doubts they might have about a particular verdict choice; instead, they must do the best they can in the face of uncertainty, just like the person faced with the shuffled deck. In other words, they must make an uncertain decision regarding events that, as a matter of historical fact, either happened or did not happen. They therefore must address the same *ontological* uncertainties that the omniscient observer faces, plus an additional set of *epistemological* uncertainties. For each factual judgment they reach, therefore, they will face some varying level of confidence regarding the likelihood that their decisions are correct.

Finally, like our hypothetical card player who wants to avoid causing harm to a loved one even if that harm will never be revealed to him, judges and juries usually care a great deal about minimizing the risk that their verdicts and judgments will be factually mistaken.[135] In other words, we wish to be careful in deciding legal cases, not merely as an abstract epistemic ideal, but in order to assign guilt or liability to those defendants who deserve it, and not to those who did not actually violate the law. Given these similarities, there seems to be no conceptual obstacle to treating the risk of fact-finding errors in probabilistic terms, and in fact doing so will (as already discussed) tend to maximize our ability to reduce outcome errors.

4.2.3   *Public perceptions of legal legitimacy.*   The final argument we must consider was first raised by Nesson: Will the public's confidence in our legal system be shaken if verdicts are rendered in accordance with classical probability theory? Nesson maintains that allowing a defense verdict to rest on a mere disjunctive likelihood would leave us 'no history', in the sense that there was no single,

---

[135] *See* Reid Hastie, *Emotions in Jurors' Decisions*, 66 Brook. L. Rev. 991, 1004 (2001).

coherent story that the jury could point to as a better explanation of the evidence than what the plaintiff had offered.[136] Cheng echoes this argument, arguing that, 'the legal system wants the jury to arrive at some narrative of the truth'.[137] These arguments touch on an important concern, because court systems lose some of their ability to deter misconduct when they appear illegitimate to the public at large.[138] Thus, even if the LLER reduces the accuracy of verdicts to some degree, we might still prefer it if the detriment to accuracy was small but its contribution to systemic legitimacy was large. Ultimately, however, there are three problems with this argument: First, it seems unlikely that the public would become aware of many disjunctive verdicts even in a system that encouraged them, which would limit any damage such verdicts could do. Second, agreement with outcomes plays a surprisingly modest role in producing judgments of systemic legitimacy, which would naturally limit any harm that occasional publicity regarding disjunctive verdicts could produce. Finally, such an approach would be ultimately self-defeating, at least if implemented transparently, because the public would probably recoil if judges or legislators were to openly announce that they were implementing a rule that increased the risk of error at trial solely to improve the court systems' public image.

To elaborate on the first point, when one considers the institutional design of both bench and jury trials, there is little reason to think that the public would often become aware that a particular verdict was disjunctive. For one thing, the applicable burden of persuasion rule is normally conveyed in the midst of a complicated set of jury instructions, which are rarely closely considered by anyone other than the participants in a particular case. For another, juries usually report general verdicts, stating only whether the defendant is liable and (if so) what amount of damages should be assessed.[139] Thus, if a typical jury were to collectively decide that although each element of a plaintiff's case was more likely than not to be true, but that their conjunction was less likely than their disjunctive negation, and award a defense verdict on that basis, the only fact that would become public is that the jury decided that the defendant was not liable. Of course, judges sometimes instruct juries to render element-by-element special verdicts, and some jurisdictions require that judges issue written findings of fact and conclusions of law at the close of a bench trial.[140] But few aside from the most obsessive of court watchers closely scrutinize such documents to extract every detail from them, so if a judge or jury occasionally publicly reported a decision in favor of a defendant after finding each element more likely than not, it is still unclear how broadly that fact would be circulated. So for all these reasons, it will be the rare case indeed where the public learns that a particular verdict rested directly on the fact that a disjunctive likelihood outweighed the conjunction of a set of probable elements.

Some attention to the psychology of decision-making makes the above even more likely: Juries, like everyone else, display a bias towards coherence in their reasoning, so that once they start to view some elements as probable they will have a general tendency to amplify the probability of all the other elements that are part of the same coherent story, while suppressing the probability of opposing elements.[141] What is more, mock jury experiments have revealed that most people, when given a sample case to decide, proceed by choosing the most coherent and appealing story that corresponds to the available evidence, and then selecting whichever verdict choice corresponds to that story.[142] For

---

[136] Nesson, *supra* note 2, at 1389.

[137] Cheng, *supra* note 2, at 1264

[138] *See generally* TOM R. TYLER, WHY PEOPLE OBEY THE LAW (2006).

[139] *See* Lombardero, *supra* note 57, at 275–76.

[140] *See, e.g.*, FED. R. CIV. P. 52 (a)(1) (mandating written findings following a federal bench trial).

[141] *See, e.g.*, Simon, *supra* note 103, at 515–18.

[142] Pennington & Hastie, *supra* note 104, at 521–29.

these reasons, conjunctive probabilities will probably be automatically and unconsciously amplified as jurors reason towards a verdict, and jurors will have to make an unusual and conscious effort to notice that a disjunctive explanation might be stronger than a conjunctive one. To put it plainly, if we gave better jury instructions regarding conjunctive and disjunctive likelihoods, it is far more likely that jurors would give such instructions short shrift than that they would start rendering so many disjunctive defense verdicts that the public would take notice.

To move on to the second point, even if the public notices individual verdicts that are disjunctively grounded and disagrees with the outcome in such cases, the impact on their overall judgments of court legitimacy should be modest. In Tom Tyler's seminal study of the factors that influence such judgments, whether participants agreed or disagreed with individual decisions had a comparatively small impact on their judgments regarding the system's legitimacy; in contrast, their perceptions regarding the efforts that courts made to treat them fairly and with dignity had a much larger effect.[143] In fact, so long as participants found the overall procedures to be fair, the fact that they found the outcome of their cases to be unfavorable had no statistically significant impact on either their affect towards legal authorities or their judgments of the legitimacy of the legal system.[144] What is more, Tyler's study focused on the judgments of people who had personal experiences with the court system, so that any unfavorable outcomes affected them personally and materially. Most of the time, any impact of disjunctive verdicts on the courts' legitimacy would be far more remote. In short, the existing empirical evidence suggests that any impact to the courts' legitimacy from allowing disjunctive verdicts would range on a scale from negligible to non-existent.

Lastly, the legitimacy argument for the LLER seems inherently self-defeating, because adopting it sends a message that public approval is more important than verdict accuracy, which would itself undermine the courts' legitimacy. If we assume—as I hope that we would—that whatever decision rule we recommend to the legal system can be publicly defended, then it seems that Nesson is urging us to send the following message to the American public: 'We know that you are too foolish to comprehend conjunctive and disjunctive likelihoods, and we assume that you will quickly assume that any verdicts you have trouble understanding indicate that the judicial system cannot be trusted. To solve this problem, we have decided to follow a rule that is easier to understand, even though that will result in finding some people liable even though they probably did nothing the law says is wrong.' I have little doubt that the public would find such a communication incredibly patronizing, and would reject the notion that it would ever be appropriate to enter wrongful convictions or findings of liability merely in order to boost the court's popularity or avoid having to publicly discuss complex questions. Indeed, it seems obvious that any such communication would undermine the judicial system's legitimacy to a much greater extent than the occasional entry of verdicts in favor of disjunctive defense stories.

5. Unravelling the paradox through clearer jury instructions

Jurors are currently encouraged to choose verdicts on the basis of element-by-element likelihood determinations, even though this violates both rules of classical probability theory as well as appropriate guidelines for choosing the best among rival explanations. As I outlined above, this likely undermines the accuracy of case outcomes, skewing a number of close cases towards the plaintiff, particularly in cases where the plaintiff's case is complex or incorporates unusual coincidences. In this

---

[143] *See* TYLER, *supra note* 138, at 101.
[144] *Id.* at 100–101.

section, I will explore the feasibility of giving plain language jury instructions that help jurors avoid errors when reasoning about the likelihoods of conjunctive and disjunctive events. Such instructions might offer a way to unravel the conjunction paradox by bringing descriptive practices into line with normative theory.

Although it might seem desirable to simply teach jurors the probabilistic formulation of conjunctive and disjunctive likelihoods, such an approach would be neither feasible nor desirable. First, we might fairly doubt that it is feasible to instruct lay jurors on the mathematical rules for assessing partially dependent conjunctions, or to rely on them to correctly employ them in the back-and-forth of jury deliberations. Moreover, adding very complicated instructions to a case raises the worry that jury verdicts will vary depending on the particular jury members' ability to understand and follow the instruction, which introduces troubling concerns about equity across cases. And given our reluctance to probe the details of jury deliberations in individual cases, it would be hard to assess whether such instructions were being successfully followed, misunderstood, or simply ignored.

Things get worse when we consider the reality that the degree of dependence among multiple story components or claim elements will be difficult to specify mathematically. In a case with four conjunctive elements, each proven to a level of 0.7 likelihood, the conjunctive probability might be anywhere between 0.7 and 0.24, depending on the degree to which knowing any one element to be true makes the others more likely. Our hypothetical jury would then find itself in an untenable position, because the range of conjunctive likelihoods includes both liability and non-liability! One rough-and-ready solution is to split the difference and choose the middle of the range, but this runs roughshod over real differences in case strength. And the ideal solution, which is to incorporate the actual levels of dependence among the elements or story components, requires a kind of judgment that is quite foreign to everyday experience. How much more likely is it, for example, that a particular car would crash, *given that* its brakes had been improperly maintained? On most such questions, the jury will not be able to draw on proof of objective frequency data or clear intuitions based on their own life experiences.[145]

A better solution would be to incorporate one of the central insights I offered above, which is that the problem of conjunctive likelihood can be reframed and analyzed within an explanationist framework as well as a probabilistic one. Happily, within the explanationist approach we need not worry about determining the exact degree of dependence among elements or story components, because the overall inquiry is a holistic comparison of the plaintiff's and the defendant's accounts. Moreover, as developed by Pardo and Allen, the explanatory approach to fact-finding permits a defendant to offer a disjunctive explanation of the evidence.[146] In an appendix to this article, I offer some proposed jury instructions along these lines, which use common-sense language to encourage jurors to make appropriate adjustments for conjunctive and disjunctive form in parties' explanations of the evidence, without the need for quantifying dependence relations.

The ideal form of such instructions is an empirical question best addressed through experimentation, so I offer no claim that my specific attempt is ideal. There are two opposing concerns in this area: Instructions may not provide enough detailed guidance to help juries appropriately adjust for

---

[145] At best, we might hope that jurors could obtain reliable frequency data on a few inferential issues in a case; it is hopeless to imagine that they could make principled decisions regarding the quantification of dependency relations among all possible conjunctions in the evidence. *Cf.* Ronald J. Allen & Michael S. Pardo, *The Problematic Value of Mathematical Models of Evidence*, 36 J. Legal Stud. 107, 119–23 (2007).

[146] *See* Pardo & Allen, *supra* note 2, at 251–52 & n.8.

conjunctive and disjunctive likelihoods, without becoming too complicated for the jury to follow. In part because of this concern, I suggest that a basic version of the instructions be offered in all cases, with more elaborate guidance available upon request by parties who wish to offer explicitly disjunctive cases. The appended instructions represent my best attempt to walk a fine line between precision and clarity, and I hope that those who find them either too precise or too vague may usefully engage them, either by proposing alternative formulations or testing the ability of mock juries to comprehend and apply them. Regardless of whether readers are convinced that this precise formulation is best, I hope to shift the conversation on the conjunction paradox away from arguments focused on justifying existing sub-optimal instructions, and towards identifying the best way to reduce mathematically sound rules to accessible language.

6. Conclusion

In this article, I have tried to show that the conjunction paradox is not an artifact of one probabilistic framework for analyzing evidence, but rather a basic mismatch between reliable standards of inference and the ways that American jurors are instructed to decide cases. Whether one prefers a Bayesian account of inference, an explanatory account, or something else entirely, it is appropriate to give more credence to disjunctive explanations than conjunctive ones, all else being equal. What is more, those who have tried to make a case for novel mathematical means of handling conjunctive proof have failed to show that such a rule would reduce error rates, better distribute the risk of errors across parties, or significantly improve the legitimacy of the justice system relative to their costs. Nearly all recommendations of this type suggest that the Least-Likely-Element Rule is superior to a pure product rule, while neglecting to analyse the wider range of probabilistic adjustments needed in cases involving neither perfect dependence or perfect independence. Once the comparison is between the LLER and a nuanced rule that takes varying levels of dependence into account, it becomes clear that the LLER would have the higher error rate of the two approaches. Nor would the LLER do much for the legitimacy of the justice system, as the public probably pays little attention to the details of proof standards, and they would likely recoil if we openly adopted a less accurate framework merely for purposes of public relations.

In the end, the best way to deal with the conjunction paradox is to instruct juries to account for the reduced likelihood of conjunctions and the increased likelihood of disjunctions when deliberating. The best such instructions would not require the jury to reason using an abstract mathematical framework, but would rather try to make the need for such adjustments a matter of common sense and ordinary intuition. I have offered some sample instructions of that type in this article, and although I have no doubt that better drafters could materially improve on them, at the very least it is a step towards unraveling the conjunction paradox and improving our system of justice.

Acknowledgements

Appendix

This appendix includes variations on the basic conjunction and disjunction instructions.

To be given in all civil cases:

(1) Members of the jury, you have the duty of deciding the case in the plaintiffs favor if all of the elements, taken together, have been proven by a preponderance of the evidence.

(2) This means that, if you believe that plaintiff's explanation of any element is less convincing than the defendant's explanation of that element, you should enter a verdict in the defendant's favor.

(3) If you think that the plaintiff has a better explanation of each individual element, you should then consider whether, given the proof the plaintiff has offered, the combination of all the elements being true at the same time is less likely than the defendant being right with respect to just one of them.

(4) You should consider this possibility with particular care if the case seems 'close' to you, if the plaintiff's story is complicated, or if the plaintiff's story seems to involve unusual coincidences or combinations of events. This is meant to account for the fact that if you make many close calls in the same direction, at some point you are likely to be wrong about at least one of those calls.

To be given in civil cases involving explicitly disjunctive defenses (at the defendant's election):

(5) Members of the jury, the defendant has offered you an explanation of the evidence that includes multiple possibilities, any one of which could separately entitle them to a verdict in [his/her/its] favor.

(6) As you deliberate, you should remember that the defendant is under no duty to offer you a single, consistent story.

(7) Indeed, you should generally consider the fact that the defense involves multiple alternative grounds to be a sign of its strength, rather than its weakness, so long as those defenses do not tend to disprove each other. To employ an analogy, the plaintiff needs multiple coins to come up 'heads' in order to win, while the defendant needs just one 'tails' to prevail.

(8) For these reasons, you need not accept the entire defense story to award a defense verdict, so long as you believe that it makes it likely that the defendant is probably right concerning at least one element in the case.

To be given in civil cases involving disjunctive alternative claims by the plaintiff that lead to overlapping claims for the same damages:[147]

(1) Members of the jury, the plaintiff is presenting multiple claims in the alternative for your decision. Each of these claims, if proven, would entitle the plaintiff to an award of damages based on the instructions I have given you.

(2) The plaintiff is under no duty to offer you a single story regarding what happened.

(3) You should first consider whether any individual claim is more likely than not to be true. If you conclude that any one claim, in its totality, is proven by a preponderance of the evidence, than you should find the defendant liable.

---

[147] A slightly adapted version of this instruction could also address criminal cases involving disjunctive charges by the prosecutor; the only adjustments necessary would be to substitute the identity of the prosecutor for the plaintiff and the criminal burden of persuasion.

(4) If you find that more than one claim offered by the plaintiff is possibly true, but do not think that any single version is more likely than not, you may still award a verdict in the plaintiff's favor if you think it more likely than not that one of these claims is true (rather than the defendant's denial of all the claims).

To be given in all criminal cases:

(1) Members of the jury, you have the duty of deciding the case in the prosecution's favor if all of the elements, taken together, have been proven beyond a reasonable doubt.

(2) A reasonable doubt is a doubt based upon reason and common sense—the kind of doubt that would make a reasonable person hesitate to act. Proof beyond a reasonable doubt must, therefore, be proof of such a convincing character that a reasonable person would not hesitate to rely and act upon it in the most important of his or her own affairs.[148]

(3) If you believe there is a reasonable doubt as to any element of the prosecution's claim, you should enter a verdict of not guilty.

(4) If you do not have reasonable doubts concerning any individual element, you should then consider whether there is reasonable doubt that all of the elements occurred together.

(5) You should consider this possibility with particular care if the case seems 'close' to you, if the prosecution's story is complicated, or if the prosecution's story seems to involve unusual coincidences or combinations of events. This is meant to account for the fact that if you make many close calls in the same direction, at some point you are likely to be wrong about at least one of those calls.

To be given in criminal cases involving disjunctive defenses[149]:

(1) Members of the jury, the defendant has offered you an explanation of the evidence that includes multiple possibilities, any one of which could separately entitle them to a verdict in [his/her/its] favor.

(2) As you deliberate, you should remember that you need not (believe/have reasonable doubt) regarding every one of these possibilities at the same time in order to find the defendant not guilty. To employ an analogy, the prosecution needs multiple coins to come up 'heads' in order to win, while the defendant needs just one 'tails' to prevail.

(3) Indeed, you should generally consider the fact that the defense involves multiple alternative grounds to be a sign of its strength, rather than its weakness, so long as those alternate grounds do not tend to disprove each other.

(4) For these reasons, you need not accept the entire defense story to award a defense verdict, so long as you (believe/have reasonable doubt) that one of the defenses that the defendant has offered is probably true.

---

[148] KEVIN F. O'MALLEY ET AL., 1A FED. JURY PRAC. & INSTR. § 12:10 (6th ed. 2015). This instruction would need to be adapted to accomodate different definitions of reasonable doubt.

[149] I left these instructions ambiguous with respect to the level of proof required for affirmative defenses because that level will vary across jurisdictions and particular kinds of defenses. *See generally* CHRISTOPHER B. MUELLER & LAIRD C. KIRKPATRICK, 1 FEDERAL EVIDENCE § 3:18 (4th ed. 2016).