SCHOOL OF LAW
CASE WESTERN RESERVE
UNIVERSITY

Case Western Reserve Law Review

2013

The Hidden Structure of Fact-Finding

Mark Spottswood

Follow this and additional works at: https://scholarlycommons.law.case.edu/caselrev

Part of the Law Commons

THE HIDDEN STRUCTURE OF FACT-FINDING

Mark Spottswood†

ABSTRACT

This Article offers a new account of legal fact-finding based on the dual-process framework in cognitive psychology. This line of research suggests that our brains possess two radically different ways of thinking. "System 1" cognition is unconscious, fast, and associative, while "System 2" involves effortful, conscious reasoning. Drawing on these insights, I describe the ways that unconscious processing and conscious reflection interact when jurors hear and decide cases. Most existing evidential models offer useful insights about the ways that juries use relevant information in deciding cases but fail to account for situations in which their decisions are likely to be affected by irrelevant stimuli. The dual-process approach, by contrast, is able to explain both probative and prejudicial influences on decision making. As a demonstration, I use the dual-process framework to explain the surprising result in *People v. Rivera*, a case in which a jury convicted a man of rape and murder despite the admission of exonerating DNA evidence. This result, I suggest, was not the product of an unusually lazy or unreasonable jury but rather illustrates the way that our ordinary cognitive processes can lead us to endorse quite unreasonable results if primed using certain common prosecutorial strategies.

After elaborating the dual-process model in a descriptive form, I then consider some of its normative implications. Many leading evidence scholars have argued that verdicts resting on "pure" or "naked" statistical evidence are problematic. Although the dual-process model of fact-finding is descriptive rather than normative, it nevertheless provides surprising insight into this debate by showing that our intuitive discomfort with verdicts that are based on purely statistical data may arise from the failure of such evidence to speak in terms that our unconscious, intuitive System 1 can process reliably. In such circumstances, intuitions about outcomes should be treated with caution. Thus, what unites the seemingly disparate examples of the

---

Rivera trial and the naked statistical evidence debate is that, in both contexts, it feels right to do wrong.

CONTENTS

INTRODUCTION

For quite some time, evidence scholars have debated which framework—among several competing alternatives—best describes the process by which judges and juries evaluate evidence and make verdict decisions.[1] This discussion has centered on two rival accounts of fact-finding. The "Bayesian" family of inferential models seeks to account for evidential analysis in probabilistic terms, with each new item of evidence contributing to changes in jurors' estimations of how likely it is that disputed facts are true.[2] A competing group of

---

1.   *See generally* Roger C. Park & Michael J. Saks, *Evidence Scholarship Reconsidered: Results of the Interdisciplinary Turn*, 47 B.C. L. REV. 949, 984–97 (2006) (discussing the "New Evidence Scholarship" and the extended debate between Bayesians and their critics).

2.   *See, e.g.*, JOSEPH B. KADANE & DAVID A. SCHUM, A PROBABILISTIC ANALYSIS OF THE SACCO AND VANZETTI EVIDENCE 26–27 (1996); Michael O. Finkelstein & William B. Fairley, *A Bayesian Approach to Identification Evidence*, 83 HARV. L. REV. 489, 498–501 (1970); John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L.

theorists asserts that fact-finding involves the construction and comparison of stories, or narratives, and the fact finder chooses the narrative that best explains the given pattern of trial evidence, without any separate consideration of individual probabilities.[3]

In this Article, I reflect on the shortcomings of both the Bayesian and story-comparison frameworks as descriptions of how jurors determine facts in the messy environment of real-world trials. Both approaches fail to account adequately for the unconscious aspects of fact-finding.[4] And, unfortunately, we cannot fully comprehend how real jurors are likely to process evidence unless we account for prejudice and associative intuition in addition to more rational styles of inference.[5]

Rev. 1065, 1083–91 (1968) (applying Bayes' rule to the fact-finding process in criminal proceedings); David Kaye, *The Paradox of the Gatecrasher and Other Stories*, 1979 Ariz. St. L.J. 101, 106–08 (analyzing statistical evidence under the Bayesian probability formula); Richard O. Lempert, *Modeling Relevance*, 75 Mich. L. Rev. 1021, 1022–1041 (1977) (applying Bayes' rule to the meaning of evidentiary rules regarding relevance).

3. *See, e.g.*, Floris Bex & Douglas Walton, *Burdens and Standards of Proof for Inference to the Best Explanation*, *in* Legal Knowledge and Information Systems: JURIX 2010: The Twenty-Third Annual Conference 37, 38–45 (Radboud G.F. Winkels ed., 2010) (advancing a "hybrid theory" as a model of reasoning for burdens of proof); Edward K. Cheng, Essay, *Reconceptualizing the Burden of Proof*, 122 Yale L.J. 1254, 1269–71 (2013) (proposing a probabilistic implementation of the story-comparison approach); Lisa Kern Griffin, *Narrative, Truth, and Trial*, 101 Geo. L.J. 281, 286–92 (2013) (evaluating the relationship between narrative theory and factual accuracy at trials); John R. Josephson, *On the Proof Dynamics of Inference to the Best Explanation*, 22 Cardozo L. Rev. 1621, 1626–27 (2001) (discussing the considerations undertaken during an abductive thought pattern); Michael S. Pardo & Ronald J. Allen, *Juridical Proof and the Best Explanation*, 27 L. & Phil. 223, 224–25 (2008) (contending that the Bayesian theory overlooks "explanation-based reasoning"); Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 Cardozo L. Rev. 519, 519–20 (1991) (introducing the "Story Model" as an alternative to Bayesian theory).

4. *See* discussion *infra* Part I.C.

5. A focus on the irrational or unconscious aspects of fact-finding has been comparatively rare, with little attention paid to it in the inferential modeling debates. The main counterexamples to this trend are Dan Simon's work on coherence networks, Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decisionmaking*, 71 U. Chi. L. Rev. 511, 512–13 (2004), and some aspects of Nancy Pennington & Reid Hastie's Story Model, *see, e.g.*, Reid Hastie, *Emotions in Jurors' Decisions*, 66 Brook. L. Rev. 991, 999–1004 (2001); Nancy Pennington & Reid Hastie, *Explaining the Evidence: Tests of the Story Model for Juror Decision Making*, 62 J. Personality & Soc. Psychol. 189, 193–96 (1992) (describing the impact on verdict outcomes after

To see the shortcomings of purely rationalist models, consider *People v. Rivera*,[6] a case in which an Illinois jury convicted a man for raping and murdering a child despite compelling forensic DNA evidence that a different, unidentified person committed the crime.[7] *Rivera* arose from the death of eleven-year-old Holly Staker in 1992.[8] Almost twenty years after her murder, an Illinois jury was called to decide, for the third time,[9] whether Juan Rivera committed the crime.[10] The facts of the case were gruesome: the killer had kicked in the door of the apartment where Holly worked as a babysitter, raped her both vaginally and anally, and then stabbed and strangled her until she was dead.[11] And the prosecution had an excellent card to play: when questioned by the police, Rivera had confessed to the crime.[12] In addition to this confession, the State offered evidence showing that Rivera had tried to concoct a fictitious alibi for the night of the crime and that other jail inmates had also heard him confess to the killing.[13] Based on the prosecution's evidence alone, it would seem like an easy case for conviction.

Once the defense had its say, however, the picture looked strikingly different. According to new and uncontradicted expert testimony, DNA samples taken from the semen recovered during Holly's autopsy did not match Rivera's DNA profile.[14] What is more, Rivera was on electronic home monitoring when the crime was committed, and there were no records suggesting that he had left his home that night.[15] To explain Rivera's confession, the defense offered

---

    modifying evidence presentation order and varying witness credibility in an experiment).

6.   No. 92 CF 2751 (Ill. 19th Cir. Ct. May 8, 2009), *rev'd* 962 N.E.2d 53 (Ill. App. Ct. 2011).

7.   *Rivera*, 962 N.E.2d at 62–63 (reversing and finding the evidence insufficient to support a verdict of guilty); *see also* Andrew Martin, *The Prosecution's Case Against DNA*, N.Y. Times, Nov. 27, 2011, § 6 (Magazine), at 42 (discussing the exculpatory DNA evidence in Rivera's case).

8.   962 N.E.2d at 55.

9.   The case was tried twice previously. The first verdict was overturned due to procedural errors, and the second was vacated in 2006 based on the newly available exculpatory DNA evidence. *Id.* at 55–56.

10.   *Id.* at 56.

11.   *Id.*

12.   *Id.* at 57.

13.   *Id.* at 57–58.

14.   *Id.* at 58–59.

15.   *See* Transcript of Trial at 017578 to 017616, People v. Rivera, No. 92 CF 2751 (Ill. 19th Cir. Ct. May 8, 2009) [hereinafter Trial Tr.]. There was some evidence that electronic home-monitoring systems

evidence that he had a low IQ that bordered mental retardation; that, during a break between police interviews, jail medical personnel had witnessed him behaving in ways suggesting psychosis; that his story had changed in response to leading questions by the police; and that his initial confession was inconsistent with the facts of the actual crime.[16]

In other words, despite the vivid evidence connecting Rivera to the crime, he was probably—indeed, almost certainly—not the man who raped and killed Holly Staker. The jury in Rivera's case, having heard all of this evidence, was instructed that Rivera was presumed innocent and should be convicted only if the jury was convinced of his guilt "beyond a reasonable doubt."[17] After four days of deliberations, they concluded that this standard had been met, finding him guilty.[18] The result was decried in the media and then eventually reversed on the ground that the evidence had been insufficient to permit any reasonable jury to convict.[19]

It is hard to know what to make of the jurors' inability to see reasonable doubt in this pattern of evidence. Although the case was disturbing, some might wish to construe this as an isolated event with little overall import.[20] The criminal jury, after all, is selected in a quasi-random fashion from members of the general public, and every so often that process might select a highly unreasonable group of people by a mere quirk of statistics. Seen in this light, the case could be viewed as the random result of a deck stacked against an innocent defendant, in which an overzealous prosecutor[21] and a judge who had

---

occasionally fail to record violations, but there was no direct evidence regarding any known malfunctions with Rivera's own monitoring system, which had, at other times, notified the government of Rivera's violations of his house arrest restrictions. *Id.*

16. 962 N.E.2d at 60; Trial Tr., *supra* note 15, at 18294.

17. Trial Tr., *supra* note 15, at 018194.

18. 962 N.E.2d at 60.

19. *Id.* at 67–68.

20. *But see* BRANDON L. GARRETT, CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTIONS GO WRONG 1–13 (2011) (providing a survey of wrongful conviction cases, paying particular attention to the frequent abuse of confessions and eyewitness evidence); *id.* at 100–02 (describing other similar cases in which a conviction was obtained despite exonerating DNA evidence).

21. *See* Martin, *supra* note 7, at 47–48 (describing lead prosecutor Andrew Mermel's staunch insistence on obtaining and defending convictions in multiple cases involving exclusionary forensic evidence or confessions by third parties).

already made up his mind that Rivera was guilty[22] allowed a weak case to go before a jury that was either unusually lazy or biased, or both. But I believe this case deserves closer study. The jury's decision in *Rivera* can be explained in terms of well-studied features of unconscious cognition that we all share.[23] This explanation, in turn, makes the trial both more understandable and more disturbing, because it implies that other juries, even if well motivated, might have reached the same perplexing result if exposed to a similar pattern of evidence.[24]

To provide an account of fact-finding that can explain cases like *Rivera*, we must depart significantly from the rationalist tradition of evidential modeling, in search of a framework that can accommodate prejudicial inferences as well as reasonable inferences. If we limit ourselves to rational inferences about probabilities and coherent narratives founded on the admitted evidence, we will be at a loss to explain why twelve ordinary people could have thought that it was not just probable, but beyond any reasonable doubt, that Rivera was Holly's killer. What we need, instead, is a similarly robust way to analyze and describe the way that the jury's views about what inferences were and were not reasonable were shaped, in significant part, by irrelevant or prejudicial factors that operated largely outside the jury's awareness.

Although evidence scholars have spent a substantial amount of time chronicling particular ways that we might fall prey to unconscious biases in the courtroom,[25] few scholars have incorporated

---

22. *See id.* at 68 (noting the defense lawyers' concern that the judge who had twice convicted Rivera was unfairly favoring the prosecution in evidentiary rulings).

23. See discussion *infra* Part III.

24. *Cf.* GARRETT, *supra* note 20, at 100 (noting fifteen other cases in which defendants were convicted despite the admission of exonerating DNA evidence).

25. *See, e.g.*, Jerry Kang et al., *Implicit Bias in the Courtroom,* 59 UCLA L. REV. 1124, 1128–35 (2012) (focusing on implicit bias based on the party's racial identification); Saul M. Kassin & David A. Garfield, *Blood and Guts: General and Trial-Specific Effects of Videotaped Crime Scenes on Mock Jurors*, 21 J. APPLIED SOC. PSYCH. 1459, 1465–70 (1991) (providing empirical evidence on the impacts of gore evidence on verdict decisions in mock trials); Janice Nadler & Mary-Hunter McDonnell, *Moral Character, Motive, and the Psychology of Blame*, 97 CORNELL L. REV. 255, 256–59 (2012) (exploring the impacts of a defendant's bad moral character on a mock juror's construal of evidence items); D. Michael Risinger, *John Henry Wigmore, Johnny Lynn Old Chief, and "Legitimate Moral Force": Keeping the Courtroom Safe for Heartstrings and Gore*, 49 HASTINGS L.J. 403, 445 (1998) (suggesting that gruesome or shocking evidence of a crime's consequences may encourage juries to subtly downplay their doubts in favor of a conviction).

these scattered insights into the broader accounts of inference at trial.[26] We lack, in other words, a framework in which we can examine inferential *prejudice* in the organized and systematic way that the Bayesian and story-comparison accounts bring to the analysis of inferential *probity*.

In this Article, I develop a systemic account of the ways that unconscious cognitive processes interact with the more familiar aspects of judge and jury decision making.[27] In so doing, I will draw heavily on the dual-process tradition in cognitive psychology. This dual-process framework, which cognitive scientists developed to explain a wide variety of seemingly conflicting results in psychological research, posits that our mental processes can be better understood if we group them into two competing cognitive "systems" operating side by side in human decision making.[28]

---

26. For the best-developed prior work in this direction, see Simon, *supra* note 5 (developing an associative-network model of inference that incorporates the phenomenon of coherence shifts) and Paul Thagard, *Why Wasn't O.J. Convicted? Emotional Coherence in Legal Inference*, 17 COGNITION & EMOTION 361 (2003) (reviewing four different computational models explaining O.J. Simpson's acquittal and arguing that the emotional coherence model is the most plausible explanation).

27. Other authors have suggested that this literature might prove useful as a means of elucidating the mechanics of fact-finding, but few have explored the question in depth. *See, e.g.*, Griffin, *supra* note 3, at 299–300 (describing how doubt and difficulty may prompt jurors to abandon logic and utilize an unconscious and associative process when making decisions); Dan Simon, *The Limited Diagnosticity of Criminal Trials*, 64 VAND. L. REV. 143, 184–85 (2011) (analyzing non-evidential aspects that pose a threat to the integrity of criminal verdicts). The most detailed attempt to date at analyzing fact-finding in dual-process terms appears in an article by Chris Guthrie, Jeffrey J. Rachlinksi, and Andrew J. Wistrich, who replicated some cognitive-bias experiments using judges as research participants. After showing that judicial intuition can sometimes lead to errors, they suggest that errors can be systematically reduced by encouraging judges to engage in more deliberative styles of processing. Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, *Blinking on the Bench: How Judges Decide Cases*, 93 CORNELL L. REV. 1, 42–43 (2007). As discussed in Parts III and IV of this Article, my own model complicates this picture substantially by showing that deliberation will often be used to justify and defend intuitive conclusions rather than to override them.

28. *See, e.g.*, DANIEL KAHNEMAN, THINKING, FAST AND SLOW 19–30 (2011) (introducing "System 1" and "System 2" as two modes of thinking); KEITH E. STANOVICH, RATIONALITY AND THE REFLECTIVE MIND 3–25 (2011) (explaining the "dual-process theory" and its role within the "Great Rationality Debate"); TIMOTHY D. WILSON, STRANGERS TO OURSELVES: DISCOVERING THE ADAPTIVE UNCONSCIOUS 17–40 (2002) (discussing the "adaptive unconscious" and its effects on decision making).

One kind of cognition, frequently called "System 1," operates unconsciously and rapidly, producing intuitions that rise to our attention without any preceding awareness of a problem-solving process at work.[29] We rely on such unconscious cognition during every moment of our waking lives, often without realizing it: when we recognize a face in a crowd, when we drive a car along a familiar route, or when we correctly interpret another person's emotional state based on subtleties of their tone of voice or posture, we are performing complicated reasoning that feels effortless.[30]

Another kind of cognition, "System 2" for short, often involves the effortful manipulation of thoughts and images in working memory, or the focus of conscious attention on a particular target.[31] System 2 is strikingly different from System 1 in several important ways: we notice and perceive the work involved in System 2 cognition, we find it effortful and tiring to sustain it for very long, and when we employ it we are constrained by a limited fund of working memory space.[32] Thus, when we do challenging arithmetic in our heads or analyze the import of an unfamiliar and poorly drafted statute, we find it necessary to focus and avoid distractions in order to finish the task, and we will grow tired the longer we persist. Shifting into System 2 thinking offers us the special power to solve problems without recourse to the patterns of our past experience. But the limited capacity of this resource means that we can only focus its power on one problem at a time and must maintain focus and motivation to use it for any significant length of time.

Dual-process cognitive theory gives us powerful tools with which to understand cases like *Rivera*. As I will show in this Article, the prosecution in *Rivera* made use of a number of standard rhetorical strategies that would have been likely to induce strong implicit associations between Rivera and guilt in the minds of the jurors. Chief among these strategies was the early use of gruesome crime scene details, the effort to paint Rivera as a generally untrustworthy person,

---

29. *See* Keith Frankish & Jonathan St. B. T. Evans, *The Duality of Mind: An Historical Perspective*, *in* IN TWO MINDS: DUAL PROCESSES AND BEYOND 1, 15–18 (Jonathan St. B. T. Evans & Keith Frankish eds., 2009) [hereinafter TWO MINDS] (providing a survey of features associated with System 1 and System 2 in the dual-process literature).

30. *See* KAHNEMAN, *supra* note 28, at 19–22 (introducing key features of System 1 and System 2).

31. *See id.* at 22–24 (providing examples such as focusing on a particular person's voice in a crowded and noisy room or parking in a narrow space).

32. *See id.* ("[In System 2 situations,] you must pay attention, and you will perform less well, or not at all, if you are not ready or if your attention is directed inappropriately.").

and the introduction of redundant and repetitive testimony regarding Rivera's confessions.[33] By using these common techniques, the prosecution generated an intuitive sense of Rivera's guilt that outstripped the real probity of the evidence it had to offer. The defense, by contrast, had analytically powerful evidence that was not nearly so well designed to generate a competing set of associative intuitions.[34] The result was that the jury most likely started deliberations with the intuition that Rivera was probably guilty, despite the defense's strong case.

The final piece of the *Rivera* puzzle incorporates analytic, System 2 processing: the jury spent four days in deliberations, most likely finding ways to minimize their doubts about their intuitively preferred result so as to accommodate both their unconsciously derived preference for a guilty verdict and the instruction requiring them to find Rivera innocent unless his guilt was proven beyond a reasonable doubt.[35] In the end, we need to consider the work done by both cognitive systems to explain why it might have felt reasonable and proper for the jury to vote to convict, or to predict when particular kinds of evidence might encourage other juries to reach similarly unreasonable decisions.

The utility of a dual-process model of fact-finding goes beyond explaining the disturbing result in one case. We can also apply its insights to evidence theory itself, as a means of identifying scenarios where widely shared, normative intuitions about the proper force of proof are likely to be unreliable. To illustrate this, I use the dual-process model to reconsider the appropriate weight to give "naked" statistical evidence at trials. Many influential scholars argue that it violates fundamental values of fairness or legitimacy to allow jurors to rest their verdicts on purely statistical evidence.[36] At some deep level,

---

33. See *infra* notes 190–210 and accompanying text.

34. *See infra* notes 211–22 and accompanying text.

35. *See* Trial Tr., *supra* note 15, at 018194.

36. *See, e.g.,* L. Jonathan Cohen, The Probable and the Provable 74–76, 270–72 (1977) (arguing that, in some cases, the verdict that has the higher probability is nevertheless unfair); Cheng, *supra* note 3, at 1265–66 ("[T]he legal system wants the jury to arrive at some narrative of the truth."); Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 Harv. L. Rev. 1357, 1378–85 (1985) (arguing that awarding proportionate judgment in situations involving statistical damages cases and developing an accompanying substantive legal rule is "the primary objective of judicial factfinding"); Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 Harv. L. Rev. 1329, 1340–41, 1346–50 (1971) ("[A] . . . similarity among the cases . . . is less easily dismissed: in all of them, making use of mathematical information first requires transforming it from evidence about the *generality* of cases to evidence about the *particular* case before us.").

these theorists urge, it feels wrong to determine liability or guilt based on a guess, even if that guess is more probable than its alternatives.[37] But this apparent tension is driven in large part by the failure of statistical forms of proof to speak to our unconscious System 1 in a way that would make us feel that the statistically stronger inference is intuitively true. In fact, the ways that our intuitive patterns of reasoning betray us when analyzing these hypotheticals closely parallel what went wrong in the *Rivera* trial. Both scenarios involve analytically powerful evidence that nevertheless lacks the sort of vivid, repetitive detail that can support strong intuitions about guilt or innocence within our underlying associative system of cognition. This insight should moderate, if not eliminate entirely, our anxieties about statistical proof. Taken in connection with the examination of what went wrong in the *Rivera* case, this suggests a broader lesson: before relying on our intuitions about proper case outcomes as a normative guide for optimal juridical decision making, we should first make sure that the process generating those intuitions is generally likely to lead to accurate outcomes.

This Article will proceed as follows. Part I introduces existing models of fact-finding. Part II relates several important aspects of dual-process cognitive theory that will form the basis of my own account of fact-finding in the trial environment. Part III presents a descriptive model of the ways that the two systems of cognition are likely to operate and interact as a fact finder processes the evidence in a disputed trial, using the *Rivera* trial as a discussion example. Finally, Part IV discusses the importance of accounting for the hidden associative component of fact-finding cognition when we consider the utility of purely statistical proof at trial. With this discussion, I hope to introduce an important new framework for future discussions of problems in evidence law, one that can accommodate not just the role of reasoned analysis in fact-finding but also the murky terrain of hunches, insights, and prejudices.

I.   Existing Models of Fact-Finding

A vast amount of scholarship has attempted to provide either formal models or informal accounts of the process by which judges and juries find facts in disputed cases. A full account of this tradition is well beyond the scope of one article and indeed would be hard to

---

37.   S*ee, e.g.,* COHEN, *supra* note 36, at 75 (illustrating how it seems unjust for *A* to lose when he is one of 1,000 people at a rodeo, of which 499 paid and 501 did not, because it is more likely than not that he did not pay). For further discussion of this hypothetical, refer to notes 80–81 and accompanying text.

cover in a lengthy book.[38] My aims in this introduction are far more modest. First, I will explore the major purposes for which these models can be used. Second, I will describe the primary tension that currently divides the literature analyzing how judges and juries reach verdict decisions—with formal models founded on probability theory on one side and, on the other side, informal models that emphasize the construction and comparison of "stories" or "narratives." Finally, I will offer a critique of the tendency in this literature to rely primarily on introspective reflection as a means of understanding how juries reach decisions, and its comparative lack of attention to other ways that we can learn about fact-finding.

A.   The Multiple Uses of Evidence Models

Before I stretch the patience of readers by describing probability networks, abductionist inference, or unconscious associative cognition, it is worth justifying the essential utility of these abstruse endeavors. Models of fact-finding inference have real-world value for anyone who wishes to determine when, and how, our current trial practices might be improved, and this value makes it worthwhile to spend time pondering these abstractions. More particularly, there are at least three different reasons we might want to model the proof process.[39]

First, we might wish to build *descriptive* models of fact-finding to better understand the fact-finding behavior of real-world juries and judges. A good descriptive model aims to describe the processes by which fact finders mentally process case evidence in order to reach verdict decisions.[40] For such a model, the criterion of success is not whether the standards of factual inference it encodes are rational, ideal, or consistent with judicial doctrine or expectations. Rather, a

---

38.   For a short introduction to this literature, see generally Park & Saks, *supra* note 1, at 984–97 (providing a background on "New Evidence Scholarship"). For an excellent book-length treatment, see generally David A. Schum, The Evidential Foundations of Probabilistic Reasoning (1994) (collecting insights about evidence from different disciplines).

39.   *Cf.* Richard Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, 66 B.U. L. Rev. 439, 440–50 (1986) (discussing various reasons why theorists would seek to construct inferential models); Michael S. Pardo, *The Nature and Purpose of Evidence Theory*, 66 Vand. L. Rev. 547, 549–57 (2013) (discussing the dichotomy between descriptive and normative accounts of proof); *see also* Peter Tillers, *Trial by Mathematics—Reconsidered*, 10 L. Probability & Risk 167, 172 (discussing various ways in which formal models of inference can be useful).

40.   *See* Lempert, *supra* note 39, at 448–50 (considering whether the Bayesian approach can be employed as a descriptive model); Pennington & Hastie, *supra* note 5, at 189–92 (elaborating one detailed descriptive account of proof, the "Story Model").

descriptive model succeeds to the extent that it improves our understanding of how fact finders translate a trial's inputs, in terms of evidence, procedures, arguments, and instructions, into outputs, in the form of findings of fact and verdict decisions. Without this understanding, we would have difficulty identifying when we need to intervene in natural reasoning processes by means of evidentiary rulings, and we would also be hard-pressed to predict how interventions would impact juries. Just as a contractor would be ill advised to start excavating a foundation before conducting a survey of the ground, we risk acting contrary to our own interests if we interfere with a fact-finding process whose mechanics we do not grasp.

Second, we might wish to construct *normative* models in order to articulate both the goals we wish to optimize in the fact-finding process and particular ways that we could go about achieving such goals.[41] If the descriptive model of evidence acts like a survey, then the normative model acts like a blueprint, illustrating a desired form that we might wish fact-finding inference to take. Sometimes a normative model might function as an ideal standard against which we measure real-world fact-finding.[42] In other circumstances, it might be adjusted to reflect the perceived limitations of real-world fact finders and to seek second-best improvements given those limits.[43] In either case, we will find it hard to know whether the existing trial process works well, or how it can be improved, without a good normative model of proof.

Third, we might wish to create *doctrinal* models that seek to describe significant features of evidentiary doctrines or judicial attitudes about the nature of fact-finding.[44] Doctrinal models can be of great use to evidence theorists. They may help us predict future

---

41. *See* Jonathan J. Koehler, Audrey Chia & Samuel Lindsey, *The Random Match Probability in DNA Evidence: Irrelevant and Prejudicial?*, 35 JURIMETRICS J. 201, 211–16 (1995) (employing a Bayesian normative model to identify inferential errors in mock-jury decision making); Lempert, *supra* note 39, at 443–48 (discussing a use of Bayes' rule as a normative model).

42. *See, e.g.*, Koehler, Chia & Lindsey, *supra* note 41, at 211–16 (showing how jurors, when weighing the value of a DNA match, tend to overestimate the importance of random match probability in comparison to a normative model).

43. *See* Pardo, *supra* note 39, at 559 (suggesting that normative models fail if they cannot provide guidelines that real-world fact finders are capable of implementing).

44. *See, e.g.*, Lempert, *supra* note 39, at 446–47 (labeling such models as "heuristic devices" that "treat[] the law's rules and procedures as normative and attempt[] to model them"); Lempert, *supra* note 2, at 1021–22 (proposing the use of a Bayesian model as a means of "clarify[ing]" evidentiary rules that "involve weighing evidence in an essentially probabilistic fashion").

rulings on unsettled questions by revealing the hidden intuitions that underlie judicial behavior.[45] Similarly, they may help us teach rules to new lawyers in a way that integrates each rule into a larger, coherent pattern, facilitating both learning and future recall.[46] And finally, doctrinal theories may be especially valuable when placed side by side with well-specified descriptive or normative models. When a good doctrinal model differs from a good descriptive model, it suggests that judges have a faulty understanding of how fact finders actually process evidence. Conversely, if a doctrinal model fails to line up with a normative model, it suggests an area in which there is a need for judicial education or rule reform.

As should be clear from the discussion above, it is quite possible that multiple models, even if dramatically different in their structures, could be useful for those who would attempt to understand or improve the law of evidence.[47] Indeed, it would be an astonishing coincidence if the best version of all three models was identical in its particulars because that would mean that the way that jurors actually reason is both perfectly captured by judicial intuitions about proof and normatively ideal. Since the real world is not likely to be such a rosy place, we will best be able to optimize evidentiary rules and practices only if we are careful to distinguish between descriptive, normative, and doctrinal models of inference.[48]

B. A Brief Taxonomy of Fact-Finding Models

Having set forth the goals of the enterprise, it is now time to explore the lay of the land. Perhaps the easiest way to chart the existing literature on fact-finding models is in terms of one mighty power arrayed against a group of loosely allied enemies. The mighty power is the Bayesian framework, which has been developed in exquisite detail during its nearly fifty years of active use. The allied enemies are the family of theories I have somewhat carelessly lumped

---

45. *Cf.* Tillers, *supra* note 39, at 172 (noting the value of predicting future behavior in litigation).

46. *See* Lempert, *supra* note 39, at 447 (discussing the success of Bayes' rule as a tool for teaching the concept of relevance to students).

47. *Cf.* Simon, *supra* note 5, at 561–62, 561 n.168 (noting that many advocates of existing models attempt to deploy them simultaneously to serve normative and descriptive functions).

48. *See id.* at 561–68 (suggesting that a holistic model functions well as a descriptive account of fact-finding, but for that very reason generates poor normative recommendations for constraining prejudicial influences on the jury); *see also* Koehler, Chia & Lindsey, *supra* note 41, at 210–16 (showing a disjunction between the evidential force of certain information about DNA testing as estimated by a normative Bayesian model and as estimated by a population of mock jurors in an research setting, and using that disjunction as support for suggested law reforms).

together as "story-comparison" frameworks. These theories do overlap in various respects, but perhaps the clearest commonality is their united objection to some or all of the assumptions of Bayesianism. For this reason, the landscape can be most easily understood if we start with the first approach and then explicate the others by comparison and contrast.

1.   The Bayesian Model

The Bayesian approach starts with a simple method—drawn from probability theory—for describing the varying confidence levels a fact finder might have regarding the probative force of a piece of evidence. Probability theory offers a way to describe the likelihood of uncertain events, by grading them on a scale that runs from 0 to 1.[49] In the evidentiary context, we can use these probabilities to represent the varying levels of subjective confidence a person might have regarding a belief.[50] Thus, the Bayesian strain of probability theory invites us to imagine a thermometer[51] of belief, which runs from 0, representing total disbelief in a proposition, through 0.5, representing maximal uncertainty, up to 1, representing total belief. We can then start to consider the effect of new items of evidence on existing confidence levels, asking whether the new evidence makes the existing belief stronger or weaker.[52] Thus, we might say that hearing eyewitness testimony placing a suspect at the scene of a crime moves us to change our degree of belief in his guilt from 0.6 to 0.9. This means that our internal feeling of confidence in the proposition has moved from "I think he did it, but I'm far from certain of it" to "I am quite confident that he is guilty."

The real power of the Bayesian approach is that we can repeatedly apply this updating process, allowing us to describe the process by which a fact finder considers multiple items of evidence before arriving at a verdict.[53] For example, consider the testimony of

---

49.   Pardo, *supra* note 39, at 575 ("[P]robability assessments fall somewhere on a scale between 1 (which equals certain truth) and 0 (which equals certain falsity) . . . .").

50.   *Cf.* Kadane & Schum, *supra* note 2, at 24 (contrasting *aleatory* probabilities that "rest entirely on counting operations involving a finite collection of possible game outcomes, all of which are assumed to be equally probable[ ]," and *epistemic* probabilities, which express a person's subjective belief that a certain event will occur).

51.   *See* 6 Jeremy Bentham, Works of Jeremy Bentham 225 (London, Simpkin, Marshall, & Co., 1843) (developing the metaphor of the "thermometer of persuasion").

52.   Terence Anderson, David Schum & William Twining, Analysis of Evidence 251 (2d ed. 2005).

53.   *See* Kadane & Schum, *supra* note 2, at 127–31 ("Bayes' rule also requires us to take account of a variety of interesting evidential interactions. By an *interaction* we mean that the evidence is

one jailhouse informant in the *Rivera* case, David Crespo, who testified that Rivera confessed to the killing one day after a Bible study class that the two attended in jail.[54] Illustrated below, Bayes' rule provides a way to describe the force that this evidence would have for a juror whose beliefs followed standard rules for combining probabilities.[55]

$$P(G/e) = \frac{P(G) * P(e/G)}{P(e)}$$

This equation tells us how to determine the degree of confidence that our juror should have in Rivera's guilt, *G*, after hearing this new piece of evidence, *e*, which we label as the "probability of *G* given *e*," or *P(G/e)*. In order to determine this quantity using Bayes' rule, we combine (1) the juror's prior confidence in guilt before hearing the new evidence, *P(G)*; (2) the hypothetical level of confidence the juror would have that Crespo's testimony was true assuming that Rivera was guilty, *P(e/G)*; and (3) the degree to which the juror is convinced that Rivera really did confess to Crespo, *P(e)*. So a juror's change in belief based on Crespo's testimony can be predicted based solely on three pieces of information: the previous strength of the juror's belief in Rivera's guilt, the strength of the juror's belief in the truth of Crespo's testimony, and the degree to which the juror believes that Crespo's testimony would be likely to arise given the assumption that Rivera was actually guilty.

Many early Bayesian models imagined the inferential process as a repeated series of these calculations. In John Kaplan's seminal paper, the model started with prior odds of guilt and then—by deciding how likely it would be to encounter that piece of evidence given either a hypothesis of innocence or a hypothesis of guilt—repeatedly adjusted those odds based on each new piece of evidence that was received using Bayes' rule.[56] This model has the advantage of structural

---

    nonindependent in various ways. What this means is that knowledge of one item of evidence may influence our judgment of the probative force of another.").

54.   People v. Rivera, 962 N.E.2d 53, 58 (Ill. App. Ct. 2011).

55.   *See* SCHUM, *supra* note 38, at 215 (explaining posterior and prior probability measures as a way to measure the probability that some hypothesis of interest, *H*, changes after additional evidence, *E**, is introduced).

56.   *See* Kaplan, *supra* note 2, at 1084–85 ("Beginning with what he believes to be Ω_0, the initial likelihood of the defendant's guilt, the decision theorist will examine the first piece of evidence and determine as best he can the value of *L*—the ratio of the probability that the piece of evidence would have occurred under the hypothesis of guilt to the

simplicity, even if quantifying the appropriate likelihood ratio for each new piece of evidence may be a subtle and difficult task.[57]

Lawrence Tribe vigorously attacked this account of fact-finding inference and pointed out that this simple version of the model assumes the conditional independence of the inferences based on individual items of evidence.[58] In lay language, this means that the model relied on an assumption that the force of earlier pieces of evidence is unaffected by subsequently received evidence. As Tribe urged, this is often false with respect to real cases. He offered the example of an armed robbery case in which the break-in took place between 3:00 a.m. and 3:30 a.m. and would have required at least fifteen minutes to accomplish.[59] Evidence placing the defendant in his car a half mile from the scene of the crime at 3:10 a.m. on the night in question might increase the odds of guilt. Similarly, evidence that he was driving in his car a half-mile from crime scene at 3:20 a.m., if evaluated on its own, might have similar force. But the two pieces of evidence, taken together, are exculpatory given the time required to commit the crime.[60]

Based on arguments like this, some modern Bayesians prefer to frame their models in networked terms rather than linear.[61] To deal with potential interdependencies among items of evidence, Bayes' rule requires us to consider the likelihood of observing a piece of evidence not only given the overall hypotheses of guilt or innocence offered by either party but also given each preceding piece of evidence seen in the case.[62] A natural way of addressing these interdependencies is to

---

probability that it would have occurred under the hypothesis of innocence. He will then calculate Ω_1. Similarly, he will find Ω_2—the likelihood of the defendant's guilt after the consideration of the second piece of evidence and before the consideration of the third piece—by considering the equation $\Omega_2 = L_2 (\Omega_1)$.").

57.   Among other considerations, each likelihood ratio depends significantly on the many judgments that coalesce into our assessments of a witness's credibility. *See id.* at 1088–91 (describing how to write an equation that incorporates all requisite potential probabilities for determining a witness's credibility).

58.   Tribe, *supra* note 36 at 1367–68, 1368 n.126.

59.   *Id.* at 1367.

60.   *Id.*

61.   *See, e.g.*, Kadane & Schum, *supra* note 2, at 242–48, 266–67 (discussing the use of inference networks for complex probabilistic analyses); Tod S. Levitt & Kathryn Blackmond Laskey, *Computational Inference for Evidential Reasoning in Support of Judicial Proof*, 22 Cardozo L. Rev. 1691, 1692 (2001) ("We use Bayesian networks ('BNs') to capture the structure of arguments and to provide a numerical representation of their strength." (footnote omitted)).

62.   *See* Kadane & Schum, *supra* note 2, at 127–31.

switch from a linear to a networked model, in which individual evidence items contribute to a final verdict decision only through their contributions to numerous intervening hypotheses about the factual events under dispute. Thus, one interesting way of implementing Bayesian analysis into an account of proof is to chart the relations between individual pieces of evidence, a variety of intermediate hypotheses of interest to a decision maker, and the final "facts of consequence" that will determine the outcome of a trial.[63] Such "Bayesian networks" can then output a single value, which is the likelihood that a party is guilty or innocent of a charged crime.[64] This complex web of inferential connections allows the model to address the connections between pieces of evidence at a granular level and then slowly build up toward the primary facts of interest in the litigation. This is not the only way to address interdependencies among evidence items—one could also address them sequentially, computing the likelihood ratio for each new piece of evidence in light of its interactions with all of the prior evidence items[65]—but the networked approach allows us to visualize those interdependencies.

2.   Bayesianism's Rivals: Stories, Comparisons, and Coherence

Having described the Bayesian approach to modeling trial inference, let us consider the alternative approaches. Perhaps the most prominent alternative account is the "Story Model," developed by Nancy Pennington and Reid Hastie to explain the results of a variety of mock-jury experiments. These authors argued that jurors do not

---

63.   *See id.* at 133–50 (combining Bayesian-probability computations with a Wigmorean-chart structure to better facilitate the separation of independent and nonindependent evidence items); *see also* Judea Pearl, Causality: Models, Reasoning, and Inference 13–21 (2000) (discussing Bayesian Networks).

64.   *See* Kadane & Schum, *supra* note 2, at 195. Joseph Kadane, David Schum, and David Kaiser employed a Bayesian-network model of the major inferences present in the famous trial of Nicola Sacco and Bartolomeo Vanzetti, *Commonwealth v. Sacco*, 151 N.E. 839 (Mass. 1926), and then computed the likelihood that Sacco was guilty based on a Bayesian combination of the inferences, which were laid out in a chart, by multiplying likelihood ratios along the graph's edges. *Id.* at 193–98. This effort produced twenty-eight charts depicting differing subsets of the hundreds of units comprising the larger inference network. *See id.* at 89–115. Nevertheless, the final product was, as the authors acknowledged, substantially simplified compared to the task facing the actual jury. *See id.* at 281 ("[T]he twelve jurors actually saw and heard all the trial evidence; we can only read verbal accounts of it in the trial transcript . . . . [Furthermore,] the [identified] chains of reasoning resulted from . . . our own imaginative reasoning . . . . [and] are not the only ones that may be reasonable or plausible.").

65.   *See, e.g.*, Richard D. Friedman, *Answering the Bayesioskeptical Challenge*, 1 Int'l J. Evid. & Proof 276, 287 (1996).

compute the probabilistic force of individual pieces of evidence but instead construct mental narratives that organize and explain the received evidence.[66] When mock jurors are asked to speak aloud while trying to determine what verdict fits the evidence, their reasoning tends to invoke "story structures" rather than other plausible structures.[67] According to Pennington and Hastie, these constructed stories guide the interpretation of subsequent ambiguities in the evidence, and jurors were more likely to select a particular verdict when the evidence supporting that verdict was presented in a fashion that facilitated the construction of a favorable story.[68] The authors then showed in a subsequent paper that when jurors are asked to provide probability estimates for individual aspects of cases, a Bayesian combination of these individual probabilities performs poorly as a predictor of their final decisions.[69]

Although in their early work they focused on the ways that constructing a single story shaped a juror's verdicts, Pennington and Hastie later incorporated the idea of story *comparison* into their theory. They maintained that jurors construct "one or more" plausible accounts of the disputed events and then choose the most acceptable story from among the generated set to guide their subsequent reasoning.[70] In cases where more than one story is under consideration, they posited that jurors would rely on three principles when choosing between alternatives.[71] First, jurors would tend to prefer stories with better "coverage," meaning those that accounted for a greater share of trial evidence than alternatives.[72] Second, jurors would prefer more "coherent" stories, meaning those that were free of internal contradictions, free of problematic gaps, and plausible based on the juror's knowledge of "real or imagined events in the real

---

66. *See* Pennington & Hastie, *supra* note 5, at 189.

67. *See* Nancy Pennington & Reid Hastie, *Evidence Evaluation in Complex Decision Making*, 51 J. Personality & Soc. Psychol. 242, 252–53 (1986).

68. *See* Nancy Pennington & Reid Hastie, *Explanation-Based Decision Making: Effects of Memory Structure on Judgment*, 14 J. Experimental Psychol. Learning Memory & Cognition 521, 528–30 (1988) ("Subjects were more than twice as likely to find the defendant guilty of murder in our stimulus case when prosecution evidence was ordered in story form and defense evidence was not, compared to when defense evidence was presented in story form and prosecution evidence was not.").

69. *See* Pennington & Hastie, *supra* note 5, at 199.

70. *Id.* at 190.

71. *Id.* at 190–91.

72. *Id.* at 190.

world."[73] In some cases, the authors suggested that one story would clearly dominate the alternatives based on these criteria and would be chosen by the fact finder for the purpose of further decision making in the case. In other cases, by contrast, both sides will offer accounts that are internally coherent and have strong connections to the evidence presented. In such cases, stories "will lack *uniqueness*, and great uncertainty will result."[74]

Others have argued for the priority of comparative frameworks over Bayesian models for quite different reasons. Jonathan Cohen objected not to the use of probabilities in an inferential model but to the *kind* of probabilities that Bayesian models employ.[75] It was a mistake to build an inferential theory using cardinal probabilities, he urged, because they failed to account for the idea of evidential *weight*.[76] In Cohen's view, we should grade the strength of an uncertain inference of fact not merely on the degree to which we currently accept it as true but also on the amount of evidence we had accumulated when evaluating the question.[77] Thus, we might think that something is probably true but also know that there was a good chance that our opinion might change in the future, given that we had examined only a small subset of available evidence. Cohen urged that instead of quantifying uncertain beliefs by reference to a fixed scale with identifiable endpoints, we should instead analyze uncertain beliefs ordinally, limiting ourselves to ranking competing beliefs in comparative terms based on the quantity of particularized evidence that supports either view.[78]

One driver of Cohen's discomfort with the Bayesian approach was his worry that, if taken as a normative model, it recommended outcomes that ill accord with common-sense intuitions about proof when applied to certain hypothetical cases. A point of particular concern involved cases in which proof of liability or guilt was founded on statistical, rather than particularized, forms of evidence.[79] Cohen developed the "Gatecrasher Hypothetical" to illustrate this concern. In this stylized example, 499 people paid for admission to a rodeo, but

---

73. *Id.* at 191.

74. *Id.*

75. *See* COHEN, *supra* note 36, at 121 (arguing that conventional features of judicial proof are incompatible with Pascalian probabilities and suggesting that an inductive conception would fare better).

76. *See id.* at 36–39 (borrowing the concept of evidential weight from John Maynard Keynes but arguing that it should assume a more central place in evidential theory than Keynes assigned it (citing JOHN MAYNARD KEYNES, A TREATISE ON PROBABILITY 71 (1957))).

77. *Id.*

78. *Id.* at 40–41.

79. *Id.* at 74–81.

1000 people were seated on the night of the show. No tickets were issued, and we will assume that a particular attendee, Andrew, is unable to testify regarding his purchase of a ticket. This implies, in a probabilistic framework, that there is a 0.501 probability that Andrew is a gatecrasher.[80] But to Cohen and numerous others in the comparative-evaluation camp, it seems "manifestly unjust" to allow the rodeo to recover its ticket price from Andrew on the basis of such proof alone, given that we would have observed the same evidence if Andrew was actually a member of the ticket-paying minority.[81] Cohen's solution to the paradox invoked his own inductive probabilities. Purely statistical evidence, he argued, failed to provide any inductive support to the idea that Andrew was a gatecrasher; as a result, the plaintiff had no more evidence supporting his case than the defendant had supporting his own, resulting in a verdict for the defense.[82] Some courts have reasoned similarly, distrusting the idea of purely statistical proof and insisting that there be a basis on which a fact finder can "actually believe" that a defendant is liable, rather than merely thinking it statistically more likely than the alternative.[83]

More recently, Michael Pardo and Ronald Allen offered another comparative account of proof as an alternative to Bayesianism, which they labeled "inference to the best explanation" or "abduction."[84] In their theory, fact-finding consists of the construction and comparison of two (or more) differing hypothetical accounts of the events giving rise to the dispute under trial. A fact finder, after hearing the evidence, evaluates which of the parties' stories better explains the evidence they have heard, and then determines which story better explains the given pattern of evidence.[85] In a civil case, the party with the stronger story will prevail, whereas in a criminal case, the defense can prevail even with a weaker story so long as its story is "plausible."[86] In this framing, fact-finding looks like a tournament

---

80. *Id.* at 75.

81. *Id.*

82. *See id.* at 271 (describing how courts, by requiring evidence to be "specifically against" a particular defendant, as opposed to generalized, help prevent injustices attributable to the standard of proof); *see also* discussion *infra* Part IV.

83. Sargent v. Mass. Accident Co., 29 N.E.2d 825, 827 (Mass. 1940); *see also* Nesson, *supra* note 36, at 1358–59 (stating that the goal of the fact-finding process is to generate "acceptable verdicts," not "mathematically probable verdicts").

84. *See* Pardo & Allen, *supra* note 3, at 223–26 (introducing their explanation-based account of judicial proof); *see also* Josephson, *supra* note 3, at 1621–22 (distinguishing among the primary meanings of "abduction").

85. Pardo & Allen, *supra* note 3, at 234–35.

86. *Id.* at 234–39.

between alternative explanations rather than an archeological reconstruction of the probable truth. By choosing the last story standing without demanding that it be probably true, abductionists believe they can avoid some difficulties faced by Bayesian frameworks.[87]

Finally, a few authors have articulated another possible alternative to the Bayesian approach, based on the concept of coherence networks. Such networks are loosely modeled on the ways that neurons interact within the brain. They are connectionist in structure: they incorporate a number of *units*, which can be activated to varying extents, and *connections* between those units, which allow activation in one unit to amplify or inhibit activation in other linked units.[88] Over time, activation levels will propagate across such networks until stable recurring patterns of activation are reached. If we think of the units as representing various beliefs that a juror could have about a case and the connections as embodying intuitive judgments of the degree to which those beliefs are coherent with each other, then this process can be used to simulate the process of choosing one set of mutually supporting facts about a case while rejecting a set of inconsistent facts.[89]

Proponents of these new connectionist inference models generally maintain that they are largely consistent with other work in the comparative tradition but are, for varying reasons, inconsistent with

---

87. For instance, Allen and Pardo believe that their account avoids the conjunction paradox and reference-class problem. *See id.*, at 253, 259–60. *But see* Edward K. Cheng, Essay, *A Practical Solution to the Reference Class Problem*, 109 Colum. L. Rev. 2081, 2084 (2009) (suggesting that there may be principled ways to choose among reference classes); Dale A. Nance, *Naturalized Epistemology and the Critique of Evidence Theory*, 87 Va. L. Rev. 1551, 1567–75 (2001) (arguing that the conjunction problem may be more apparent than real). *See generally* Ronald J. Allen & Michael S. Pardo, *The Problematic Value of Mathematical Models of Evidence*, 36 J. Legal Stud. 107, 111–14 (2007) (describing the reference-class concern); Nesson, *supra* note 36, at 1385–86 (describing the conjunction paradox).

88. *See* Simon, *supra* note 5, at 520–21; Paul Thagard, *Causal Inference in Legal Decision Making: Explanatory Coherence vs. Bayesian Networks*, 18 Applied Artificial Intelligence 231, 235 (2004) [hereinafter Thagard's *Causal Inference*]; Thagard, *supra* note 26, at 362–67. *See generally* William Bechtel & Adele Abrahamsen, Connectionism and the Mind: An Introduction to Parallel Processing in Networks 1–65 (1991) (contrasting networks and symbol systems as two approaches to modeling cognition and then introducing a third approach involving connectionist architecture). For further discussion and application see *infra* Part III.A.

89. *See, e.g.*, Thagard's *Causal Inference*, *supra* note 88, at 237–39 (describing the use of such a network to analyze the competing inferences in the *von Bulow* trial).

formal Bayesian accounts of proof.[90] But despite these claims, the connectionist approach also differs in significant ways from the other anti-Bayesian models that I have described. First, coherence networks do not isolate and compare specific explanatory accounts but instead analyze all the evidence items simultaneously. Second, coherence networks can accommodate and describe some aspects of fact-finding cognition that prior models cannot. Dan Simon used a coherence model to explain the tendency of mock jurors to respond to new items of seemingly incriminating evidence by inflating their assessments of the probative force of other, entirely unrelated evidence against that party. Such "coherence shifts" lie mostly outside the awareness of the jurors themselves and seem to add problematic and irrational aspects to the fact-finding process.[91] Likewise, Paul Thagard has used such models to illustrate how an emotional preference for one conclusion over another might interact with evidentiary strength, using the O.J. Simpson case as an example.[92] Although the utility of connectionist models is to date little explored in the evidence literature, it is a promising new direction on which I hope to build in this Article.[93]

Thus, there remains a hotly contested division between Bayesians and proponents of competing alternatives. The Bayesian approach, on the one hand, provides a simple formalization of uncertainty and characterizes the process of fact-finding in terms of repeated revisions of internal probability estimations based on each new item of evidence. The competing models, by contrast, mostly converge on the idea that evidential evaluation is holistic rather than atomistic; that it requires the consideration of particularized explanatory narratives, not just mere statistical likelihoods; and that narratives can and should be evaluated by comparison with one another rather than in terms of their absolute force.

C.   *What the Existing Discussion Has Taken for Granted*

Having explored the contrast between these two approaches, it is worth stepping back a bit and seeing some of their commonalities. When one surveys this literature, two common assumptions seem

---

90. *See e.g.*, Simon, *supra* note 5, at 562 (suggesting that the global coherence shifts exhibited by such models are inconsistent with the atomistic, piece-by-piece assessments of probative force required in a Bayesian model); Thagard's *Causal Inference*, *supra* note 88, at 242–43 (suggesting that the coherence network is more intuitively plausible because Bayesian networks require the specification of a large number of obscure and unintuitive probabilities, which is not a feature that one observes introspectively when considering the inferential force of evidence).

91. Simon, *supra* note 5, at 547.

92. *See* Thagard, *supra* note 26, at 372–75.

93. *See* discussion *infra* Part III.

widespread. The first assumption is that fact-finding is largely an explicit process that is accessible to introspective analysis. Thus, when we speak of probabilities in a descriptive Bayesian model, the usual assumption seems to be that these represent the explicit beliefs held by fact finders.[94] Likewise, when Pennington and Hastie criticized the Bayesian model for lacking descriptive realism, they focused on conscious reasoning processes almost exclusively. To demonstrate that juror decisions are poorly predicted by Bayes' rule, they relied on mock jurors' self-reported probabilistic estimates of the strength of various trial inferences.[95] Similarly, when they wished to suggest that story construction played a more important role, they relied on the thoughts that mock jurors reported aloud while deliberating about a case.[96]

The second assumption is that evidential reasoning is best modeled in ways that largely correspond with our common-sense assumptions about the reasonable force of various kinds of proof. Although this assumption is rarely defended explicitly, it forms the bedrock of the standard set of approaches to answering questions in this field. Neither the Bayesian probabilistic framework nor the various comparative approaches provide much of a basis for explaining why judges and jurors sometimes act in ways that seem unreasonable when analyzed from a reflective distance. The two primary exceptions to this trend have been the connectionist models recently deployed by Simon and Thagard,[97] but these authors' attempts to incorporate motivated cognition and emotional bias into models of the fact-finding process have had distressingly little impact on the continuing debates between Bayesians and Bayesioskeptics.

Of course, to the extent that these models are used normatively, as a means of recommending best practices rather than describing what judges and juries are likely to do in real cases, we are unlikely to be troubled by either of the above assumptions. It would be difficult, after all, to imagine a useful way of identifying ideal inference processes without making those processes both explicit and reliant on defensible inferential arguments. But even if these two assumptions are virtues when applied normatively, they limit the reach of our theorizing when applied in a descriptive context. First, it seems to be

---

94. *See, e.g.*, Kadane & Schum, *supra* note 2, at 24 (noting that they used subjective, or "epistemic," probabilities as the basis of their analysis of the Sacco and Vanzetti case); Nance, *supra* note 87, at 1600–06 (characterizing Bayesian models as consistency constraints on the subjective probability estimates of fact finders).

95. *See* Pennington & Hastie, *supra* note 5, at 197–201 (describing an experiment designed to test the Story Model).

96. Pennington & Hastie, *supra* note 67, at 245–46.

97. Simon, *supra* note 5, at 520–49; Thagard, *supra* note 26, at 372–76.

an incontrovertible fact that judges and juries sometimes act indefensibly.[98] But unless we employ accounts of fact-finding that explain such behavior, any rule-reform suggestions we might make based upon our models will be ill equipped to deal with our most problematic cases. Second, as I discuss below, an ever-larger body of research suggests that a substantial component of real-world judgments and inferences are derived from an unconscious system of reasoning that does not obey the common-sense rules of explicit reasoning. It is my hope that by incorporating this hidden structure of thought into our conception of fact-finding behavior, we can also make headway toward understanding why judges and juries sometimes act in ways that seem to defy reflective common sense.

II. A Primer on Dual-Process Cognition

In recent years, research programs in psychology and neuroscience have produced a consistent pattern of results, suggesting that our minds can process information in two dramatically different ways.[99] On some tasks, answers come quickly, effortlessly, and automatically, while at other times, we consciously experience a slow, effortful process of reasoning to a conclusion.[100] Increasingly, researchers have come to identify these differences as indicative of two different modes of reasoning and judgment, working side by side in our minds. Sometimes these modes work collaboratively, but other times they conflict with each other. Many psychologists refer to these two modes of processing as "System 1" and "System 2,"[101] and I will follow that

---

98. *See, e.g.*, People v. Rivera, 962 N.E.2d 53, 55–60 (Ill. App. Ct. 2011) (describing a jury conviction after the presentation of highly exculpatory evidence); Griffin, *supra* note 3, at 282–85 (discussing the unsettling case of *State v. Jensen*, 794 N.W.2d 482, 490 (Wis. Ct. App. 2010), in which a jury convicted a man of murdering his wife despite a number of facts suggesting that it could have been a suicide); *see also* Susan L. Gellis, *Reasons for Case Reversal in Texas: An Analysis*, 16 St. Mary's L.J. 299, 307, 318 (1985) (finding, in a sample of Texas appellate decisions, that a substantial fraction of reversals were based on the appellate court's conclusion that there had been insufficient evidence, or no evidence at all, supporting the jury's verdict).

99. *See, e.g.*, Jonathan St. B. T. Evans, *Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition*, 59 Ann. Rev. Psychol. 255, 256–58 (2008); Frankish & Evans, *supra* note 29, at 1; Keith E. Stanovich & Richard F. West, *Individual Differences in Reasoning: Implications for the Rationality Debate?*, 23 Behav. & Brain Sci. 645, 658–59 (2000).

100. *See* Kahneman, *supra* note 28, at 19–21.

101. *See, e.g.*, *id.*; Peter Carruthers, *An Architecture for Dual Reasoning, in* Two Minds, *supra* note 29, at 109, 109–12; Keith Frankish, *Systems and Levels: Dual-System Theories and the Personal-Subpersonal Distinction, in* Two Minds, *supra* note 29, at 89, 96–102 (2009); *see*

convention in this Article. Others are wary of the metaphor of structurally distinct systems and prefer to think of the two modes of processing as a family of "Type 1" processes, which share many important features, and a family of "Type 2" processes, which do not share the features of the "Type 1" family.[102]

At the outset, it is important to be as clear as possible about what, precisely, we are claiming when we describe cognition as involving two systems or processes. We need not think that a two-system mind uses two physically separable brain structures; rather, these terms merely denote two categories of mental processes that can be usefully separated because the individual processes share certain common features.[103] To say that we have dual-system minds is to say that the ways in which we learn, encode, and use knowledge when solving problems can take two very different forms and that, at times, these differing processes can produce conflicting outputs.[104] In this section, I will outline the terrains in which these two systems operate, their unique methods of processing information, and the ways that they interact with each other.

---

also Jonathan St.B. T. Evans & David E. Over, Rationality and Reasoning 141 (1996) (providing an early use of the "dual systems" terminology); Frankish & Evans, *supra* note 29, at 15 (tracing the origin of the "System 1/System 2" locution to Keith E. Stanovich (citing Keith E. Stanovich, Who Is Rational?: Studies of Individual Differences in Reasoning (1999))).

102. *See, e.g.*, Stanovich, *supra* note 28, at 18–19 (explaining his preference for a Type 1/Type 2 terminology to capture a "dual-*process* theory" as opposed to a "dual-*system* theory"); Evans, *supra* note 99, at 270–71 ("We might be better off talking about type 1 and type 2 processes since all theories seem to contrast fast, automatic, or unconscious processes with those that are slow, effortful, and conscious."). Some theorists further complicate the account by bifurcating System 2 into multiple systems. For instance, Stanovich recently proposed a tri-partite model, which maintains the distinction between Type 1 and Type 2 processes while further subdividing the latter category into algorithmic and reflective processes. Stanovich, *supra*, note 28, at 25–26. But these refinements of the model's terminology should not distract from the basic, widely shared insights that (1) there is more than one type of mental processing and (2) at least one important analytic cleavage is captured by the dichotomy between Type 1 processes and other styles of processing. *Id.* at 33 (quoting Daniel T. Gilbert, *What the Mind's Not*, *in* Dual-Process Theories in Social Psychology 3, 3–4 (Shelly Chaiken & Yaacov Trope eds., 1999)).

103. *See* Kahneman, *supra* note 28, at 28–29 (explaining that System 1 and System 2 are "fictional characters" used for explanatory purposes).

104. *See* Stanovich, *supra* note 28, at 16–17, 19–22.

A.   System 1 Is a Hidden, Unconscious Information Processor

System 1 processing is implicit rather than explicit. This means that the work it does occurs outside conscious awareness, so that we are aware of our conclusions but not the process that produced them.[105] For some, this may come as a startling suggestion; "how," you might wonder, "can I think about something without knowing that I am thinking?" But reflection shows that we are actually unaware of most of our own cognition. Take, as a simple example, the act of driving to work on a typical, uneventful morning.[106] Over every second of our journey, we are performing a simply astonishing quantity of computation. Reflected light enters our eyes in a jumble of sensation, and we process it extensively to form a model of the world around us, in which objects are distinct from one another and related to one another in a three-dimensional spatial model. Additionally, we are automatically sorting objects in this model so that we give prioritized attention to signals that relate to the task at hand. We mostly do not notice the trees and the businesses that line the side of the road, the details of the cars surrounding us, or the clouds in the sky, but we respond to traffic signals or brake lights with automatic ease.[107] When we wish to move our vehicle in a particular direction, we must move our hands and feet in precise sequences to manipulate vehicular controls, but we do not have to think about what we will do before we do it.

Sometimes, we can infer that System 1 is at work just based on the fact that we can perform some mental tasks without felt effort or focused awareness; thus, when we drive a car without thinking about it, we know that some unconscious mental activity must be doing the heavy lifting.[108] But System 1 does not only step up to the plate when

---

105. *See* WILSON, *supra* note 28, at 17–41 (2002) (describing the multitude of roles performed by implicit System 1 processes, which he collectively labels the "adaptive unconscious"); Ron Sun et al., *The Two Systems of Learning: An Architectural Perspective*, *in* TWO MINDS, *supra* note 29, 239, 239–42 (discussing the distinction between implicit and explicit learning and knowledge).

106. KAHNEMAN, *supra* note 28, at 29.

107. *Id.* at 21–22 (noting that System 1 processes can, with effortless ease, direct our attention where it is needed from moment to moment).

108. With a little introspection, we can identify many other situations where we perform impressive computational feats with little to no conscious attention. Other examples include walking down a crowded street, recognizing familiar faces, performing routine tasks like getting dressed, and understanding the meaning of sentences written in a familiar language. When we feel an impulse to walk, we just move from point *A* to point *B* without thinking about our balance or the need to avoid obstacles. Likewise, when we see a simple sentence written in our native language, we usually know what it means rapidly and without conscious analysis. *See id.*

we stop paying attention; it is constantly working in the background of our minds and shaping the way we think even when we are consciously attending to a particular task.[109] So to study the nature of this hidden but pervasive form of mental processing, it is necessary to measure how people process information that they are unaware of or that they think is irrelevant to the task at hand.

An extensive literature has sprung up that does exactly this, showing that people's decisions are regularly impacted to a significant extent by factors that they do not consciously perceive.[110] Some of the clearest examples of such implicit cognition involve people with brain disorders. For example, people with blindsight are unable to perceive certain regions of their visual field consciously. If you display objects to them so that they are only visible in the blinded part of their vision and then ask them what they saw, they will reply that they have no idea. But if you let them try to pick the objects out of a larger group, they will choose the right ones at levels well above random chance, all the while asserting that their choice was a mere guess.[111] Nor are such phenomena limited to people with brain disorders; when ordinary people view words that have been flashed on a screen too quickly to be noticed consciously, they will change their behavior in response to the subliminal cue without realizing that it even occurred.[112] Our eyes take in, and our brains interpret, more information than we ever consciously perceive.

What is perhaps most striking about this automatic and implicit processing is the extent to which it goes on, without our awareness, even regarding stimuli that we are processing consciously in different ways. As we encounter objects, words, or concepts in our daily lives, features of these objects will affect our subsequent behavior even when we are not aware of any connection. A full catalogue of such effects is impossible, but a few examples may illustrate the point:

- When evaluating ambiguous vignettes, people were more likely to see immoral behavior in those vignettes if they had

---

109. *See* WILSON, *supra* note 28, at 14, 40–41 (describing the extent to which unconscious associative processes mediate nearly all significant human activities—including learning, evaluation, and goal-setting).

110. *See id.* at 44–48 (noting that even what we perceive as consciously determined choices can be the product of prior unconscious cognition).

111. Daniel L. Schacter, *Implicit Knowledge: New Perspectives on Unconscious Processes*, 89 PROC. NAT'L ACAD. SCI. U.S. 11113, 11113–15 (1992).

112. *See* WILSON, *supra* note 28, at 30–31 (describing an experiment in which participants, after being primed with extremely short displays of words like "hostile," "insult," and "unkind," subsequently interpreted other people's behavior in more negative ways than did a control group, even though they did not report any awareness they had seen the words).

recently been exposed to unrelated disgusting stimuli. Thus, bad smells, a disorderly room, and being asked to recall disgusting experiences all made it more likely that people would view particular behavior as immoral.[113]

- When subjects are asked to evaluate a speaker's credibility, they will inflate their ratings of that speaker's honesty if the speaker is attractive.[114] Similarly, when research participants were asked to monitor the competence of workers who were performing a maze-solving task, they generally believed that the more attractive workers were better at solving the maze, even though they were not.[115]

- People who have been primed with images of money become more independent and less helpful. They are less likely to ask for help on difficult tasks and less likely to provide help when others request it. They even choose to sit farther away from other people than people who were not similarly primed.[116]

As these examples show, unconscious cognitive processing extends well beyond the domain of behaviors we think of as habitual or automatic. Rather, implicit System 1 processing plays a role in determining our beliefs and actions even in the domains we ordinarily think of as embodying rational or moral behavior.

B. The Associative Structure of System 1 Unconscious Thinking

Having explored one means of differentiating System 1 from System 2 processes—the fact that System 1 processes occur outside of our conscious awareness—let us now consider a second point differentiation: the associative structure of System 1 thinking. System 1 processes seem to do much of their work by sensing familiar patterns in sensory data and assimilating those patterns to known concepts, forming an unconscious associative network.[117] Sometimes,

---

113. Simone Schnall et al., *Disgust as Embodied Moral Judgment*, 34 Personality & Soc. Psychol. Bull. 1096, 1097–104 (2008).

114. Marc-André Reinhard & Siegfried L. Sporer, *Content Versus Source Cue Information as a Basis for Credibility Judgments: The Impact of Task Involvement,* 41 Soc. Psychol. 93, 95–97 (2010).

115. Markus M. Mobius & Tanya S. Rosenblat, *Why Beauty Matters,* 96 Am. Econ. Rev. 222, 228–34 (2006).

116. Kathleen D. Vohs et al., *The Psychological Consequences of Money*, 314 Science 1154, 1154–56 (2006).

117. *See* Kahneman, *supra* note 28, at 50–58 (providing examples of priming and the "complex constellation of responses [that can] occur[ ] quickly, automatically, and effortlessly" as a result of System 1 processes); Wilson, *supra* note 28, at 24–27 (discussing "the adaptive unconscious as the pattern detector"); Steven A. Sloman, *The Empirical Case for*

activating these concepts directly prompts us to action. At other times, the activation of one mental pattern or concept may spread, without our knowledge, to other related concepts, which then affect our subsequent behavior.[118] The common thread that ties these influences together lies in System 1's ability to automatically form connections between stimuli that tend to occur together, which it uses to guide our behavior when those stimuli reoccur. System 1, in other words, learns by association and uses associative connections between unconsciously represented concepts to prime us toward beliefs, decisions, and actions that have been associated with similar stimuli in the past.[119]

First, let us explore the idea that a fundamental aspect of System 1 cognition involves assimilating sensory information to familiar patterns. We are all familiar with one dramatic example of System 1's associative power, which is the effortless ease with which we can recognize the faces of people we know even under poor viewing conditions.[120] As an example, gaze for a minute at Figure 1.

Two Systems of Reasoning, 119 Psychol. Bull. 3, 4 (1996); Eliot R. Smith & Jamie DeCoster, *Dual-Process Models in Social and Cognitive Psychology: Conceptual Integration and Links to Underlying Memory Systems*, 4 Personality & Soc. Psychol. Rev. 108, 111 (2000). *But see* Evans, *supra* note 99, at 261 (doubting that some versions of the dual-process framework, such as the heuristic/systematic processing account, can be reconciled with associationism). For further discussion of associative networks and modeling see *infra* Part II.B.

118. *See* Timothy P. McNamara, *Theories of Priming: II. Types of Primes*, 20 J. Experimental Psychol.: Learning Memory & Cognition 507, 514–16 (1994) (providing evidence that priming effects can exhibit spreading activation patterns).

119. Smith & DeCoster, *supra* note 117, at 113–14.

120. *See* Keith E. Stanovich & Richard F. West, Authors' Response, *Advancing the Rationality Debate*, 23 Behav. & Brain Sci. 701, 715 (2000) (including face-detection on a list of routine System 1 cognitive tasks).
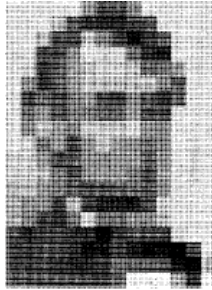
Figure 1: A field of shaded squares, arranged to form a particular pattern.[121]

At first, many viewers who are unfamiliar with this illusion may see nothing but arbitrary noise. If they persist, however, most will eventually perceive the outline of a face (it may help to hold the image at a distance or to slightly blur one's vision), which they will recognize as belonging to Abraham Lincoln. Indeed, once we see the image, it is hard to make it go away or to see the squares as anything but a portrait.[122] At the same time, though, it is quite hard to put into words what about the squares, exactly, that makes them so evocative of Lincoln.

For our next example, consider the poster in Figure 2.

Figure 2: Poster for a collegiate drama program.[123]

Did you notice anything wrong? If not, look again!

---

121. From Leon D. Harmon & Bela Julesz, *Masking in Visual Recognition: Effects of Two-Dimensional Filtered Noise*, 180 Science 1194, 1194 (1973). Reprinted with permission from the American Association for the Advancement of Science (AAAS).

122. Howard Margolis, Patterns, Thinking, and Cognition: A Theory of Judgment 37–39 (1987).

123. John P. Frisby & James V. Stone, Seeing: The Computational Approach to Biological Vision 201 fig.8.32 (2d ed. 2010). Reprinted with permission from John P. Frisby.

For many of us—including the people who initially had this poster printed and distributed before noticing that anything was wrong with it—the misspelling of the word "Spring" is initially surprisingly hard to see. Our brains see most of the letters in the word "Spring" in a context in which that word would be appropriate and fill in the gap to make the message make sense. But once we have recognized the mistake, we no longer find it difficult to notice; if anything, it is hard to take our eyes off it![124]

These two images, taken together, help illustrate some profound facts. First, we can recognize familiar patterns, even when they are obscured by an astonishing amount of noise. We constantly depend on this ability without realizing how much mental work is actually involved in making it possible.[125] Indeed, some have suggested that many dazzling expert performances, such as a chess master's ability to quickly recognize a bad move, should be attributed more to the expert's familiarity with a large number of situationally significant patterns than with the expert's application of rule-like knowledge.[126] System 1 allows us to recover meaning even in garbled communications. If the garbling is small, our associative minds may recover the appropriate meaning without our ever noticing that something is wrong. Indeed, if we relax and let System 1 do its associative work, we can often recover meaning in astonishingly degraded information streams.[127]

Second, these two images illustrate the fact that System 1 does not just allow us to notice patterns; it bosses us into seeing those patterns to the exclusion of other possibilities.[128] The drama poster shows this most dramatically. At first glance, we may automatically read the word as "Spring" without noticing anything wrong because System 1 automatically associated a related pattern of letters to a known pattern. But once we have seen the misspelling, there is no going back; System 1 now associates the cluster of inputs to the somewhat odd and unfamiliar word, "Sring," and cannot see it any other way.

System 1, it seems, can recognize patterns with fluid ease, but it tends to see only one pattern at a time, even when stimuli are

---

124. MARGOLIS, *supra* note 122, at 10–11.

125. *See* Sharon Gilad-Gutnick et al., *Recognizing Degraded Faces: The Contribution of Configural and Featural Cues*, 41 PERCEPTION 1497, 1497, 1506–08 (2012).

126. *See* KAHNEMAN, *supra* note 28, at 11.

127. Idneed, msot radeers wlil hvae reltaievly lttile tourble decdoing tihs sreioulsy gabrled snetnece bceuase tehy wlil rpadily aissmliate each mispeleld wrod to a fmaliiar pttaern.

128. *See* MARGOLIS, *supra* note 122, at 39–41 (describing this tendency as "pattern-seeking, pattern-dominated cognition").

somewhat ambiguous between multiple possibilities. This is illustrated most dramatically by the classic illusion of the Necker cube, included below as Figure 3.
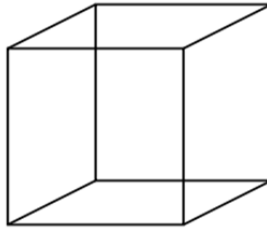

Figure 3: The Necker cube.

The Necker cube shows us the edges of a three-dimensional cube projected onto a two-dimensional plane, and as such, it is ambiguous. We can view it so that either the top right or the lower left vertex are closer to us, with the other on the "back" of the cube. But what is striking is that we cannot see the cube as ambiguous between the two states. Seeing the cube as "in-between"—or seeing it as merely a flat drawing of lines meeting at odd angles on a sheet of paper—is nearly impossible to do. System 1, it seems, strongly assimilates inputs to a single pattern and resists ambiguity in classification.[129]

Perhaps surprisingly, System 1 can perform its pattern-recognizing function even when we cannot perceive any patterns consciously. In one experiment, participants in a gambling game were offered a choice among several decks of cards from which they could draw, with rewards and penalties distributed seemingly at random among the cards in each deck.[130] This apparent randomness concealed a pattern: over time, the gains and losses in some decks averaged out

---

129. *See* Paul Thagard, *Explanatory Coherence*, 12 Behav. & Brain Sci. 435, 438–39 (1989); *see also* Keith E. Stanovich, *Distinguishing the Reflective, Algorithmic, and Autonomous Minds: Is It Time for a Tri-Process Theory?*, *in* Two Minds, *supra* note 29, at 55, 68–70 (exploring the concept of focal bias, in which an intuitively attractive representation becomes focal and dominates further cognition, leading participants to hunt for confirmatory evidence rather than seek further information in a neutral way).

130. Antoine Bechara et al., *Deciding Advantageously Before Knowing the Advantageous Strategy*, 275 Science 1293, 1293 (1997). Further evidence of unconscious pattern-learning effects have been developed in experiments whenever a certain stimulus arises in an ongoing sequence. *See generally* Arnaud Destrebecqz & Axel Cleeremans, *Can Sequence Learning Be Implicit? New Evidence with the Process Dissociation Procedure*, 8 Psychonomic Bull. & Rev. 343 (2001) (finding that when data are structured in a way that incorporates complex patterns, study participants are often able to exploit the pattern unconsciously, achieving faster reaction times than in control conditions, even though they are not aware that a pattern exists).

to a net gain, while in others they averaged out to net losses. Intriguingly, participants seemed to learn which decks were favorable before they could verbalize that knowledge; they grew more likely to choose the winning decks, and they showed signs of stress when drawing from the losing decks even while they insisted that the decks seemed equivalent. Thus, they had unconsciously discovered a pattern without any conscious access to that knowledge.[131] To the extent that our associative minds can notice patterns more subtle than we can consciously perceive, they may produce intuitions that outpace our ability to consciously explain.

Now that we have seen the surprising power of System 1, let us consider some of the ways it may lead us astray. One key difference between System 1's associative engine and explicit System 2 cognition is that the former is unable to ignore particular inputs merely because they are "irrelevant." The result is that inputs we would confidently say are irrelevant to our current tasks sometimes affect our behavior without our knowing.[132]

The phenomenon of associative "priming" demonstrates System 1's inability to screen out some stimuli as situationally irrelevant. A standard way to demonstrate priming effects involves first exposing a participant to a concept in a manner too subtle for them to perceive and then measuring its impact on their reaction time in other, related tasks.[133] For example, studies show that if we test people on the speed with which they can distinguish English words from meaningless garbled letters, a person who had been primed with the concept "chair" via a quick, subliminal flash of that word on a screen would show faster reaction times when identifying related words like "table," but not unrelated words like "cow."[134] Because of this variance in reaction times, we know that the unconscious representation of the "chair" concept has been activated, even though it was without the person's awareness.

Associative activations can also spread through intervening mental concepts, causing some stimuli to have truly surprising effects on judgment and behavior. Thus, for instance, seeing the word "gift" has primed study participants to more swiftly recognize the word "pie," even though gifts and pies have no direct experiential relation

---

131. Bechara et al., *supra* note 130, at 1293.

132. *See generally* WILSON, *supra* note 28, at 93–115 ("How well do people know the causes of their judgments, feelings, and actions? There are cases in the psychological literature of people who are so ignorant of why they respond the way they do that they have to invent explanations.").

133. Schacter, *supra* note 111, at 11115.

134. *Id.*

for most people, because gifts associate to birthdays, which associate to cake, which associates to pie.[135]

Priming may have strong adaptive value, allowing us to respond more quickly to stimuli that predictably occur in particular settings. But priming has a darker, stranger side as well. By placing ourselves in a state of readiness to encounter certain patterns in our environments, we may also increase the frequency with which we will perceive those patterns in ambiguous stimuli, and thus introduce a biasing tendency into our perceptions, choices, or behavior.[136] Once we have started to see a situation in a particular way, we may have trouble seeing it any other way, regardless of additional evidence.[137]

In the simplest form of this phenomenon, stereotypical primes that we encounter in our environments may shape our perceptions or decisions. Thus, in a disturbing set of experiments, participants primed with cues that activate mental patterns associated with African American stereotypes, such as hearing violent rap music, were more likely to associate an individual in a story who has an African American–sounding name with hostility, sexism, and low intelligence.[138]

When one adds spreading associative connections into the mix, the picture gets even stranger. Consider, for instance, the "halo effect," which describes our tendency to associate one type of positive trait with other, unrelated traits that also have a positive valence.[139] For instance, most people exhibit a "beauty bias": they tend to assume that attractive people will exhibit heightened forms of a variety of positive characteristics, such as kindness, intelligence, and honesty.[140] Perhaps of greatest relevance to the legal system, we unconsciously expect pretty people to be truth-tellers and thus more likely to believe what they say, even if we have other means of assessing their credibility.[141] Most of us would find the idea that we

---

135. *See* McNamara, *supra* note 118, at 507, 514–16.

136. *See* Wilson, *supra* note 28, at 29–30.

137. *See* Stanovich, *supra* note 28, at 68–70.

138. *See* Laurie A. Rudman & Matthew R. Lee, *Implicit and Explicit Consequences of Exposure to Violent and Misogynous Rap Music*, 5 Group Processes & Intergroup Relations 133, 147 (2002) (finding that negative stereotypes of African American "may be perpetuated by (some) rap music, at least in college-aged subjects").

139. *See* Andrew J. Wistrich, Chris Guthrie & Jeffrey J. Rachlinski, *Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding*, 153 U. Pa. L. Rev. 1251, 1266–67 (2005).

140. Saul M. Kassin & Lawrence S. Wrightsman, The American Jury on Trial: Psychological Perspectives 100 (1988); Reinhard & Sporer, *supra* note 114, at 95–97.

141. *See* Reinhard & Sporer, *supra* note 114, at 101–02.

are systematically more likely to find attractive people as credible hard to defend, but such behavior follows a straightforward associative logic. Beauty, like honesty, is something we praise culturally, something we generally find pleasing, and something we view as good rather than bad. Thus, activating our mental representation of attractiveness primes us to see good, pleasant, praiseworthy things, and honesty is another pattern matching that description. For this reason, if we are facing a close question regarding a person's believability, their attractiveness may play a role in determining the answer we find intuitively plausible.[142]

One further, surprising aspect of associative inferences is that our physical and emotional feelings can also activate associations that exert an unconscious influence on our choices and behavior. One example of this phenomenon, mentioned above, involves the role of disgust in moral evaluation. Part of our emotional reaction to serious moral violations can involve a physical feeling of disgust or revulsion. But in the logic of associative processing, such connections can also occur in reverse: because we have felt disgusted when sensing norm violations in the past, we become primed to see norm violations whenever we feel disgusted, regardless of the reason why.[143] Thus, researchers have found that inducing disgust by clearly irrelevant means (such as by making a room look disorderly or introducing a foul smell into the air) makes research participants more likely to condemn ambiguous behavior as immoral.[144]

So it turns out that associative cognition can affect our judgments and decisions in ways that are far from obvious. Although I have set out some ways in which the associative system can let us down from time to time, I do not wish to be alarmist. Most of the time, associative cognition is a powerful and useful tool, allowing us to recognize patterns based on subtle features even when those features are blurred by noise or hard to put into words. But if we are to provide a useful account of the role of such processing in fact-finding

---

142. In another striking example of spreading associative activity, participants who were exposed to physical warmth became more likely to see other people as emotionally warm. *See* Lawrence E. Williams & John A. Bargh, *Experiencing Physical Warmth Promotes Interpersonal Warmth*, 322 SCIENCE 606 (2008).

143. Schnall et al., *supra* note 113, at 1105.

144. *Id.* at 1097–105. Similarly, we might feel cross or angry when we ponder moral violations, and thus we might expect feelings of irritation to produce similar effects. This is just what we observe. In a recent study, Israeli judges turned out to be stricter in their sentencing decisions if they had been sitting on the bench for a long time without a break. *See* Shai Danziger et al., *Extraneous Factors in Judicial Decisions*, 108 PROC. NAT'L ACAD. SCI. U.S. 6889, 6889 (2011).

cognition, we must keep both its strengths and its weaknesses in mind.

C. *System 2 Processes Information in a Conscious, Effortful Manner*

Having considered System 1 processes in some depth, let us now turn our attention to System 2 cognition. One of the first key hallmarks of System 2 processing is that it takes place within the explicit view of conscious awareness and is thus available to introspective awareness.[145] When System 1 gives us an answer, we know what the answer is but not how we arrived at it. By contrast, when System 2 is doing the work, we are aware of the work that goes into producing the outcome and experience a series of internal steps involved in getting there.[146]

A second, closely related hallmark of System 2 cognition is that it is ego depleting, meaning that we find it mentally taxing to employ and need to use willpower in order to keep it going.[147] All of us, I am sure, have experienced this feeling, such as when trying to do a complicated bit of mental math.[148] In this respect, System 2 cognitive processes seem analogous to physical exercise. People generally save their mental strength if they know they will have to use it in the future, and if they have recently expended their willpower to focus on a task, they will find it harder to repeat the performance until they have rested.[149] So, while our associative cognition happens automatically and effortlessly, our System 2 is activated only with an effort, and we will find ourselves weary if we use it for too long without taking a break.

A third significant feature of System 2 is that it often involves the manipulation of words, symbols, or images held in working memory.[150]

---

145. *See* Evans, *supra* note 99, at 257–58; *see also* Carruthers, *supra* note 101, at 109–10; Frankish, *supra* note 101, at 96–97.

146. *See* Evans, *supra* note 99, at 258–59 (noting that System 2 functions by means of a sequential flow of conscious thinking through a limited-capacity working memory system).

147. *See* Martin S. Hagger et al., *Ego Depletion and the Strength Model of Self-Control: A Meta Analysis*, 136 Psychol. Bull. 495, 496 (2010); E.J. Masicampo & Roy F. Baumeister, *Toward a Physiology of Dual-Process Reasoning and Judgment: Lemonade, Willpower, and Expensive Rule-Based Analysis*, 19 Psychol. Sci. 255, 259 (2008). *See generally* Rebecca Hollander-Blumoff, *Crime, Punishment, and the Psychology of Self-Control*, 61 Emory L.J. 501, 538–43 (2013) (reviewing literature on the strength model of self-control).

148. A simple demonstration of the sensations involved can be obtained by mentally adding together any two randomly chosen four-digit numbers.

149. *See* Hagger et al., *supra* note 147, at 496–97; Masicampo & Baumeister, *supra* note 147, at 259–60.

150. *See* Evans & Over, *supra* note 101, at 154; Frankish, *supra* note 101, at 92–93.

A bit of introspection shows that many of our conscious thoughts take linguistic or imagistic form. If we try to monitor the nature of our own thoughts, we may hear words in our mental ears or see images in our mind's eye. For some types of System 2 cognition, such sounds and images may be an integral part of mental processing.[151] For instance, when we add a column of numbers in our heads or consider the validity of a logical argument, we often hold symbols in mind in sequence, mentally applying rules in order to transform them in steps toward a final result.

To be clear, the fact that symbols or mathematics is at stake is not a necessary sign of System 2 processing; if we are familiar enough with a particular problem, we may have integrated its answer into our underlying associative system. Thus, any adult who has learned his multiplication tables as a child knows the answer to the question, "What is 4 times 5?" without having to manipulate any numbers in active memory. But if the question shifts to, "What is 105 times 32?," no answer will spring to mind, and our associative system will be dumbfounded. Instead, we can solve the problem only by applying a series of learned rules, and, in so doing, we will experience all of the hallmarks of System 2 cognition. Our conscious minds will be occupied, and we will experience the task as effortful. It is this sort of symbolic manipulation that is another core, identifying feature of System 2 processing. System 1 can do many things well, but applying explicitly learned rules to truly novel situations is simply beyond its capabilities.

One final benefit of System 2's ability to reason by way of the sequential manipulation of symbols and images is that it enables us to more easily consider hypothetical scenarios and their likely consequences. As we explored above, System 1's pattern recognition feature seems to rapidly assimilate even ambiguous stimuli to the closest known pattern.[152] Thus, once we have seen Lincoln's face in the squares, it is hard to see other potential patterns that they might make. This difficulty illustrates the general tendency of System 1 processing to default to a single model of reality.[153] But for many problems that we face—including many problems in the law—the ability to reason counterfactually is critical. This ability requires simultaneously knowing the true state of the world and imagining what would happen if the facts were otherwise. System 2 provides us with a means to achieve this by letting us consciously manipulate symbols and images in order to engage in hypothetical reasoning.[154] Using these explicit semantic representations, we can describe events

---

151. *See* Frankish, *supra* note 101, at 92–93.

152. *See* discussion *supra* Part II.B.

153. *See* Stanovich, *supra* note 28, at 67.

154. *See id.* at 21–22, 61–63.

that we have not observed or which we do not currently believe and then imagine the consequences that would follow if our imagined scenarios were true, without—as System 1 would—immediately treating our descriptions as the true state of affairs. Thus, this System 2 feature enables us to pose hypothetical or counterfactual questions and explore their answers.

D. The Complex Interplay of Dual-Process Reasoning

Although the manipulation of symbols and hypothetical scenarios is an important function of System 2, it is not the whole story. A significant volume of System 2 activity involves the exercise of self-control, in which we try to guide or restrain the results of automatic System 1 processing.[155] If I tell you to stare intently at a particular word on a page for a long stretch of time, you will find that your conscious mind is occupied in an effortful way, indicating that System 2 is involved in the task. But this sort of effortful focusing of attention surely does not involve symbol manipulation. Similarly, if you have committed yourself to losing weight, you may sometimes find yourself in a common conundrum, in which a tasty treat is available to you but your goal is to avoid eating it. The common element in both settings is that we must resist our automatic impulses, which prompt us to pay attention to distractions and to eat anything that we associate with pleasurable tastes. And because System 2 resources require the active use of conscious willpower, we will find these innate impulses harder to resist as we become distracted or tired.

This self-control function, combined with System 2's potential to use semantic reasoning to avoid associatively driven errors, is potentially very attractive as a means of improving the judgments and choices that would arise from an unmonitored System 1. Thus, for instance, there is a well-documented feature of Type 1 processing known as "belief bias," in which people who are trying to work out whether a logical argument is technically correct will often approve of faulty arguments if those arguments are used to support true conclusions.[156] It turns out, however, that the extent to which people fall prey to belief bias hinges on the degree to which System 2 processes are actively involved in solving the problem. Thus, when

---

155. *See* Masicampo & Baumeister, *supra* note 147, at 256. *See generally* Klaus Fiedler et al., *Exerting Control over Allegedly Automatic Associative Processes*, *in* PSYCHOLOGY OF SELF-REGULATION: COGNITIVE, AFFECTIVE, AND MOTIVATIONAL PROCESSES 249, 249–69 (Joseph P. Forgas et al., eds., 2009) (reviewing literature on executive overrides of automatic System 1 processes).

156. *See* Evans, *supra* note 99, at 264–65; *see also* DAN SIMON, IN DOUBT: THE PSYCHOLOGY OF THE CRIMINAL JUSTICE PROCESS 22–25 (2012) (describing this as "conformation bias").

research participants must allocate working memory to other tasks or answer too quickly to think through the problem sequentially, errors are more likely; conversely, when people are encouraged to reason deductively, such errors are reduced.[157]

This potential for these conscious *overrides* of System 1 responses suggests that a good way to avoid System 1 errors would be to employ System 2 reasoning as often as possible. This approach, however, is not the panacea it may seem. System 2 processing is not always engaged in order to monitor or override our intuitive impulses. It can also be engaged to defend or explain such instincts in ways that make them appear more reasonable than they really are.[158]

The role of System 2 as an unwitting apologist for the results of unconscious cognition has been well documented. One type of evidence comes from studies of "split-brain" patients, who underwent surgeries to sever the fibers that connected the two hemispheres of their brains (the corpus callosum) in order to treat chronic seizures.[159] The split-brain patients allowed researchers an unusual opportunity because the left hemispheres of their brains, which typically manage language functions, had no direct inputs from the right hemispheres, which lack speech-producing capabilities. In studies of such patients, researchers have observed a surprising phenomenon: the left, speaking side of the brain will invent reasons for actions initiated by the right, silent side, even though the former is ignorant of the stimuli that motivated the latter's action.[160] Even more strikingly, patients who engage in such confabulation do not seem to be aware that their fabricated reasons were not the real reason for their actions.[161]

---

157. *See* Evans, *supra* note 99, at 264–65 (reviewing literature).

158. *See* Jonathan Haidt, *The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment*, 108 Psychol. Rev. 814, 818 (2001). *See generally* Richard E. Nisbett & Timothy DeCamp Wilson, *Telling More Than We Can Know: Verbal Reports on Mental Processes*, 84 Pscyhol. Rev. 231 (1977) (collecting many examples of unconsciously confabulated explanations of associatively determined behaviors).

159. Wilson, *supra* note 28, at 95.

160. *Id.* at 96–97.

161. One such patient, P.S., was shown two images simultaneously, one in each eye, and then asked to choose a card with a related image. The left side of his brain was shown a picture of a chicken claw, while the right side was shown a snow scene. When offered the chance to choose a card with his right hand, he chose a chicken card, which matched what his left brain had seen; when choosing with his left hand, he chose a shovel, matching the snowy scene that was shown to his right brain. He found it easy to explain both decisions: He chose the chicken because it matched the claw, and he chose the shovel because "you have to clean out the chicken shed with a shovel." *Id.* at 96. But his explanation, of course, had nothing to do with his choice of the shovel, because his left hand was guided by his right brain, which had been shown a snowy

This sort of confabulation is not confined to split-brain patients, however. One striking example of System 2's ability to generate false reasons for unconsciously induced choices was seen in an experiment by Richard E. Nisbett and Timothy DeCamp Wilson, who arranged four identical pairs of nylon stockings on a table and then asked passersby to give them comparative ratings as part of a fictitious marketing survey. Despite the fact that the stockings were identical, viewers expressed a clear statistical tendency to prefer the rightmost item (chosen by 40%) and to disfavor the leftmost item (12%).[162] The exact reasons for this "right-side bias" are somewhat mysterious,[163] but the main point is that none of the participants believed that the location of the stockings had anything to do with their preferences, and all but one participant (who happened to be a psychology student) denied that any such influence was possible when asked directly. Rather, they invented fictitious differences between the stockings in terms of "knit, weave, sheerness, elasticity or workmanship" to justify their right-skewing preferences.[164]

One final example may serve to drive the point home as starkly as possible. Thalia Wheatley and Jonathan Haidt used hypnosis to prime participants to feel disgust when reading certain words and then had the participants evaluate several stories.[165] One of these stories involved no plausible moral violation: it described a student council leader who was "in charge of scheduling discussions about academic issues" and who tried to choose topics that "appeal to both professors and students in order to stimulate discussion."[166] Some of the hypnotically disgusted participants nevertheless rated the student council member's actions as morally wrong, even though no unprimed participant agreed.[167] Significantly, some hypnotized participants invented strained reasons that they claimed explained their peculiar

---

scene rather than anything to do with chickens. Despite this disconnect, P.S. "seemed perfectly comfortable with his answer and had no idea that it was a confabulation." *Id.*

162. Timothy de Camp Wilson & Richard E. Nisbett, *The Accuracy of Verbal Reports of Stimuli on Evaluations and Behavior*, 41 Soc. Psychol. 118, 123 (1978).

163. *Cf.* Daniel Cassanto, *Embodiment of Abstract Concepts: Good and Bad in Right- and Left-Handers*, 138 J. Experimental Psychol.: Gen. 351, 365 (suggesting that right- and left-handed people tend to associate their dominant side with positive affect ideas and their non-dominant side with negative affect ideas).

164. Wilson & Nisbett, *supra* note 162, at 124.

165. Thalia Wheatley & Jonathan Haidt, *Hypnotic Disgust Makes Moral Judgment More Severe*, 16 Psychol. Sci. 780, 780–83 (2005).

166. *Id.* at 782.

167. *Id.* at 782–83.

moral judgments: the council leader, they said, was "up to something" or was a "popularity-seeking snob."[168]

To sum up, our perceptions, decisions, and judgments are the product of two different, sometimes interacting, modes of cognition. System 1 processes information swiftly, effortlessly, and automatically; it has an almost magical ability to detect and recognize patterns. But the same associative structure that allows System 1 to detect patterns amidst the noise can fail us when it has been trained on a misleading set of examples or when some of its associative levers have been primed into activity by irrelevant stimuli. System 2, by contrast, works slowly, requires mental effort, and is capable of abstract and hypothetical analysis. With these capabilities, System 2 plays a dual role, sometimes overriding System 1's mistakes and at other times working to justify our associative intuitions.

III. A Dual-Process Account of Fact-Finding

Now that we have explored dual-process models of ordinary cognition, it is time to consider how these processes affect legal fact-finding. In this Part, I will attempt to describe the ways in which our two distinct styles of information processing interact when judges and juries decide cases.

My account places System 1 at center stage: fact finders arrive with preexisting networks of associations between various potential facts and concepts of guilt and criminality. Then, fact finders supplement their networks with new associations based on the case's facts and evidence. During the trial itself, System 2 is primarily engaged in the task of maintaining focus on the evidence items because of the largely passive nature of the jury and judicial instructions encouraging them to refrain from decision until the parties have finished presenting their cases.[169] By the end of a trial,

---

168. *Id.* at 783.

169. *See* Ronald J. Allen, *Rationality, Algorithms and Juridical Proof: A Preliminary Inquiry*, 1 Int'l J. Evid. & Proof 254, 261–62 (1996). (suggesting that juries have preconceived notions about facts and laws, which may or may not be accurate, and cannot appropriately evaluate information as they receive it). System 2 is unlikely to significantly constrain System 1's associative activity, except through attentional focus processes—that is, by trying to resist the pull of these distracting thoughts and to maintain focus on the witnesses' testimony. In fact, excessive explicit processing would tend to distract the jury from listening to the evidence, based on the well-documented phenomena of inattentional blindness, in which focused mental work generally limits our ability to attend to environmental stimuli. *See, e.g.*, Christopher F. Chabris et al., *You Do Not Talk About Fight Club If You Do Not Notice Fight Club: Inattentional Blindness for a Simulated Real-World Assault*, 2 Perception 150 (2011) (showing a majority of research participants who were placed in a cognitive-load condition failed to observe a violent

most fact finders will have developed intuitive feelings about the right answer in the case, which will be a product of an unconscious associative model. At the final stages of a case, however, System 2 processes will also play an important role. By fitting the pieces of evidence into stories, analyzing the fit between factual judgments and verdict categories, and applying semantically structured burdens of proof, judges and juries may use their deliberative faculties either to protect their intuitions or to override them. Figure 4 may help to clarify some of these interactions.
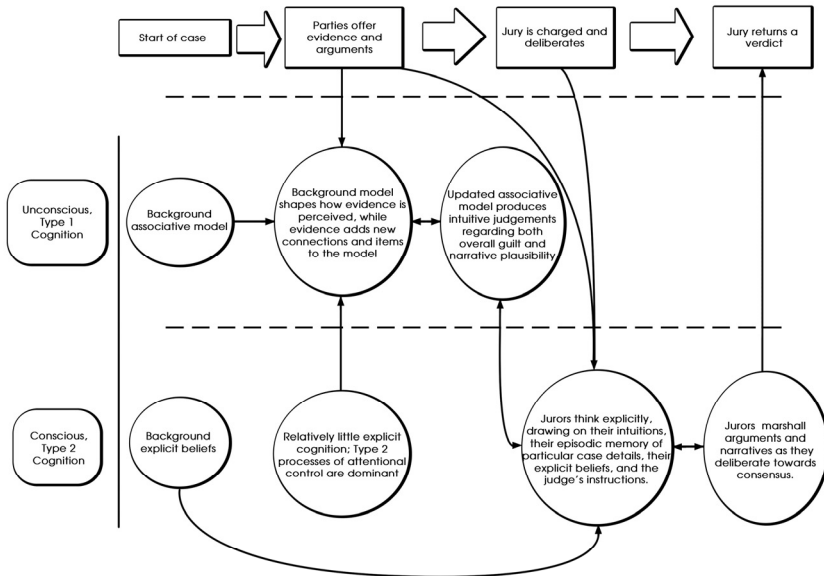

Figure 4: Chart depicting the interactions between System 1 and System 2 cognition processes during a trial.

To make this framework as clear as possible, I will walk through the ways that it might be applied to understand the factual inferences that allowed the jury to reach a guilty verdict in Juan Rivera's case. This exercise is, by necessity, speculative in nature; I do not have access to the private thoughts of the jurors in that case,[170] and even the jurors themselves would be unable to report the unconscious

---

physical confrontation in their environment). An implication of this is that, when one needs to absorb and retain large quantities of information, explicit mental processing may undermine—rather than facilitate—learning.

170. None of the jurors were interviewed in connection with any media reports following the verdict or the appeal.

aspects of their fact-finding process. My analysis, moreover, will necessarily simplify the messy details of a month-long trial. But these limits should not worry us excessively because the goal of this discussion is not to implement a fully specified formal model nor to perfectly describe the reality of fact-finding in one case.[171] Rather, my goal is to provide a descriptive account of the role that the two systems of cognition might play in a real-life case, and we can make progress toward this goal without a perfectly specified model or a godlike window into the minds of individual jurors. As this discussion will illustrate, a dual-process approach to modeling inference, unlike standard formulations of the Bayesian or story-comparison approaches, can explain why the pattern of evidence in cases like *Rivera* are likely to induce even well-meaning jurors to make seemingly indefensible verdict choices.

A.   System 1 Cognition at Trial: The Background Associative Model

To visualize the role of associative networks in fact-finding, it will help to first establish a useful metaphor: the interactive connectionist network. Previously described as an alternative to Bayesianism models, connectionist networks are systems comprising many simple units, which are connected by edges, or *connections*, into a larger network.[172] Each unit might be connected with many others, and the connections may vary in terms of their *weight*. Over time, the units can take on various *activation values*, indicating that the unit is currently in active use. When one unit is active, the connections between it and other units will impact the activation values of the connected units, varying in effect based on the weight of the connection.[173] If the weight is positive, the connection will have an excitatory effect, such that one active unit will amplify the value of the unit to which it is linked. If the weight is negative, that unit will have an inhibitory effect and act to lower the activation value of the connected unit instead of amplifying it.[174] These weighted connections

---

171. *See* discussion *supra* Part II.

172. Bechtel & Abrahamsen, *supra* note 88, at 23. I am not the first researcher to use connectionist models to describe fact-finding cognition. *Cf.* Thagard's *Causal Inference*, *supra* note 88, at 232 (exploring judicial reasoning in a trial according to computational model using both explanatory coherence and Bayesian networks); Simon, *supra* note 5, at 520–23 (evaluating a connectionist model to understand coherence shifts in mock juror cognition). To my knowledge, however, I am the first to suggest combining a connectionist account of System 1 activity with a nonconnectionist, motivation-driven account of System 2 activity. For a more general discussion of the associative cognitive process that informs the connectionist model see *supra* Part II.B.

173. Bechtel & Abrahamsen, *supra* note 88, at 24.

174. *Id.* at 24–25.

allow the network to express, in an intuitive way, the idea that we have learned to expect to encounter some concepts in combination (like thunder and lightning), while other concepts tend to imply each other's absence (like rain and sunshine).

Finally, such networks may also evolve and learn through the implementation of rules that allow the weights of connections to change in response to shifting patterns of unit activations over time.[175] Donald Hebb developed a learning rule that can help a connectionist system learn to more accurately model the associations that are present in its environment. In his simple model, when two units are simultaneously activated, the weight of the connection between them will be strengthened; when one is active and the other suppressed, the weight will become weakened or even inhibitory.[176] Over time, this means that units that "fire together, wire together."[177]

Once a model has been constructed and trained on a set of inputs, it can simulate certain kinds of cognition. Figure 5 provides a visual illustration of such a very simple connectionist network, which James McClelland designed to encode the demographic characteristics of the members of the two gangs from *West Side Story*, the "Jets" and the "Sharks."[178]

---

175. *See generally id.* at 66–105 (discussing learning strategies that have been developed for connectionist networks).

176. *Id.* at 72. *See generally* Peter Dayan & L.F. Abbott, Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems 281–93 (2001) (providing background information about the Hebbian Rule and learning).

177. Bernard J. Baars & Nicole M. Gage, Cognition, Brain, and Consciousness: Introduction to Cognitive Neuroscience 83 (2d ed. 2010).

178. James L. McClelland, *Retrieving General and Specific Information from Stored Knowledge of Specifics*, Proc. Third Ann. Conf. Cognitive Sci. Soc'y 170, 171 (1981). Astute readers will notice that the features ascribed to members of the Jets and the Sharks bear virtually no relation to any characters in the actual musical.
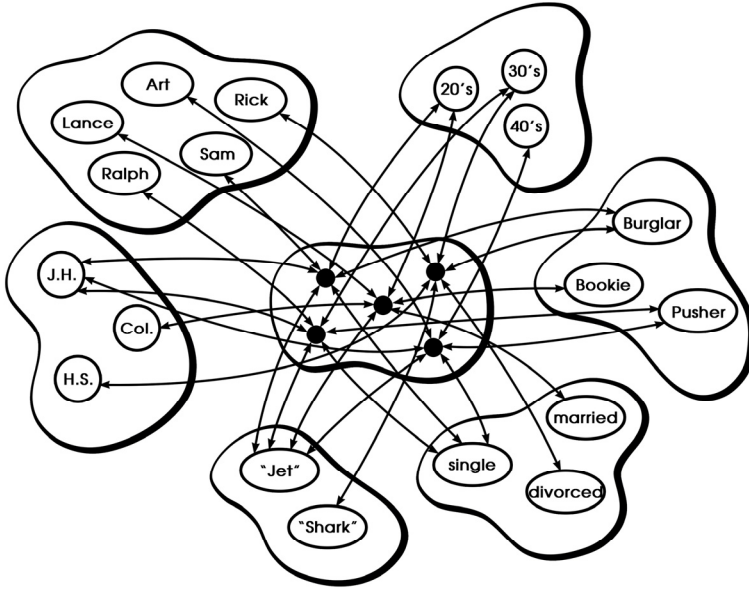
Figure 5: A visual depiction of a connectionist model, illustrating the characteristics of five hypothetical gang members.[179]

With the help of Figure 5, the nature of connectionist models should become clearer. If one activates a unit representing one of the individuals' names, the connection flowing away from that unit will activate a central unit, which will in turn activate other units that encode various features of that individual.[180] Thus, activating the name "Lance" will lead to activation of his central conceptual representation unit,[181] which will in turn activate the units encoding for "20s," "Bookie," "married," and so on. These other features, in turn, will propagate activation further. Thus, by thinking of Lance, one activates the "Jets" unit, which in turn leads to the weaker activation of other Jets. This neatly captures the idea that, by

---

179. From James L. McClelland, *Retrieving General and Specific Information from Stored Knowledge of Specifics*, Proc. Third Ann. Conf. Cognitive Sci. Soc'y 170, 171 fig.1 (1981). Reprinted with permission from James L. McClelland.

180. *See generally* Bechtel & Abrahamsen, *supra* note 88, at 21–47 (describing the features and function of the Jets v. Sharks network).

181. This central unit is a necessary feature of the network because no single feature, including the name "Lance," could uniquely describe the individual in question. Rather, "Lance" is coded as a connected network of related concepts, including his name, age, gang membership, etc., and the central unit is what links all those concepts together.

thinking of Lance, you are primed to think of other members of his gang.[182]

These connectionist networks can be implemented as formal models, but they can also be quite useful as visualization tools, helping us imagine the nature of networked associative processing, even as they necessarily simplify away some of the rich details of true human cognition.[183] These models usefully illustrate the ways that associative activation propagates through the unconscious mind. The pattern-recognition power of the associative machine, its tendency to exhibit strange priming effects, and its resistance to updating based on certain kinds of inputs are all relatively easy to describe in this framework, regardless of whether we could crack open an individual's skull and point to physical structures that map neatly onto the descriptions below.

With the analytic framework introduced, let us consider the nature of unconscious factual cognition at trial. The first point that must be emphasized is that judges and juries do not begin a trial as a *tabula rasa*, but rather start the trial with preexisting associative networks based on their past experiences.[184] This means that, before we can start to describe how a juror might process the *Rivera* evidence, we should first consider the existing set of patterns that she will be predisposed to see in any pattern of evidence.

Thus, let us imagine our hypothetical juror before she has heard any evidence or arguments in the *Rivera* case. Before the lawyers and witnesses have begun to speak, our juror will have already heard about many criminal cases, whether on the news or in fictional accounts, some of which may involve similar elements to the case she is about to hear. From those cases, she may have unconsciously

---

182. *See* discussion *supra* Part II.

183. *See* Bechtel & Abrahamsen, *supra* note 88, at 101–03 (suggesting that connectionist models can integrate associative cognition with rule-based cognition); Sloman, *supra* note 117, at 6–8, 19–20 (suggesting the use of connectionist models as a means of capturing associative aspects of cognition); Thagard, *supra* note 129, at 438–39 (using a connectionist network to model features of the Necker cube illusion). Such networks radically simplify some features of the biological brain, which may represent concepts in distributed form rather than as isolated and identifiable units within a larger network. *Cf.* Dayan & Abbott, *supra* note 176, at 230 (suggesting firing-rate models to simulate neural networks); 2 James L. McClelland & David E. Rumelhart, Parallel Distributed Processing: Psychological and Biological Models, 164–66 (1986) (discussing the benefits of connection information distribution); Thagard, *supra* note 129, at 456–57 (distinguishing connectionist models using distributed representations from those using localist representations). This simplification is a feature, not a bug, when we are trying to find ways to visualize the complex and alien terrain of unconscious cognition.

184. *See* Simon, *supra* note 5, at 536.

learned that certain facts tend to occur in conjunction with other facts. She starts the trial, in other words, with preexisting associative schemas connecting the concepts of "rape," "murder," and "guilt," with other events or ideas. Figure 6 provides a way to visualize the operation of a small subset of these possible associations, dealing with concepts that associate closely with rape.[185]
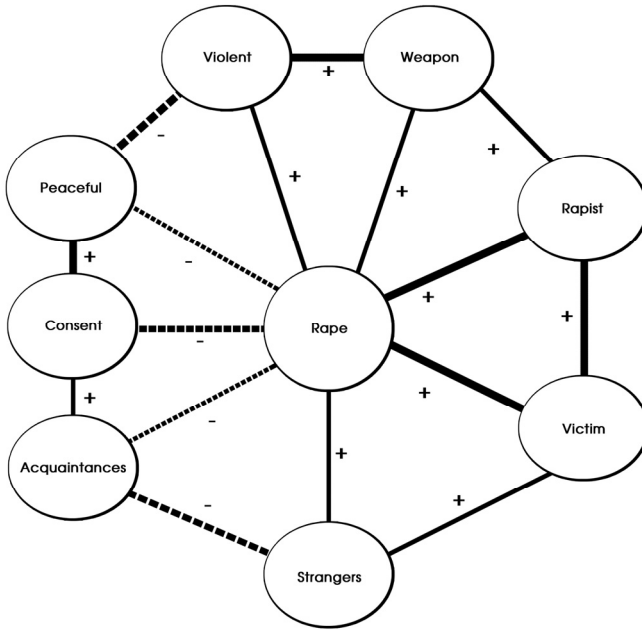

FIGURE 6: A simple network model illustrating a juror's possible associations with rape.

This simple network model is designed to illustrate a few ideas. First, note that some of the lines between units are solid, which indicates excitatory associative activation, while other lines are dotted, suggesting an inhibitory connection. These lines have varying weights, indicating that some of these relations are more powerful than others. This is designed to reflect the idea that activating some concepts in the associative unconscious may activate some related concepts while inhibiting other, dissonant concepts. For example, because it is hard to simultaneously intuit that conduct is peaceful and violent, the activation of one concept suppresses the other. Other aspects of the network may facilitate the intuitive recognition of guilt or innocence simultaneously, creating internal tension in the network.

---

185. *See* Henry F. Fradella & Kegan Brown, *The Effects of Using Social Scientific Rape Typologies on Juror Decisions to Convict*, 31 LAW & PSYCHOL. REV. 1, 4–6 (2007) (reviewing factors that may influence juror decisions in rape trials).

Thus, if a case involved both the use of a weapon and an alleged assault by an acquaintance, activation would be propagated simultaneously to the violent unit and the peaceful unit, which would then actively work to suppress each other, as well as the other units that are closely related. In such scenarios, the two families of related units would act competitively, and the model might well exhibit unstable cycles of fluctuating activations until one set of units eventually outcompetes the other and the network settles into a coherent and stable state.[186]

Figure 6 also illustrates that our preexisting associative schema may sometimes diverge from what is normatively ideal. Although some of the connections built into this model are clearly sound—such as the inhibitory connection between rape and consent—other connections do not correspond to real-world rape-case patterns. For example, this model incorporates a widely held, but false, assumption that rapes are correlated more with encountering strangers than known associates.[187] In fact, the reverse is true: the Department of Justice estimates that sixty-two percent of rapes are committed by "intimates" who have a prior sexual relationship with the victim.[188] But if the juror is operating with the set of associations described above, perhaps because that is the pattern of rape cases that the juror has encountered in the news and on TV, then that juror will tend to see rape more easily in cases involving strangers rather than intimates.

B. The Impact of Party Appearance and Identification on the Background Associative Model

Having visualized some of the details of a basic, preexisting associative model, let us now add the information that becomes available to jurors at the very start of a case. Before any evidence has been presented, jurors will be placed in the same room with the parties; there, the jurors will automatically begin integrating the parties into their associative matrices without realizing that they are "thinking" about anything at all. Figure 7 illustrates some of the associations that the *Rivera* jurors might have formed after sitting

---

186. *See* Sloman, *supra* note 117, at 8 (referring to this state as a point of "*minimum energy* or of *maximum harmony* or *coherence*" which provides the desired inference) (citations omitted).

187. *See* Fradella & Brown, *supra* note 185, at 4 (noting that stereotypical beliefs among the public about rape, rape victims, and those accused of rape are often flawed).

188. PATRICIA TJADEN & NANCY THEONNES, NAT'L INST. JUSTICE, FULL REPORT OF THE PREVALENCE, INCIDENCE, AND CONSEQUENCES OF VIOLENCE AGAINST WOMEN: FINDINGS FROM THE NATIONAL VIOLENCE AGAINST WOMEN SURVEY 44 (2000).

through jury selection and receiving initial instructions from the judge but before the lawyers started their opening statements.
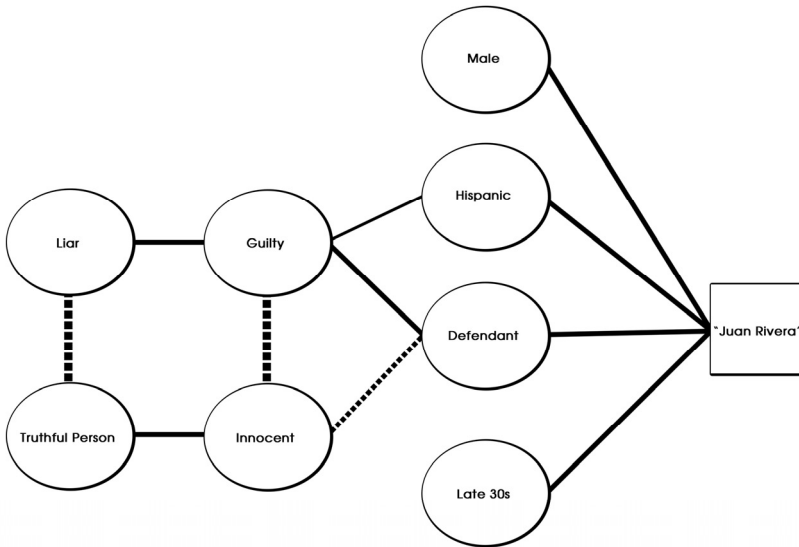

FIGURE 7: A network of pretrial associations a juror may have had with the defendant, Juan Rivera.

At this stage, all our juror knows is Juan Rivera's name, what he looks like, and his status as an accused defendant in a rape-murder case. Our juror may also know the official rule of our criminal justice system, which is that Rivera should be presumed innocent until he is proven guilty beyond a reasonable doubt. But our simple model suggests that, at least at the level of unconscious associative processing, the jurors likely associated Rivera with guilt rather than innocence. First, news reports, water-cooler gossip, and crime dramas on TV are all likely to facilitate a strong associative link between "defendant" and "guilty." This association is especially likely given the rarity with which innocent defendants are encountered in daily life. Second, the defendant's ethnicity may also be used, more weakly, to associatively prime the concept of guilt.[189] Finally, it is quite

189. *Cf.* Justin D. Levinson et al., *Guilty by Implicit Racial Bias: The Guilty/Not Guilty Implicit Association Test*, 8 OHIO ST. J. CRIM. LAW 187, 189–90, 206 (2010) ("[S]tudy participants held strong associations between Black and Guilty, relative to White and Guilty, and these implicit associations predicted the way mock jurors evaluated ambiguous evidence."); Russell K. Robinson, *Perceptual Segregation*, 108 COLUM. L. REV. 1093, 1170–71 (2008) (noting the problem of implicit preference for white over black, except by African American participants). I am not aware of research demonstrating a Hispanic-Guilty association as carefully as has been done in studies of implicit associations regarding African-Americans, but it seems reasonable to

plausible to expect that our juror will associate lying with guilt, given that telling lies is itself a norm violation and (sometimes) a crime as well. Because our juror will be weakly primed by Rivera's name, role, and appearance to intuit his guilt, our juror will also be primed to disbelieve what Rivera will say at trial. Thus, even before the lawyers begin their opening statements, and despite the nominal presumption of innocence, we can expect that the jurors will be developing an unconscious model of the case that connects his identity with guilt and deception.

C.   *System 1 Cognition: Adding Case-Specific Units to the Associative Network in Response to Trial Evidence*

Next, let us consider how a juror's associative model is altered by the parties' presentations of evidence. In *Rivera*, the prosecution spent a substantial portion of its case-in-chief asking multiple witnesses to describe in great detail the scene of the murder. After both police witnesses and neighbors testified to the sequence of events that culminated in the discovery of Holly's body, the jury saw numerous photographs of her corpse and the gruesome scene,[190] as well as a video walk-through of the apartment taken by police while the body was still present.[191] The prosecution also offered detailed autopsy testimony, including photos of each stab wound that the killer had inflicted on Holly, as well as close-up details of injuries to her genitals and anus from the rape.[192] Formally, all of this evidence was only weakly relevant to the case. Both the prosecution and the defense agreed that Holly had been brutally raped and murdered. And although the prosecution wished to suggest that the dirty scene or careless autopsy practices might have led to contamination of the forensic DNA evidence, that contention was rejected by all the expert witnesses who testified on that subject.[193] Indeed, perhaps for this reason, the defense did not even bother to cross-examine many of these initial witnesses.

---

assume that similar associations exist. And even if there were no direct associations of that type, we would expect spreading activation to follow a Hispanic-Minority-Black-Guilty path, so that even in the best-case scenario, Rivera's ethnicity would have weakly primed the idea of guilt.

190. *See, e.g.*, Trial Tr., *supra* note 15, at 013868 (describing the admittance of a crime scene photo of Holly's nude and bloody corpse to the jury); *id.* at 013908–21 (describing the admission of many additional photos, including shots of blood-covered toys and the victim's bloody handprints on the wall).

191. *Id.* at 013897–901.

192. *Id.* at 015760–812.

193. People v. Rivera, 962 N.E.2d 53, 58–59, 62–63 (Ill. App. Ct. 2011).

Although this res gestae evidence might seem inconsequential if viewed from a detached, analytical standpoint, it may have been quite powerful in its ability to structure the associative networks that the jurors were unconsciously constructing. In terms of the testimony itself, the repeated references and images of Holly's body and its condition would create associations only between that girl and concepts like violence, blood, and death. But the evidence in the case was not the only sensory input that was training this fledgling associative network; rather, the jury was also subject to the ongoing stimulus of the defendant's presence in the room. Figure 8 illustrates these associations.
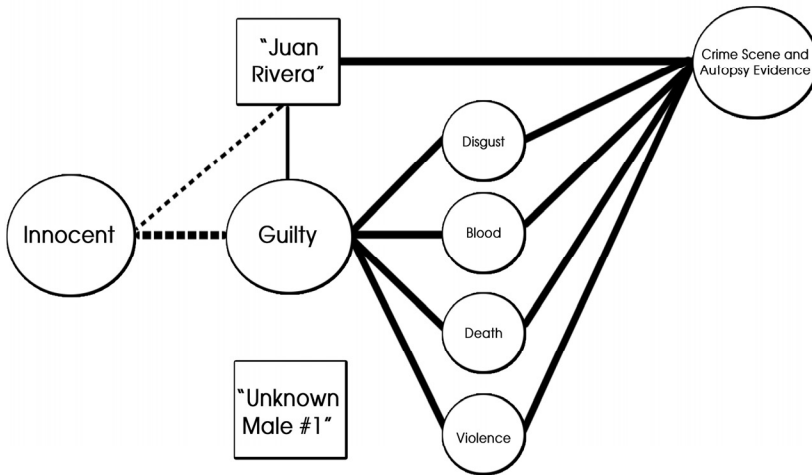

FIGURE 8: An associative network that jurors may have been constructing during the trial of Juan Rivera.

Thus, with an associative network already somewhat primed to associate Rivera, a non-Caucasian defendant, with guilt generally and therefore with rape and murder more specifically, the jurors started the case by encountering a steady stream of images of violent death with Rivera in the room as a continuing co-stimulus. And although there were many other individuals in the room, it is quite likely that the jurors gave special focus to Rivera, in part because it would be all too human to ask, under such circumstances, whether the person in front of you could have done such awful things. Thus, by having the jurors share a room with Rivera and pictures of gruesome death for several days, the prosecution was likely able to induce a reasonably strong unconscious association between Rivera, the scene of the crime, and the violent nature of Holly's death. As a result, the jurors would have had a natural tendency to interpret subsequent ambiguous

evidence in a manner that was consistent with these associations.[194] Here we see a key way that the dual-process model differs from prior approaches: evidence that no one would think of as relevant to the defendant's claimed alibi defense may, at the unconscious level, be priming jurors to view that defense with skepticism when it is raised later in the trial.

After exposing the jury to extensive images and testimony about the crime itself, the prosecution next called multiple witnesses who recounted various instances where Rivera lied to officers during the early phases of the murder investigation. According to several witnesses, Rivera had, at various times, claimed he attended, on the night of the crime, a party that never occurred;[195] alleged another individual was guilty of the crime, even though that individual was in fact not involved;[196] and asked an acquaintance to help him concoct an alibi for the night of the crime.[197] This phase of the prosecution's case had one obvious, evidentially proper purpose: to suggest that because Rivera was telling lies relating to the crime, he may have been conscious of his own guilt and trying to deflect investigators' interest onto other targets.

This evidence, however, would also have a more nuanced impact on the jury's fact-finding in a second, more subtle way. Connecting Rivera with lying may have significantly influenced the jurors' unconscious associative models of the case, triggering other associations that weakened Rivera's subsequent positions in nonobvious ways. Thus, because we generally associate lying with bad, immoral, guilty behavior, the jury was likely further primed to view Rivera as a person likely to engage in other kinds of bad

---

194. *See* Simon, *supra* note 5, at 537–38 (describing the impact of adverse information about a defendant on a coherence network); *cf.* Old Chief v. United States, 519 U.S. 172, 180–81 (1997) (discussing the prejudicial effect of evidence of a defendant's past violent crime); Risinger, *supra* note 25, at 442–46 (suggesting that prosecutors use heartstrings-and-gore evidence in part because seeing the "horror of the crime" may make jurors less willing to bear the risk of a wrongful acquittal).

195. Rivera had claimed that he went to a party at the home of a friend, Shanita Craig, located near the scene of the crime. *Rivera*, 962 N.E.2d at 56. In fact, there was no party at Craig's house that night. *Id.*

196. Rivera had reported to investigators that he saw a Hispanic or African American individual leave Craig's party for an hour and then return to the party in a disheveled state. The man, according to Rivera, had been breathless and sweaty when he returned, with disheveled clothing and a bloody scratch on his face. *See, e.g.*, Trial Tr., *supra* note 15, at 013967–68. Rivera later admitted that this, like his attendance at the party, was a fabrication. *Id.* at 014212.

197. According to a man named Michael Jackson, Rivera had made this request while the two were in the Lake County jail. *Rivera*, 962 N.E.2d at 58.

behavior, including the rape and murder at issue.[198] But this strong link between Rivera and lying would also, quite likely, help prime the jurors to view other people connected with Rivera as liars themselves. Thus, by strongly painting Rivera as a liar, the prosecution also effectively planted the suggestion—in the jury's unconscious associative network—that his counsel, his family, and other witnesses on his behalf might also be lying. Even though he never took the stand, Rivera's lack of credibility may have spread like a virus to inspire instinctive distrust of his whole "team."

Now we arrive at the heart of the prosecution's case: Rivera's various confessions. This was, in one sense, the strongest evidence of guilt in the case. Rivera had told multiple people, including police and (allegedly) various other jail inmates, that he was responsible for Holly's death.[199] He even signed two written statements that purported to recount his confessions to police officers.[200] For most people, it is common sense to think that an innocent person would not willingly admit to committing a crime unless the confession was coerced.[201] But despite the apparent power of this evidence, it also presented serious difficulties for the prosecution. Rivera's confessions to the police initially involved multiple incorrect details that the killer would seemingly have gotten right.[202] It was only after further interrogation that a second statement, which resolved many of the inaccuracies, was produced, raising the possibility that the police were feeding Rivera details of the crime scene to make his account more believable.[203] Rivera, moreover, behaved so erratically during the short window between his two confessions that he was confined to a padded

---

198. *Cf.* Ill. R. Evid. 404(a) (purporting to bar these sorts of inferences); Ill. R. Evid. 404(b) (permitting evidence of prior acts to be introduced for non-propensity purposes).

199. *Rivera*, 962 N.E.2d at 57–58.

200. *Id.*

201. *But cf.* Garrett, *supra* note 20, at 8 (concluding that over half of the forty falsely confessing exonerees who succumbed to police pressure were juvenile or mentally disabled); Saul M. Kassin, *Inside Interrogation: Why Innocent People Confess*, 32 Am. J. Trial Advoc. 525 (2009) (discussing the "phenomenon of false confession" and the influence of interrogation).

202. Rivera, for instance, stated initially that Holly had changed into a nightgown before their encounter, *Rivera*, 962 N.E.2d at 57, but this was doubly improbable. Her corpse was wearing a T-shirt, not a nightgown, when it was found, and the home in which she was murdered was not her own, so there would have been no reason for her to have a nightgown there in any event. *See* Trial Tr., *supra* note 15, at 013868, 014773.

203. *See Rivera*, 962 N.E.2d at 66–67 (concluding that the police had provided evidence to Rivera during the investigative process).

cell for his own protection, placed on suicide watch by a jailhouse nurse, and prescribed antipsychotic medications by a physician.[204] And to the extent that the prosecution claimed he had confessed in less pressurized settings, the prosecution was forced to rely on the testimony of jailhouse inmates who could be impeached with both their prior criminal records and their interest in receiving reduced sentences for coming forward.[205]

Despite these inconsistencies, the prosecution used the confession evidence in a way that would have maximized its associative impact. A central feature of the prosecution's strategy was the relentless repetition of details from the confessions. Rather than present each confession once or twice, the jury instead heard each confession from multiple witnesses' perspectives, some of whom repeated the confession details several times over. For instance, the first witness to describe the initial confession, Detective Donald Meadie, started by relating the confession as Rivera told it to him, to lay the foundation for admitting the written summary Rivera signed. Before concluding, Detective Meadie then read that summary aloud to the jury, effectively reporting the same confession twice.[206] Another witness, Sergeant Lou Tessman, went through the same lengthy and repetitive process with the subsequent confession.[207] Three more officers were then called to recount what they had heard and seen during Rivera's two confessions.[208] Finally, the prosecution layered on the far less detailed confessions related by the two jailhouse informant witnesses.[209] Thus, the jury was taken through various versions of the same story, in various forms and levels of detail, at least nine distinct times.

With this manner of presenting the evidence, the common elements among the initial and "corrected" confessions would have forged strong associative links through sheer repetition.[210] Thus, Rivera would have become very strongly linked with the jurors' unconscious representations of Holly, the murder weapon, the stabbings, and the scene of the crime. Thinking of any one of these

---

204. *E.g.*, Trial Tr., *supra* note 15, at 017138, 017263–64, 017306–08.

205. *Rivera*, 962 N.E.2d at 58, 64–65.

206. Trial Tr., *supra* note 15, at 014611–61.

207. *Id.* at 014822–93.

208. *Id.* at 015446–015518, 015851–66, 015917–67.

209. *Rivera*, 962 N.E.2d at 58, 64–65.

210. *See, e.g.*, Bechara et al., *supra* note 130, at 1293–94 (demonstrating implicit learning of complex conditional information in a noisy environment based on frequent repetition of stimuli); Destrebecqz & Cleeremans, *supra* note 130, at 347–49 (same); *cf.* Simon, *supra* note 156, at 136 (noting that jurors will often use the "richness" of a confession narrative as a proxy for its believability).

things would cause a juror to readily and unconsciously bring Rivera's image to mind. In contrast, other details of the confessions conflicted and would not have formed equally strong associations. For example, Rivera reported staging a break-in in his second statement but not in his first one; thus, the jurors would be much less likely to automatically associate Rivera with the damage to the door. But through punctuated repetition of links between Rivera, the crime scene, and Holly, the prosecution made it easy for jurors to imagine scenarios of Rivera holding the murder weapon and stabbing Holly with it, while making it harder to imagine scenarios where those elements did not occur in combination.

A very different story follows for the defense's evidence. Although it was quite strong analytically, the defense's case was almost certainly much less powerful than the prosecution's case in terms of its ability to forge strong associative links between Rivera and innocence or between the crime and another perpetrator.[211]

First, the defense tried to paint a contrasting portrait of Rivera for the jury, with testimony that portrayed him as a troubled and frightened boy rather than a violent, strategic man. As framed by an expert psychiatric witness, Rivera's depression might have induced him both to behave recklessly, because he did not value himself very much, and to try very hard to please others, because of his low self-esteem.[212] Coupled with an IQ that bordered mental retardation[213] and a lengthy, stressful police interrogation that left him sleep-deprived and nearly psychotic,[214] the defense tried to suggest that the confession was a desperate attempt to please the investigators by telling them whatever they wanted to hear.

Although this evidence may have provided good reasons to doubt the prosecution's theory of the case, it was less powerful in its ability to build a strong set of competing associative connections suggesting innocence. Some of the new suggested links might well have helped Rivera: connecting him with youth and family,[215] or with mental

---

211. To be clear, I do not intend to impugn the quality of the defense team, who operated in a professional and impressive manner throughout the trial. Rather, the difficulties arose due to the nature of the defense's evidence, the wide latitude that the prosecution had been given early on, and the inability to introduce expert testimony that might have weakened the associative force of the confession evidence.

212. Trial Tr., *supra* note 15, at 017552–53.

213. *Id.* at 017396.

214. *Id.* at 017410–23, 017456, 017554.

215. The defense called Rivera's parents to provide him with an alibi on the night of the murder, and this allowed some discussion of his home life with them. *See id.* at 017693–017710 (questioning Rivera's father about his childhood). The defense also admitted into evidence a picture of Rivera as a youth. *Id.* at 017700.

retardation,[216] might have served him well because those concepts may be associatively linked with innocence. But associating Rivera with mental illness or psychosis could have undermined those intuitive links because jurors might well connect mental illness and recklessness with guilt and antisocial behavior rather than with innocence. Additionally, the mere fact that Rivera was required by standard practice to present this evidence *after* the jurors had developed an associative network that favored the prosecution's case meant that the defense had an uphill battle to win. Any time that an ambiguous fact could be associated with guilt rather than innocence, the structure of the preexisting network would make the guilt connections easier to form than the innocent ones.[217]

Second, the key element of the defense's case, the exclusionary DNA profile, was likewise limited in its ability to speak in terms that System 1 could process. The defense offered evidence, from two experts, that the semen that had been deposited in Holly's vagina shortly before her death bore a single, unique DNA profile that was not attributable to Juan Rivera but rather to an individual the defense labeled as "Unknown Male #1."[218] At the level of System 1, this evidence built a competing associative link between the commission of the crime and this new figure, Unknown Male #1, and degraded the existing link with Rivera that had been formed by the repetitive confession testimony. But whereas Rivera's side of the

---

216. *Id.* at 017396.

217. Dan Simon, Chadwick J. Snow, and Stephen J. Read observed a pattern like this in a study they conducted to investigate cognitive consistency theories. They first exposed participants to a substantial portion of evidence in a hypothetical case, while instructing them that more would be coming later on. They then obtained from their participants an indication of which way the participants were "leaning," either toward a guilty or not-guilty verdict. Finally, they exposed the participants to new evidence that was either strongly inculpatory (placing the defendant near the scene of the crime) or strongly exculpatory (establishing a solid alibi). Although this was not the focus of the study, it is noteworthy that two-thirds of the participants who learned information strongly inconsistent with their initial leanings ended up rendering a verdict decision that aligned with their initial leanings rather than with the strong new evidence. Strikingly, this occurred even though the initial field of evidence was ambiguous enough to induce a balanced number of participants to "lean guilty" and "lean not guilty," suggesting that even prior associations that are relatively weak can resist displacement by strong new evidence. Dan Simon, Chadwick J. Snow & Stephen J. Read, *The Redux of Cognitive Consistency Theories: Evidence Judgments by Constraint Satisfaction*, 86 J. PERSONALITY & SOC. PSYCHOL. 814, 824–27 (2004). I am grateful to Dan Simon for pointing out this interpretation of their data in correspondence.

218. People v. Rivera, 962 N.E.2d 53, 58–59 (Ill. App. Ct. 2011).

network was richly detailed and thus mutually reinforcing,[219] this competing family of units would necessarily be weak because only minimal detail could be attached to the unknown contributor of the semen. Moreover, the strength of the competing connections would not be determined merely by the rational probative force of the new evidence but also by the frequency and conviction with which key witnesses suggested the new associations. On these metrics, the defense operated at an inherent disadvantage. Its experts were few in number, and its arguments hinged on one single, powerful observation rather than a multitude of less probative connections.

Furthermore, these new connections between the crime and Unknown Male #1 were likely weakened significantly by the prosecutor's repeated attempts to elicit from defense witnesses some acknowledgment that the exculpatory results could be attributed to contamination.[220] If conceived as an attempt to elicit evidence favorable to the state's theory of the case, this strategy was a failure because no expert was willing to agree that contamination was a likely explanation of the forensic results.[221] But merely by asking the question again and again, the prosecutors built one more connection into this new family of units—an unconscious associative link between the DNA testimony and the concept of contamination. Contamination associates to various negative concepts, such as unreliability, and invokes a generally negative emotional effect in connection with the DNA evidence. As a result, our jurors would likely have been primed to view the DNA evidence as less reliable than it otherwise would be and to find thinking about the evidence as slightly more aversive and unpleasant than they otherwise would. Thus, despite the judge's repeated instructions to the contrary,[222] the arguments and questions of counsel would have been integrated into the jurors' unconscious associative models just as readily as actual items of evidence in the case.

At the close of the evidence, the jurors would likely have constructed a rich associative model connecting the various items of evidence, individuals, and events in the case, based in large part upon the frequency with which those elements were encountered in combination. Because of this internal tension in the unconscious network, it would have been possible for jurors, depending on which elements of the case they were focusing on at any one time, to find it intuitive to link either Rivera or Unknown Male #1 with Holly's death. But for the reasons described above, the family of units that tended to activate "Rivera" and "guilt" simultaneously would tend to

---

219. *See supra* Figure 8.

220. *Rivera*, 962 N.E.2d at 59.

221. *Id.*

222. *See, e.g.*, Trial Tr., *supra* note 15, at 018191–93.

dominate the network most of the time. As a result, despite the serious weaknesses in the prosecution's case, which are plain from a reflective distance, a juror who sat through the *Rivera* trial could easily have found a verdict of guilty more intuitively plausible than a verdict of acquittal.

D. The Role of System 2 Reasoning in Shaping the Rivera Verdict

In the account I just outlined, a judge or juror listening to a case can be thought of as building an unconscious associative network, which links the people, events, and concepts involved in the case in a network of excitatory and inhibitory relationships. This associative model, once formed, functions by outputting intuitive feelings about the facts of the case. But this System 1 cognitive activity is far from the whole picture of how a fact finder will reach a verdict. In addition to relying on their intuition, fact finders also employ the conscious and effortful thinking styles that characterize System 2 cognition. Depending on their motivations, fact finders might employ this type of thinking either to advance their intuitive preferences or to resist them.[223] In this section, I explore some of the roles that this form of cognition might have played as the jury deliberated the right result in Rivera's case, and, in particular, how System 2 may have been employed to override any remaining doubts as the jury deliberated in Rivera's case.

We should begin by recognizing that some cases may require little System 2 involvement by jurors. If, after hearing the evidence, the right answer seems obvious and the law does not require a reasoned explanation for the decision, a juror might vote for her strong intuitive preference without engaging in much explicit thought or discussion. We know, however, that the *Rivera* jury members did not find the case easy or obvious: they deliberated for several days and sent multiple notes to the judge during deliberations indicating that they were finding it difficult to reach agreement.[224] Thus, many of the jurors were likely experiencing either an internal associative conflict or a conflict with the intuitions of other jurors, and would have needed to deploy System 2 resources in order to establish both internal and external consensus about the verdict.

System 2 cognition may take several forms during the fact-finding process. First, a fact finder will often find it useful to construct a narrative to assist in deciding a case. Such narratives will organize the facts of the case into "episode structures" that focus on chains of cause and effect—particularly on the goals, choices, and actions of key individuals in the case—rather than on the simpler correlational information that forms the basis for the associative mechanisms

---

223. *See* discussion *supra* Part II.

224. *E.g.*, Trial Tr., *supra* note 15, at 018205, 018255, 018269–70, 018301.

described earlier.[225] In some cases, these narratives might be articulated only in a fact finder's mind, as she strings together individual aspects of the case in her working memory. In other cases, the fact finder might communicate a narrative to others, either in oral form (such as during jury deliberations) or in writing (as in the case of a judicial opinion following a bench trial).[226] Given the secrecy of juror deliberations under modern law,[227] we will never know for sure whether the *Rivera* jurors found it useful to organize their thoughts about the case into narrative form, but, given the case's complexity, it seems quite likely.

Narrative explanations will generally integrate both aspects of the evidence actually heard during trial and "gap-filling" material that helps to complete a particular aspect of a case's story.[228] Such gaps might arise for many reasons. Sometimes the evidence may not have included information that is necessary to help explain or justify a decision. At other times, necessary information may have been presented but then forgotten in the deluge of details that so often accompany a complicated case. Thus, a story favoring the prosecution that attempted to be consistent with the DNA forensic testimony would require a juror to rely on assumptions regarding how, exactly, an eleven-year-old girl had managed to die with semen in her vagina that was not attributable to the person who raped and killed her. By contrast, a story favoring the defense would require gap-filling regarding why Rivera would have fabricated a fake alibi if he had a real one available. As with any difficult case, neither side's narrative was free of gaps or inconsistencies.

When fact finders engage in gap-filling, we can expect cooperative involvement by both cognitive systems. System 2 constructs an explicit story based on the fact finder's memory of the evidence, while System 1 fills the gaps by supplying intuitions about what the "missing pieces" are likely to be. Similarly, there may be a need to address conflicting testimony when constructing explanatory stories. Such choices will often be made based on an intuitive sense of which

---

225. *See* Pennington & Hastie, *supra* note 67, at 248–51 (describing the episode structure of cause and effect); *cf.* Sloman, *supra* note 117, at 4 (describing an "associative system" of reasoning for comparison with a "rules based" system).

226. *See* Fed. R. Civ. P. 52(a)(1) (requiring findings of fact to be set forth in the record when a judge rules following a bench trial); Fed. R. Crim. P. 23(c) (requiring findings of fact to be set forth in the record following a bench trial if a party requests the judge to do so before the verdict is rendered).

227. *See* Ill. R. Evid. 606(b) (prohibiting courts from requiring jurors to testify regarding jury deliberations).

228. *See* Pennington & Hastie, *supra* note 67, at 249 (noting that jurors fill in gaps in testimony to construct complete episodes).

witnesses are believable or which narrative chains are more likely to occur in the real world. Once again, we might expect a fact finder to resolve such disputes in part by choosing the version that intuitively feels right.[229]

Saying that System 1 may help to shape the narrative that a fact finder will find acceptable is not to say that it will so dominate the process that the construction of a narrative will never causally affect the outcome. A different jury hearing the *Rivera* case might also have felt an associatively derived intuition that Rivera "felt" like a guilty person. But that jury might have found that the hypothesis of guilt was impossible to square with any articulable story, perhaps due to the difficulty of constructing a plausible narrative of guilt that explains the DNA evidence. In such a circumstance, the hypothetical jurors might have the persistent intuition that the defendant is somehow "getting away with something" yet still decide to vote for acquittal.

In addition to constructing solitary narratives, some fact finders might take this process a step further by engaging in an explicit process of comparison between narratives favoring the prosecution and defense. Thus, in a case like Rivera's, a juror might feel that neither the State's nor the defense's theory of the case can be rejected outright because both are supported by some persuasive evidence. In such cases, one sensible way to proceed is to engage in abductive reasoning—imagining alternatively that each party's story is actually true and considering how well those stories would explain the given pattern of evidence.[230] Thus, Professors Allen and Pardo's explanatory conception of the proof process, which was offered as an epistemic and normative model for what proper evidential reasoning should look like,[231] will in some cases form an important part of a juror's reasoning process, as a descriptive matter.

---

229. *See* Pennington & Hastie, *supra* note 5, at 190–91 (noting that judgments of story coherence will depend on whether the story is consistent, plausible, and complete).

230. *See* Cheng, *supra* note 3, at 1266–68 (proposing the use of the preponderance standard as a ratio test that compares the probability of the narratives offered by the plaintiff and defendant); Pardo, *supra* note 39, at 596–97 (discussing an alternative conception that focuses on potential explanations of the evidence).

231. Although Pardo and Allen sometimes labeled their model as "descriptive," *e.g.*, Pardo & Allen, *supra* note 3, at 226, they do not use that term in the way that I am employing it in this Article. Whereas I conceive of descriptive models as attempting to describe real-world reasoning practices, they initially developed their model as an attempt to "expla[in] the legal proof process and its features," *id.* at 268, in a primarily *epistemic* sense. *Compare id.*, *with* discussion *supra* Part I. Thus, as Pardo clarifies in later work, he does not believe that the explanatory theory of the *proof process* could be falsified by contrary empirical data regarding real-world fact-finding practices; rather, it

When placed within the context of a dual-process framework, abductive reasoning might work either to further associative intuitions or to restrain them, depending both on the cognitive complexity of the case and the extent to which the fact finder is motivated to either defend or resist her own intuitions. Like the simple story construction discussed above, the process of abduction will necessarily demand System 2 resources. The associative model, after all, is unitary and holistic, and it cannot be instantly restructured by a simple act of imagination.[232] Rather, our fact finder must employ mental effort to imagine a world in which the key facts alleged by one party are true and focus on the likely pattern of evidence she would expect to see in that world.[233]

Some elements of this process might involve further System 2 thinking in the form of explicit logical inferences. Often, however, the abductive process will also depend on associative intuitions. It seems likely that many jurors, if forced to choose between competing plausible explanations, would proceed by imagining alternatively that each story was true and then gauging which one feels more "right" at a gut level. If the process takes this intuitive form rather than a more analytical style, then the process of abduction will be substantially constrained by the preexisting structure of the fact finder's associative network. In such instances, the use of abduction will tend to act in a justificatory fashion rather than as a meaningful constraint on the answer produced by System 1. Therefore, in cases where jurors engage in an explicit comparison of stories, that process may have no causal impact on the result except when jurors are motivated to question and reconsider their intuitive judgments.

System 2 cognition may also give effect to legal rules that constrain fact finders to reach results that do not necessarily accord with their straightforward intuitions, such as burdens of proof. In civil

---

provides normative guidance when it conflicts with such practices. *See* Pardo, *supra* note 39, at 598–99. This means that their theory would be more clearly classified as normative—rather than descriptive or doctrinal—when situated in my own typology. Understood in this light, the discussion above is meant to explore the extent to which Pardo and Allen's normative conception of epistemically proper fact-finding processes are realized in practice, an inquiry which is related to, but distinct from, the central goals of their work. I am indebted to Mike Pardo for useful correspondence on this point.

232. *See* STANOVICH, *supra* note 28, at 66–67; Jonathan St. B. T. Evans et al., *A Theory of Hypothetical Thinking*, *in* THINKING: PSYCHOLOGICAL PERSPECTIVES ON REASONING, JUDGMENT, AND DECISION MAKING 3, 4–10 (David Hardman & Laura Maachi eds., 2003).

233. *Cf.* Pardo & Allen, *supra* note 3, at 234–35 (noting that fact finders decide based on the relative plausibility of the versions of events put forth by the parties and additional ones constructed by themselves or fellow jurors).

cases, which are typically resolved by a preponderance of the evidence, this will rarely have a significant impact on a juror's verdict choice because a juror whose associative intuition favors guilt will generally feel that guilt is more probable than not.[234] In criminal cases, the requirement of proof "beyond a reasonable doubt" will create a more complicated dialogue between the two systems of reasoning.[235] A juror who wishes to follow this instruction faithfully may be motivated to resist a weakly dominant associative impulse regarding a case's outcome.

Thus, even if Rivera felt more guilty than innocent to most of the jurors in his case, there were likely enough competing associations raised by the defense that the jurors would have initially felt some discomfort going along with their dominant intuition. By giving a reasonable doubt instruction, a judge effectively urges that if a juror has conflicting intuitions that are both strong enough to count as "reasonable," then the juror should choose innocence, even if she feels a stronger intuitive pull toward finding the defendant guilty.[236] In some cases, this instruction may be effective in initiating a System 2 override of the dominant System 1 impulse, but we face a competing possibility: jurors might try to resolve the conflict not by overriding their dominant intuitions but instead by crafting arguments that help to put their doubts to rest. If jurors were to spend a significant amount of time exploring reasons to discount the defense's theory of the case, they might, through a new process of repetition, strengthen their intuitions concerning guilt and weaken the associative connections with innocence.[237] With sufficient time and effort, the jurors might discover that they no longer felt any "reasonable doubts" after all.

This framework might well explain how the *Rivera* jury, after four days of deliberation, was able to arrive at a result that seems impossible to square with the reasonable doubt instruction. Much of the time may have been spent not in trying to craft reasons to resist their intuitions, but instead trying to minimize the reasons the defense had given for doubting that their intuitions were correct. For instance, with sufficient discussion of the idea that the forensic evidence was contaminated, the jurors might have slowly degraded the associations between the DNA evidence and Rivera's innocence,

---

234. *See* 1 Christopher B. Mueller & Laird C. Kirkpatrick, Federal Evidence § 3:5 (3d ed. 2007).

235. *Id.* at § 3:17.

236. *See* Pardo & Allen, *supra* note 3, at 238–39.

237. *See* Patricia G. Devine, *Stereotypes and Prejudice: Their Automatic and Controlled Components*, 56 J. Personality & Soc. Psychol. 5, 15–16 (1989) (discussing the use of repeated invocation of contrary belief structures to weaken a preexisting associative stereotype).

making it easier to silence their internal doubts regarding his guilt. By repeating a story to themselves with sufficient frequency, in other words, the jurors may have come to believe it despite its lack of foundation in evidence.[238]

When we pull these different strands of System 2 cognition together, the resulting picture may seem murky. System 2 can play a role in evidential inference through the construction of narratives or arguments, through the analysis of hypothetical scenarios framed around the parties' theories, through explicit burden-of-persuasion rules, or through the implementation of limiting instructions. In all of these different guises, however, there is a common theme. Two interacting forces will shape System 2's effects: (1) the case-specific intuitions that derive from the fact finder's associative model and (2) the motivational forces that encourage the fact finder to either follow those intuitions wherever they might lead or, alternatively, to resist the intuitions using one of the above-described strategies. Globally, this interaction means that the standard approaches to modeling fact-finding, which focus on the probabilistic or plausible force of evidence to the exclusion of motivational influences, will often fail to explain the results that jurors are likely to reach on a given set of evidence. It also implies that we will not be able to dissolve the link between flawed intuitions and erroneous outcomes simply by encouraging more reflection, as some have argued.[239] The *Rivera* jury, after all, reflected for four days before returning a deeply flawed verdict. Rather, we have two options. We can try to shape the fact-finding process so that problematic intuitions are less likely to arise in the first place. But if that fails, we must do more than just induce fact finders to reflect before deciding—we must motivate them to view their own intuitions from a detached distance.

IV. Professor Cohen's Gatecrashers, or: How I Learned to Stop Worrying and Love Statistical Evidence

In Part III, I offered a general account of the roles that both System 1 and System 2 cognitive processes play in the realm of legal fact-finding, using the case of *People v. Rivera* as an extended example. Examining the case in these terms, I believe, helps to explain the seemingly inexplicable willingness of twelve jurors to find a man guilty beyond a reasonable doubt of rape and murder when exculpatory DNA evidence clearly implicated an unknown third party instead. This initial use of the dual-process model focused on its unique *descriptive* power in explaining the ways in which jury

---

238. *See id.*

239. *See* Guthrie et al., *supra* note 27, at 29–40.

intuitions might diverge from normative rationality, and the ways in which they might use their deliberative faculties to further those intuitions rather than resist them. Although the primary goal of the dual-process model is descriptive, achieving better understandings of the sources of our intuitive judgments about evidence has important normative benefits as well. Going forward, I hope to show that thinking of fact-finding in dual-process terms can also be helpful when we address some of the common puzzles that have frustrated evidence theorists when they have attempted to formulate models of normatively justifiable trial results. In this final section, I will focus on one such conundrum—the longstanding debate over the proper probative force of naked statistical evidence.[240] The dual-process model of fact-finding does not directly tell us whether such evidence should be trusted, but it does suggest that the widespread intuitive distrust of verdicts that rest on purely statistical inferences about guilt or liability may arise from the inability of such evidence to speak in terms that our intuition can understand. As a result, theorists may find themselves seeking to justify normatively suboptimal trial outcomes out of a desire to defend intuitions that are not themselves particularly trustworthy.

Recall Cohen's paradox of the gatecrasher.[241] A thousand people went to the rodeo, but only 499 of them bought tickets. We know, therefore, that 501 of them crashed the gate. Unfortunately, we assume in this hypothetical that this is the limit of our knowledge; no further evidence exists to discriminate between paying customers and gatecrashers.[242] As Cohen argues, "the balance of probability" would support finding any randomly chosen attendee guilty if the rodeo organizer sued to collect the ticket price, because there is a 0.501 chance that any randomly chosen patron was ticketless that night. But, he maintains, this result would be an "absurd injustice," with the result that no judge or jury would ever convict a person under such circumstances.[243] Indeed, some courts have disallowed parties

---

240. By "naked" statistical evidence, I am referring to evidence that arrives in a plainly statistical form; that is, it involves transparent and quantifiable uncertainty. Although all sources of evidence do involve some inherent uncertainties, and although neither explicitly statistical or facially non-statistical evidence can be said to be superior as an *a priori* matter, Pardo, *supra* note 39, at 573, the admissibility and probative force of naked statistics in trials has been a subject of lengthy and heated debate in the evidence literature. *See generally* Amit Pundik, *What is Wrong with Statistical Evidence? The Attempts to Establish an Epistemic Deficiency*, 27 CIV. JUST. Q. 461 (reviewing literature).

241. COHEN, *supra* note 36, at 75.

242. *Id.*

243. *Id.*; *see also* Charles R. Nesson, *Reasonable Doubt and Permissive Inferences: The Value of Complexity*, 92 HARV. L. REV. 1187, 1192–94 (1979) (developing the similarly structured "Prison Yard" hypothetical);

from using purely statistical evidence when parties seek to prove whether a specific event in the past has actually occurred.[244]

A number of scholars have attempted to justify this sort of intuition through analytic argument. Some have focused on the idea that we can rationally draw some sort of "negative inference" from the fact that more evidence was not offered, effectively concluding that the plaintiff is hiding evidence.[245] This seems, in the end, to dodge the hypothetical, which is designed to force us to decide what a fact finder should do when statistical evidence is the only source of information.[246] In that world, the rodeo organizer is not spoliating any evidence and could not have found more even if motivated to do so, so it would be senseless to punish the plaintiff for the first failure and pointless to incentivize him or others to try harder in future, similar cases.

Other theorists, including Jonathan Cohen, Alex Stein, and, more recently, Edward Cheng, relying on the idea that purely statistical proof lacks any individualized force against a particular defendant, argue that the balance in such cases is purely even.[247] As Cheng puts it, we would encounter the same evidence whether the defendant was innocent or guilty, so we have an equal probability of observing it either way.[248] But this argument suffers from two flaws. First, is it really the case that we would be equally likely to observe the same ratios either way? Perhaps not. The situation is analogous to our position after being given a coin with an unknown bias, flipping it once, and getting heads. It seems more likely, in such a circumstance, that the coin is biased toward heads than tails, even though either is within the realm of possibility. Similarly, it seems more likely that we would observe a 501/499 ratio of crashers to payees in the case of a particular defendant's liability than innocence, given that we should

---

Tribe, *supra* note 36, at 1340–41 (developing the similar "Blue Bus" hypothetical).

244. *See, e.g.*, Sargent v. Mass. Accident Co., 29 N.E.2d 825, 827 (Mass. 1940) (finding that proving a proposition is mathematically favorable is not enough to prove preponderance of the evidence); People v. Risley, 108 N.E. 200, 203 (N.Y. 1915) (holding the lower court erred in allowing testimony regarding probability).

245. *See, e.g.*, Kaye, *supra* note 2, at 106; Richard A. Posner, *An Economic Approach to the Law of Evidence*, 51 Stan L. Rev. 1477, 1508–10 (1999); Tribe, *supra* note 36, at 1349.

246. *See* Ronald J. Allen, *A Reconceptualization of Civil Trials*, 66 B.U. L. Rev. 401, 411–12.

247. *See* Cohen, *supra* note 36, at 271; Alex Stein, Foundations of Evidence Law 82–83 (2005); Cheng, *supra* note 3, at 1269–71; *cf.* Allen, *supra* note 246, at 414 (critiquing this line of analysis).

248. Cheng, *supra* note 3, at 1270.

expect more guilty "flips" in a world biased toward gatecrashers than in a world biased toward payees.

The second flaw in the "no individualized force" argument is that, even if we thought that the likelihood of observing this evidence was equivalent either way, this still does not mean that the information is useless. As Professors Cheng and Cohen both acknowledge, it is more likely than not that even a randomly chosen defendant did, in fact, crash the gate.[249] In that sense, we do know something about the defendant as an individual, although it happens to be the case that we know the same thing about all the other rodeo attendees. The fact that this logic, if applied *in the aggregate*, will lead to far too much compensation for the rodeo is immaterial; after all, if we applied Cohen's preferred rule of no-liability the outcome would be even more unjust in the opposite direction.[250]

My point in referencing this ongoing discussion is not, however, to present a knock-down argument as to whether we should or should not find a defendant guilty in the scenario Professor Cohen envisions (or others like it). Instead, I suggest that thinking of this and other similar examples in dual-process terms may reframe the debate and reveal the potential dangers of relying on intuition as a guide toward optimal decision making. Imagine, for a moment, a judge or juror who heard a case that is as simple and short as Cohen described, with the entire case amounting to little more than a recitation of the facts of the number of tickets sold and the number of rodeo attendees, plus some proof that the defendant was an attendee. Such testimony, given its relatively abstract nature, would support only a minimal associative model, as illustrated in Figure 9.

---

249. Cohen, *supra* note 36, at 75; Cheng, *supra* note 3, at 1270.

250. *See* Allen, *supra* note 246, at 413 (finding that Professor Cohen's rule would result in a situation where "denying recovery results in a larger windfall, overall, to the defendants").
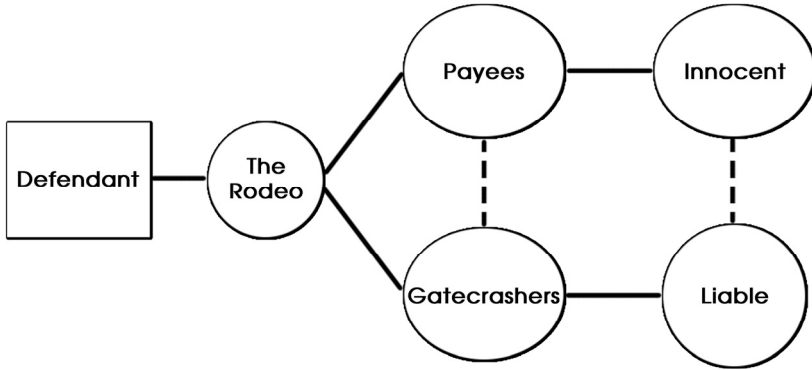
Figure 9: An associative model of the gatecrashers scenario with neither sub-network dominating.

In a model like this, neither the liable-gatecrasher nor the innocent-payee sub-networks would tend to dominate, because neither the one nor the other occupied a greater length of time at trial or was mentioned more frequently in connection with the defendant. Rather, they both are connected to the defendant with approximately equal associative weight. As a result, a fact finder trying to decide what to do in such a case would likely find that their intuitions are silent. Either finding the defendant liable or not liable feels unsupported, so that the jurors do not feel like they know what happened to any level of certainty.

The only answer that can be given in such a scenario must arise from System 2 cognition. Perhaps, like Cohen, a judge or juror would likely be inclined to preserve the status quo in the absence of a felt conviction that the defendant did something wrong. Or perhaps, as some courts have suggested, it is proper to defer to a mathematical calculation indicating guilt or liability so long as that calculation is rational and supported by a proper foundation in evidence, even if it feels intuitively unsupported.[251] So, unfortunately, just initiating System 2 processing will not determine a unique outcome.

But one thing is clearer now than it might have been before: we might not want to give any special deference to our intuitive judgments in a case like this. The example, by construction, is novel and foreign to our prior experience, so our past associations are unlikely to provide any useful guide to resolving the case. The fact that our associative network would not produce a dominant and coherent result may not reveal any deep truths about the proper resolution of a case like this. Instead, it might merely demonstrate

---

251. *See, e.g.*, People v. Collins, 438 P.2d 33, 38–39 (Cal. 1968).

System 1's inability to compute appropriate answers to problems that are framed probabilistically.[252]

To see why this is the case, imagine an individual, Isabelle, who rarely flies and rarely reads about safe air travel but watches a steady diet of nightly news. Now imagine further that the local nightly news show always covers airplane crashes, wherever they occur in the world, because they make for dramatic viewing. What result is likely to arise? Well, Isabelle will have relatively few associations between air travel and safety (because she does not fly very often), but she will have developed relatively strong associations between air travel and danger because she often thinks about air travel in the context of the disaster reports. The plausible result is that she feels tense and nervous when she thinks about air travel and easily imagines scenes of disaster—despite the fact that air travel is statistically quite safe and that those reported disasters compose a vanishingly small fraction of airplane trips. And because the unconscious system is trained by patterns of co-occurring stimuli rather than explicit semantic instructions, this could be the case even if Isabelle has been informed that air travel is safer than driving, as a statistical matter.

Isabelle's case illustrates a more general truth about the limits of the intuitive judgments that our dual-process minds can generate. Isabelle's intuition was not a reliable guide to choosing a safe means of air travel because it was trained on a non-representative sample of co-occurrences between airplanes and accidents. This suggests a broader principle: intuition may be a good guide when it incorporates a fair quantity of representative and domain-specific experience, but in other cases, it may be useless or even actively wrong.

Applying this principle to Cohen and company's intuitions about the gatecrasher case, we should see cause for concern. Logically, an individual fact finder would have a similarly indeterminate set of intuitions about the outcome of a gatecrasher case—whether the payee/crasher ratio was 499/501, 400/600, or 600/400—because the numbers alone would not lead either party to elaborate its case with more detail, repetition, or vividness. But those differences, it seems, *should* matter, at least if we want to minimize the social costs of errors in the trial process. It is meaningfully different, from the perspective of social cost, whether we are erring in forty percent versus sixty percent of similarly situated cases, even if our associatively generated intuitions were the same across the board.[253]

---

252. *See generally* Kahneman, *supra* note 28, at 109–18 (cataloging the limited ability of System 1 to generate useful intuitions to statistically framed questions).

253. *See* Allen, *supra* note 246, at 413 (noting that awarding victory to the defendant in the rodeo case results in a suboptimal allocation of the risk of error).

As a result, it might be worthwhile to treat our intuitions about purely statistical proof with care and caution. The dual-process framework, as a descriptive matter, provides strong support for the idea that real-world fact finders will generally be reluctant to sanction a defendant based on statistical evidence alone because their associative intuitions will provide no encouragement toward such a result. But this descriptive account cannot teach us what the best normative approach toward handling such cases would be.[254] To decide what jurors *ought* to do, we would ideally decide to defer to some source of information that does incorporate relevant considerations of social costs, such as the rate of errors likely under different frameworks, the costs of those errors, and the impact of differing rules on the system's legitimacy or efficiency.[255]

Our intuitions about naked statistical proof are likely to be inapposite to such goals, unfortunately. First, there are not many cases in ordinary trial practice where naked statistics play a significant role. Second, our system of trial by a lay jury further reduces the extent to which such cases can be used to develop reliable intuitions. Third, and most importantly, even if jurors and judges were exposed to such cases with some frequency, they would not get reliable feedback about outcome accuracy, so there would be little means of training their intuitions to produce reliable outcomes.[256] In short, it seems that the gatecrasher case is unable to generate strong intuitions regarding the defendant's guilt or innocence, but it is doubtful that this intuitive void has any deep meaning in terms of the accuracy or fairness of the outcome. This suggests that the only "paradox" that the gatecrashers present is this: Why does it feel so wrong to do what is right, which is to ignore our intuitions, do the math, and require a defendant who probably jumped a turnstile to pay for his seat at the rodeo?

CONCLUSION

In this Article, I have tried to elaborate a dual-process account of fact-finding and to illustrate why such an approach is valuable. The dual-process view lays bare a hidden feature of fact-finding that has received little attention in most existing formal and informal models of the proof process—the significant role played by unconscious

---

254. *See* Simon, *supra* note 5, at 550–59 (urging that a coherence network model might function well descriptively but that its outputs were not necessarily a good guide to normatively correct inference).

255. *See* Pundik, *supra* note 240, at 463 (urging that we need a normative model, not a descriptive one, if we are to draw inferences about the desirability of using naked statistical evidence).

256. *See* Mark Spottswood, *Evidence-Based Litigation Reform*, 51 U. LOUISVILLE L. REV. 25, 27 (2012).

associative cognitive processes in constraining and generating verdict decisions. System 1 can make a wrong decision feel right, as seen with *Rivera*, and a potentially right decision feel wrong, as seen with naked statistical data. But System 1 processes fly under the radar because they are not subject to our conscious awareness or control. Their effects can be hard to predict because they operate according to a logic that sometimes defies common-sense ideas about fair and accurate decision making. In the universe of the unconscious associative system, two concepts can become connected based on their mere co-occurrence in our sensory environments, and those associations will cause us to expect to see those stimuli together in the future. This simple fact of human nature has radical implications for legal fact-finding: it means that jurors will draw inferences in ways that disregard the legal separation between what is and what is not "evidence" in the courtroom; treat the frequency of evidential combinations as significant, even if it bears no real relation to that evidence's probative force; and be more receptive to evidence and arguments when they go along with our (sometimes biased) preexisting, unconscious expectations.

These new ways of understanding fact-finding cognition are not mere academic curiosities, however. *Rivera* shows us that associative cognition can play a powerful role in shaping our decisions and can lead to results that we might not endorse from a reflective distance. In that case, the prosecutor took a case that seemed like a loser and presented it in a way that powerfully captured a jury's intuitions, arriving at a result that few observers would have predicted. Whether we wish to exploit this effect as lawyers, understand it as scholars, or prevent it from recurring as regulators of the trial environment, we all share an interest in understanding what kinds of evidence are likely to bring about unreliable or unfair intuitive judgments. To that end, I hope that considering the jury's view of the case from a dual-process point of view has been illuminating.